ISSN 2085-4552

ULTIMATICS

Jurnal Teknik Informatika

IMPLEMENTASI OF YOLOV8 IN OBJECT RECOGNITION SYSTEM FOR PUBLIC AREA SECURITY IN KEBUN RAYA
Prihandoko Prihandoko, Sri Agustina Rumapea, Muhamad Faishal Fawwaz (1-10)

OPTIMIZING CV MATCHING WITH JOB VACANCIES USING THE BOYER-MOORE ALGORITHM
Andrew Christofer Hadiwinata, Eunike Endariahna Surbakti (11-16)

SI-MAMAS: DESIGN AND DEVELOPMENT OF A MOBILE-BASED MOSQUE MANAGEMENT INFORMATION SYSTEM
Muhamad Alda (17-27)

USING CONVOLUTIONAL NEURAL NETWORK AND SALIENCY MAPS FOR CIREBON BATIK RECOGNITION
Yoga Aditiya, Marlinda Vasty Overbeek, Suwito Pomalingo (28-33)

APPLICATION OF FUZZY AHP-TOPSÄ°S HYBRID METHOD IN FACILITY LOCATION SELECTION FOR SOFTWARE SYSTEMS
Kamala Rafig Aliyeva (34-41)

DEVELOPMENT OF RESTFUL API MENTAL HEALTH APPLICATION WITH MICROSERVICES ARCHITECTURE USING GOOGLE CLOUD PLATFORM
Muhamad Zidan, Is Mardianto, Agus Salim (42-53)

ASSOCIATION RULE MINING OF CONSUMER BEHAVIOR AT MOY SUPERMARKET USING APRIORI ALGORITHM
Eka Putri Oktavia, Amalia Nur Alifah, Mediliadita Dwi Dimastia (54-62)

COMPARISON OF COSINE SIMILARITY, RABIN-KARP, AND LEVENSHTEIN DISTANCE ALGORITHMS FOR PLAGIARISM DETECTION IN DOCUMENT
Jasman Pardede, Agil Yudistira (63-71)

A COMPARATIVE STUDY: PREDICTING CUSTOMER CHURN IN BANKING USING LOGISTIC REGRESSION & RANDOM FOREST
Mhd Basri, Muhammad Iqbal Pradipta, Krisna Aditya (72-81)

ENHANCING INTELLIGENT TUTORING SYSTEMS THROUGH SVM-BASED ACADEMIC PERFORMANCE CLASSIFICATION AND RULE-BASED QUESTION RECOMMENDATION
Fenina Adline Twince Tobing, Toto Haryanto (82-89)

PREWEDDING LOCATION SELECTION RECOMMENDATION SYSTEM USING COUNT VECTORIZATION AND COSINE SIMILARITY
Dede Kurniadi, Ilham Ahmad Maulana (90-98)

CLUSTERING STUDENT COMPETENCIES USING THE K-MEANS ALGORITHM
Ratih Friska Dwi Andini, Febri Liantoni, Aris Budianto (99-106)

WEB-BASED ONLINE LEARNING PLATFORM WITH RAD AND LARAVEL METHODS
Mohammad Reza Fahlevi, Saeful Mu'minin (107-113)

CALCULUS CALCULUS SCORES AND SLEEP QUALITY: A STUDY OF UMN INFORMATICS 2022
Bonifacius Martin Wibawa, Revan Revian, William Lo Channiko, Fransiscus Ati Halim (114-122)

DEVELOPING AN ANDROID-BASED TOUR GUIDE APPLICATION
Imanuel Vicky Sanjaya, Restyandito Restyandito, Maria Nila Anggia Rini, Andhika Galuh Prabawati (123-134)

Engineering | Vol. 17 | No. 1 | Hal. 1-134 | June 2025 | ISSN 2085-4552

UMN

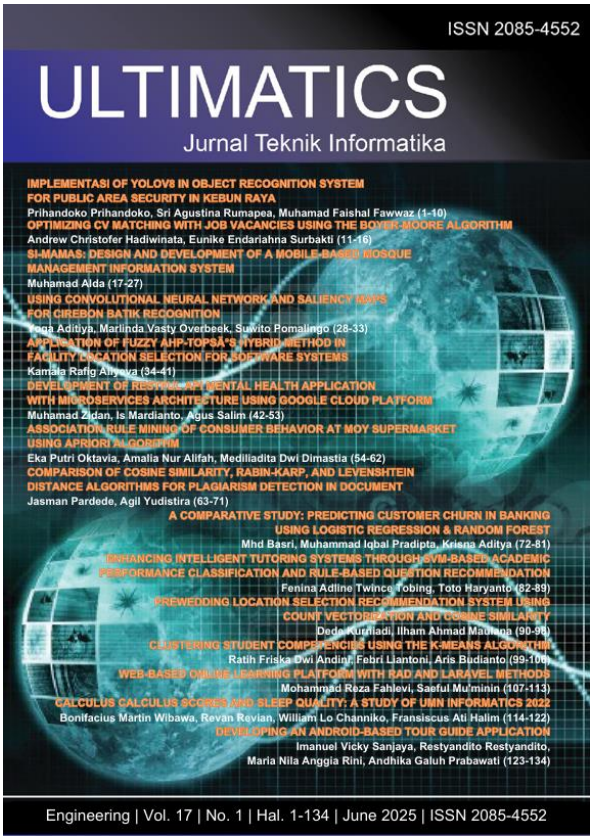
UNIVERSITAS
MULTIMEDIA
NUSANTARA

ISSN 2085-4552

EDITORIAL BOARD

EDITORIAL ADDRESS

Universitas Multimedia Nusantara (UMN)
Jl. Scientia Boulevard
Gading Serpong
Tangerang, Banten - 15811
Indonesia
Phone. (021) 5422 0808
Fax. (021) 5422 0800
Email: ultimatics@umn.ac.id



Ultimatics : Jurnal Teknik Informatika is the Journal of the Informatics Study Program at Universitas Multimedia Nusantara which presents scientific research articles in the fields of Computer Science and Informatics, as well as the latest theoretical and practical issues, including Analysis and Design of Algorithm, Software Engineering, System and Network security, Ubiquitous and Mobile Computing, Artificial Intelligence and Machine Learning, Algorithm Theory, World Wide Web, Cryptography, as well as other topics in the field of Informatics. Ultimatics: Jurnal Teknik Informatika is published regularly twice a year (June and December) and is managed by the Informatics Study Program at Universitas Multimedia Nusantara.

Call for Papers

Ultimatics: Jurnal Teknik Informatika is the Journal of the Informatics Study Program at Universitas Multimedia Nusantara which presents scientific research articles in the fields of Analysis and Design of Algorithm, Software Engineering, System and Network security, as well as the latest theoretical and practical issues, including Ubiquitous and Mobile Computing, Artificial Intelligence and Machine Learning, Algorithm Theory, World Wide Web, Cryptography, as well as other topics in the field of Informatics. ULTIMATICS is published twice a year by Faculty of Engineering and Informatics of Universitas Multimedia Nusantara in cooperation with UMN Press.

International Journal of New Media Technology (IJNMT) is a scholarly open access, peer-reviewed, and interdisciplinary journal focusing on theories, methods and implementations of new media technology. Topics include, but not limited to digital technology for creative industry, infrastructure technology, computing communication and networking, signal and image processing, intelligent system, control and embedded system, mobile and web-based system, and robotics.

Ultima Computing: Jurnal Sistem Komputer is a Journal of Computer Engineering Study Program, Universitas Multimedia Nusantara which presents scientific research articles in the field of Computer Engineering and Electrical Engineering as well as current theoretical and practical issues, including Edge Computing, Internet-of-Things, Embedded Systems, Robotics, Control System, Network and Communication, System Integration, as well as other topics in the field of Computer Engineering and Electrical Engineering.

Ultima InfoSys: Jurnal Ilmu Sistem Informasi is a Journal of Information Systems Study Program at Universitas Multimedia Nusantara which presents scientific research articles in the field of Information Systems, as well as the latest theoretical and practical issues, including database systems, management information systems, system analysis and development, system project management information, programming, mobile information system, and other topics related to Information Systems.

ISSN 2085-4552

FOREWORD

ULTIMA Greetings!

Ultimatics : Jurnal Teknik Informatika is the Journal of the Informatics Study Program at Universitas Multimedia Nusantara which presents scientific research articles in the fields of Computer Science and Informatics, as well as the latest theoretical and practical issues, including Analysis and Design of Algorithm, Software Engineering, System and Network Security, Ubiquitous and Mobile Computing, Artificial Intelligence and Machine Learning, Algorithm Theory, World Wide Web, Cryptography, as well as other topics in the field of Informatics. Ultimatics: Jurnal Teknik Informatika is published regularly twice a year (June and December) and is published by the Faculty of Engineering and Informatics at Universitas Multimedia Nusantara.

In this June 2025 edition, Ultimatics enters the 1st Edition of Volume 17. In this edition there are fifteen scientific papers from researchers, academics and practitioners in the fields of Computer Science and Informatics. Some of the topics raised in this journal are: Implementation of YOLOv8 in Object Recognition Systems for Public Area Security in Kebun Raya Bogor, Optimizing CV Matching with Job Vacancies Using the Boyer-Moore Algorithm, SI-MAMAS: Design and Development of a Mobile-Based Mosque Management Information System, Using Convolutional Neural Network and Saliency Maps for Cirebon Batik Recognition, Application of Fuzzy AHP-TOPSÄ°S Hybrid Method in Facility Location Selection for Software Systems, Development of Restful API Mental Health Application with Microservices Architecture Using Google Cloud Platform, Association Rule Mining of Consumer Behavior at MOY Supermarket Using Apriori Algorithm, Comparison of Cosine Similarity, Rabin-Karp, and Levenshtein Distance Algorithms for Plagiarism Detection in Document, A Comparative Study: Predicting Customer Churn in Banking Using Logistic Regression & Random Forest, Enhancing Intelligent Tutoring Systems through SVM-Based Academic Performance Classification and Rule-Based Question Recommendation, Prewedding Location Selection Recommendation System using Count Vectorization and Cosine Similarity, Clustering Student Competencies Using the K-Means Algorithm, Web-Based Online Learning Platform with RAD and Laravel Methods, Calculus Calculus Scores and Sleep Quality: A Study of UMN Informatics 2022, Developing an Android-based Tour Guide Application.

On this occasion we would also like to invite the participation of our dear readers, researchers, academics, and practitioners, in the field of Engineering and Informatics, to submit quality scientific papers to: Ultimatics : Jurnal Teknik Informatics, International Journal of New Media Technology (IJNMT), Ultima Infosys: Journal of Information Systems and Ultima Computing: Journal of Computer Systems. Information regarding writing guidelines and templates, as well as other related information can be obtained through the email address ultimatics@umn.ac.id and the webpage of our Journal here.

Finally, we would like to thank all contributors to this December 2025 Edition of Ultimatics. We hope that scientific articles from research in this journal can be useful and contribute to the development of research and science in Indonesia.

June 2025,

David Agustriawan, S.Kom., M.Sc., Ph.D.
Editor-in-Chief

ISSN 2085-4552

DAFTAR ISI

No	Title	Authors	Pages
1	Implementation of YOLOv8 in Object Recognition Systems for Public Area Security in Kebun Raya Bogor	Prihandoko Prihandoko, Sri Agustina Rumapea, Muhamad Faishal Fawwaz	1-10
2	Optimizing CV Matching with Job Vacancies Using the Boyer-Moore Algorithm	Andrew Christofer Hadiwinata, Eunike Endariahna Surbakti	11-16
3	SI-MAMAS: Design and Development of a Mobile-Based Mosque Management Information System	Muhamad Alda	17-27
4	Using Convolutional Neural Network and Saliency Maps for Cirebon Batik Recognition	Yoga Aditiya, Marlinda Vasty Overbeek, Suwito Pomalingo	28-33
5	Application of Fuzzy AHP-TOPSÄ°S Hybrid Method in Facility Location Selection for Software Systems	Kamala Rafig Aliyeva	34-41
6	Development of Restful API Mental Health Application with Microservices Architecture Using Google Cloud Platform	Muhamad Zidan, Is Mardianto, Agus Salim	42-53
7	Association Rule Mining of Consumer Behavior at MOY Supermarket Using Apriori Algorithm	Eka Putri Oktavia, Amalia Nur Alifah, Mediliadita Dwi Dimastia	54-62
8	Comparison of Cosine Similarity, Rabin-Karp, and Levenshtein Distance Algorithms for Plagiarism Detection in Document	Jasman Pardede, Agil Yudistira	63-71
9	A Comparative Study: Predicting Customer Churn in Banking Using Logistic Regression & Random Forest	Mhd Basri, Muhammad Iqbal Pradipta, Krisna Aditya	72-81
10	Enhancing Intelligent Tutoring Systems through SVM-Based Academic Performance Classification and Rule-Based Question Recommendation	Fenina Adline Twince Tobing, Toto Haryanto	82-89
11	Prewedding Location Selection Recommendation System using Count Vectorization and Cosine Similarity	Dede Kurniadi, Ilham Ahmad Maulana	90-98
12	Clustering Student Competencies Using the K-Means Algorithm	Ratih Friska Dwi Andini, Febri Liantoni, Aris Budianto	99-106
13	Web-Based Online Learning Platform with RAD and Laravel Methods	Mohammad Reza Fahlevi, Saeful Mu'minin	107-113
14	Calculus Calculus Scores and Sleep Quality: A Study of UMN Informatics 2022	Bonifacius Martin Wibawa, Revan Revian, William Lo Channiko, Fransiscus Ati Halim	114-122
15	Developing an Android-based Tour Guide Application	Imanuel Vicky Sanjaya, Restyandito Restyandito, Maria Nila Anggia Rini, Andhika Galuh Prabawati	123-134

ISSN 2085-4552

Implementation of YOLOv8 in Object Recognition Systems for Public Area Security in Kebun Raya Bogor

Prihandoko[1], Sri Agustina Rumapea[2], Muhamad Faishal Fawwaz[3]
[1,3]Departemen of Informatics, Gunadarma University, Depok, Indonesia
[2] Faculty of Computer Science, Universitas Methodist Indonesia, Medan, Indonesia
[1]pri@staff.gunadarma.ac.id, [2]sriagustina78@gmail.com, [3]muhamadfaishalfawwaz44@gmail.com

Accepted 29 April 2025
Approved 29 April 2025

***Abstract*— Pedestrian areas often serve as centers of high public activity, requiring intelligent monitoring systems to ensure user safety and comfort. Computer vision technology, especially object detection, provides a promising method for identifying and counting individuals in public spaces. This study implements YOLOv8 to develop a human detection and crowd counting model in the pedestrian zones of Bogor Botanical Garden. Researchers collected images and videos from three strategic locations and annotated them using Roboflow with a single class labeled "person." They trained the model on Google Colab using a Region of Interest (ROI)-based method and evaluated it through metrics like precision, recall, F1-score, confusion matrix, and mean Average Precision (mAP). The model achieved a precision of 0.846, recall of 0.858, F1-score of 0.85, and mAP@50 of 0.951, though mAP@50-95 dropped to 0.586. These findings show YOLOv8 provides strong real-time pedestrian detection, though enhancing precision in complex environments remains challenging.**

***Index Terms*— *crowd counting; deep learning; object detection; pedestrian surveillance; YOLOv8*.**

I. INTRODUCTION

Crowds in public spaces, including pedestrian areas, town squares, open markets, and sports venues, present significant challenges related to security and safety. High-density human gatherings not only impact comfort levels but also elevate the risk of incidents such as accidents and criminal activities. Poorly managed large crowds have occasionally led to tragic outcomes, exemplified by the crowd crush disaster in Itaewon, Seoul, in 2022, which resulted in over 150 fatalities [1]. With the continuous increase in public events and activities, effective crowd management and monitoring have become imperative. In response to these challenges, artificial intelligence (AI) technologies, particularly object detection and crowd counting methods, offer promising avenues to enhance safety and security.

Numerous prior studies have investigated various deep learning approaches, including CNN, SSD, YOLO, and R-CNN, for human detection and counting tasks. Among these methods, the YOLO (You Only Look Once) algorithm, specifically its latest iteration YOLOv8, has gained recognition due to its superior real-time detection capabilities and high accuracy under complex environmental conditions [2]. YOLOv8 significantly improves human behavior detection accuracy by approximately 4.2% compared to previous versions [3]. Furthermore, YOLOv8 achieved high precision (94.32%) and recall (91.17%) in complex scenarios such as passenger detection in elevators [4]. Despite these advancements, existing research identifies persistent limitations in accurately detecting individuals within densely crowded and complex environments. Architectural modifications to the YOLOv8 model are necessary to improve its mean Average Precision (mAP) [5]. Additionally, although YOLOv8 offers promising performance, it still struggles with distant or partially obstructed objects [6]. Therefore, there remains a critical need to further evaluate and refine the YOLOv8 algorithm under real-world, challenging conditions.

The YOLOv8 method is highly suitable for implementation in research related to object recognition systems for public area security at Kebun Raya Bogor, as this model achieves an optimal balance between detection speed and accuracy—both critical aspects for real-time scenarios. Compared to YOLOv9, which, despite offering slightly better accuracy, introduces higher complexity resulting in increased inference times and computational demands, YOLOv8 provides a more practical and resource-efficient solution for real-world deployments. Meanwhile, Faster R-CNN, although known for high detection accuracy, employs a two-stage detection process that significantly increases inference latency, making it less effective for real-time monitoring in public spaces. Thus, YOLOv8 emerges as the most appropriate choice in this study, effectively

addressing the need for rapid, efficient, and sufficiently accurate object detection in real-time public surveillance applications.

Addressing the existing challenges and building upon previous research findings, the present study aims to develop and evaluate a deep learning-based model employing the YOLOv8 algorithm specifically for the detection and counting of individuals within crowds at the pedestrian areas of Bogor Botanical Gardens. The primary objective of this study is to create a reliable and real-time monitoring system capable of enhancing public safety by promptly identifying human density levels.

To achieve this objective, the research employs an experimental approach involving the collection of image and video datasets from pedestrian areas within Bogor Botanical Gardens. The collected datasets will undergo preprocessing stages, including image resizing and annotation with bounding boxes, facilitated through the Roboflow platform. Subsequently, data will be partitioned into training, validation, and testing subsets to provide comprehensive model evaluation.

II. LITERATURE REVIEW

2.1 Pedestrian Zones and Urban Life

Pedestrian zones are a key part of modern urban planning. These spaces are designed to encourage people to walk, interact, and feel safe and comfortable in the city. When done right, pedestrian zones don't just improve the walkability of a place—they also reduce traffic congestion and lower air pollution. To make pedestrian zones work effectively, planners need to consider how people move through the area, how safe and accessible it is, and how well it connects with other parts of the city [7-9].

But as more people gather in public spaces—especially during events or busy times—it becomes more difficult to manage and monitor those areas. That's where technology, like crowd monitoring and computer vision, starts to play an important role.

2.2 Counting Crowds

Crowd counting is all about estimating how many people are in a space and understanding how they're distributed. It's used in many settings—like public safety, traffic control, and event planning. But this task is far from simple. When crowds are dense or people are partially blocked from view (what we call occlusion), it becomes much harder to accurately count them. The problem gets even trickier when people are moving or spread out unevenly [10].

Crowd counting can be approached in two main ways. The first is supervised learning, where the system is trained using data that's already labeled—so it knows, for example, what a "person" looks like. The second is unsupervised learning, where the system has to figure out for itself how to group and interpret the data without any prior labeling [11].

2.3 Computer Vision

Computer vision is a branch of artificial intelligence that teaches machines to "see" and make sense of visual data, just like humans do. Thanks to powerful models called convolutional neural networks (CNNs), machines can now detect patterns, shapes, and movements in images and videos with impressive accuracy. These systems can recognize what's in a scene, understand how things relate to each other, and even spot unusual activity [12-14].

For crowded areas, computer vision allows systems to automatically scan camera footage and detect how many people are present—without needing a human operator to watch every frame.

2.4 Object Detection

At the heart of many computer vision systems is object detection. This technology lets a computer recognize and locate multiple objects—like people, cars, or bicycles—in a single image. It does this by drawing bounding boxes around the objects and labeling them. There are two types of object detectors:

- Two-stage detectors, such as Faster R-CNN, first identify possible object locations, then analyze each one to decide what's inside.
- One-stage detectors, like YOLO and SSD, skip the proposal step and predict everything at once, making them faster and more efficient for real-time use.

To measure how well these models perform, we use metrics like precision (how many detections were correct), recall (how many real objects were found), and IoU (Intersection over Union)—a measure of how closely the predicted box matches the actual object.

2.5 YOLOv8

One of the most impressive object detection models today is YOLOv8. It's the latest version of the popular "You Only Look Once" family of models, and it's packed with upgrades. Unlike older versions that rely on predefined "anchor boxes," YOLOv8 is anchor-free, which makes it simpler and more flexible when detecting objects of different sizes.

It also uses something called decoupled heads—basically, separate parts of the model for deciding what an object is and where it is. This helps it make better predictions. YOLOv8 also adds smarter loss functions like CIoU and DFL, which help the model learn more efficiently and predict bounding boxes more precisely.

What's more, YOLOv8 isn't limited to object detection. It can also do instance segmentation, pose estimation, and image classification, making it a versatile choice for many real-world applications. It runs fast—up to 60 frames per second—and handles

cluttered, crowded scenes better than previous models [2, 4].

III. METHODOLOGY

This research utilizes the YOLO version 8 or YOLOv8 method. The use of YOLO in machine learning aims to detect and classify specific objects in images or videos. The results of detection and classification using YOLO can be used to count the number of objects that pass through a specific area or that have been defined in an image or a video frame. This research aims to detect people in images or videos at pedestrian areas in Kebun Raya Bogor.

This study employs an experimental approach to evaluate the performance of the YOLOv8 algorithm for detecting pedestrians in crowded urban areas. The overall stages of the research are presented in Fig. 1.

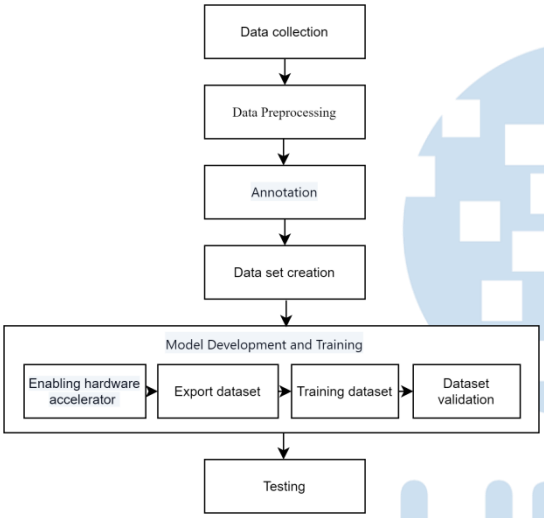


Fig. 1. Research Framework

a) Data collection

The research begins with the collection of data in the form of images and videos. After data collection, pre-processing is carried out to prepare the data for further analysis. Once the pre-processing stage is completed, the data are annotated and divided into three categories: training data, validation data, and test data. An example of the data collection location can be seen in Fig. 2. After the annotation and data division processes, the design and coding of the detection model are conducted. The final stage involves testing, where the developed model is evaluated to assess its accuracy in detecting people within Kebun Raya Bogor.

Fig. 2. Image of Pedestrian Area near Gate 3

b) Data Pre-processing

After collecting images and videos at three designated points, a project was established within the Roboflow Workspace, a software tool designed for data storage and annotation to facilitate the creation of a dataset. The initial step in the workflow involved defining a Class in Roboflow, specifically designated for "person," which serves as the identifier for the object to be detected in the study. Table 1 presents the details of the images per frame that were successfully collected from the three capture points."

Subsequently, the collected image and video files were uploaded to the platform. A crucial step in processing the video data involved extracting frames at a rate of one frame per second from each video to transform these into a uniform image format for further analysis. This frame rate was chosen to balance the need for detailed temporal resolution while managing the volume of data processed and stored, ensuring efficient handling during the subsequent stages of model training and validation.

TABLE 1. NUMBER OF IMAGES FROM CAPTURE POINTS

No	Collection poin	Number of images
1	Pedestrian near Gate 3	416
2	Pedestrian near Gate 4	418
3	Sempur Park	367
Total		**1201**

c) Annotation

This stage begins after the collection of image-formatted data and focuses primarily on identifying the "person" object within the images by using the annotation feature on Roboflow. During the annotation process, each image is carefully marked with a bounding box around each detected person. This labeling is performed manually or under supervision to ensure accuracy and consistency in object identification. The annotation process can be seen in Fig. 3.

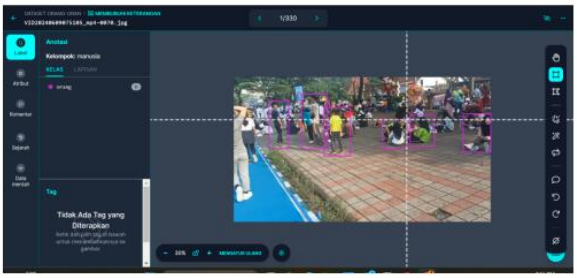


Fig. 3. Image Annotation

After labeling all images, the dataset is divided into three distinct subsets: training data, validation data, and test data. The training data is used to adjust the model's parameters and optimize its performance during the learning process. The validation data serves to periodically evaluate the model throughout training, helping to detect signs of overfitting and ensure generalization. Meanwhile, the test data is reserved for assessing the model's final performance after training is complete. The detailed distribution of the dataset across these categories is presented in Table 2, ensuring clarity and reproducibility for further analysis.

TABLE 2. IMAGE DATA DIVISION

No	Data type	Percentage	Amount of data
1	Training Data	70%	840
2	Validation Data	20%	241
3	Test Data	10%	120
Total		100%	1201

After specifying the percentages for data division, Roboflow automatically partitions the image data into three categories at random. The results of this data segmentation can be observed in Fig. 4, where each image is clearly marked to indicate its classification as training, validation, or test data, noted at the bottom left corner of each image.

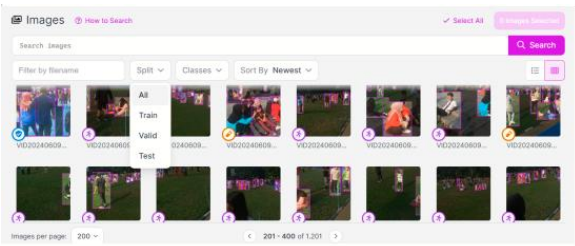


Fig. 4. Results of the Image Data Division

d) Dataset Formation

After dividing the data into three parts, the next step is to create a dataset for research purposes. However, before constructing the dataset, a preprocessing stage is required. During this stage, the images are resized to ensure that they all have the same resolution, facilitating consistent data handling during the subsequent analysis. In the preprocessing stage, all images are resized to a uniform resolution of 1280 x 720 pixels. The fitting format 'Fit (black edges) in' is selected to ensure that images with different original resolutions do not distort—either by stretching or compressing—when compiled into the dataset.

e) Model Development and Testing

In the Model Development and Testing phase, a systematic approach was undertaken to build an effective person detection system tailored for the pedestrian zones of Bogor Botanical Gardens. This phase began with the preparation of the dataset, which included image annotation and segmentation into training, validation, and test sets. Following this, the focus shifted to the design and coding stage, where the YOLOv8 algorithm—renowned for its high performance in real-time object detection—was used to train the model. The model was trained using annotated data to recognize and count individuals, addressing a key challenge in maintaining public safety in densely populated areas.

After training, the model was tested to evaluate its accuracy and reliability under real-world conditions, such as varying lighting and crowd density. This step is critical to ensure that the model performs effectively in dynamic urban environments. The development phase highlights two essential components: training the YOLOv8 model and testing its performance using the annotated dataset. The detailed steps involved in this stage are described in the following sections

1. Enabling hardware accelerator GPUT4

This research was conducted using Google Colab as the computational platform. By default, Google Colab operates on a Central Processing Unit (CPU), which tends to be slower in executing machine learning processes. To address this limitation, a T4 GPU was utilized, which is more effective in accelerating the training and inference of models in machine learning and deep learning. GPUs, particularly the T4 model provided for free by Google Colab, offer better energy efficiency and superior optimization for matrix and vector operations compared to CPUs.

2. Export Dataset

To export the dataset that has been created, it is necessary to copy the API code provided at the dataset formation stage. This code contains the necessary information to access and manage the organized dataset.

3. Training Dataset

The training dataset plays a crucial role in the development of machine learning models, serving as the primary resource for teaching the model to accurately recognize and predict outcomes. This dataset consists of pre-processed and annotated images used to adjust the model's parameters through supervised learning techniques. During training, the

model iteratively learns from this dataset by comparing its predictions against the actual outcomes, continuously improving its accuracy. In the code shown below, the dataset is trained using 100 epochs with an image size of 800 pixels. One epoch means the model has learned or recognized each data sample in the training dataset once. Increasing the number of epochs allows the model to become more familiar with the form of the object being trained. The adequacy of the training dataset directly influences the model's ability to generalize to new, unseen data, making the quality and diversity of the training examples paramount. Consequently, ensuring a comprehensive and representative training dataset is essential for the successful application of the model in real-world scenarios.

```
!yolo task=detect mode=train model={model} data={dataset_location}/data.yaml epoch=100 imgsz=800 plots=True
```

4. Dataset Validation

The validation dataset is a collection of data used for an objective evaluation of a model's performance during the training process. Not involved in adjusting model parameters, the validation dataset plays a critical role in identifying issues such as overfitting, where the model perfectly fits the training data but fails to generalize to new data. The code shown below is used to validate data, ensuring that the trained model is tested using this validation dataset. This validation process is crucial for assessing how effectively the model can predict new data and for adjusting the model to achieve an optimal balance between learning and generalization capabilities. Using this validation dataset allows developers to optimize the model, ensuring that the best-trained model performs well on the same data used during training.

```
!yolo task=detect mode=val model={HOME}/runs/detect/train/weight/best.pt data={dataset.location}/data.yaml
```

f) Testing

In the testing phase, the trained system is evaluated to assess its object detection and classification capabilities under real-world conditions. This process involves the use of video files that were previously uploaded or stored in the Google Colab directory. The test results are presented in video outputs that display bounding boxes around "person" objects, accompanied by confidence scores ranging from 0.25 to 1, indicating the level of prediction accuracy. Evaluation is conducted not only through visual analysis of the output videos but also by measuring performance metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive overview of the model's effectiveness in generalizing and detecting new, unseen objects beyond the training process. Therefore, the testing phase is a critical step in objectively validating the model's performance across various real-world conditions and scenarios.

The following outputs are used to evaluate the accuracy of the person detection model in this study:

- Confusion Matrix
  A performance evaluation table that shows the number of correct and incorrect predictions made by the model for each class. This matrix allows for detailed identification of classification errors.
- Normalized Confusion Matrix
  A normalized version of the confusion matrix, where each value is divided by the total number of predictions for the corresponding class. It provides proportions that are easier to compare across classes.
- F1-Score
  An evaluation metric that combines precision and recall into a single harmonic mean value. The F1-score is particularly useful when there is an imbalance between the number of positive and negative classes.
- Precision
  Indicates the proportion of positive predictions that are truly relevant (correct). In other words, precision measures how accurate the model is when making positive predictions.
- Recall
  Measures the model's ability to correctly identify all actual positive cases. Recall shows how many of the total positive cases were successfully detected by the model.
- Precision-Recall Curve
  A graph that illustrates the trade-off between precision and recall at various threshold levels. This curve is useful for evaluating model performance, especially in cases with imbalanced data.
- Train Batch
  A subset of the training dataset used to update the model's weights during one training iteration. In object detection, train batch images are often visualized to observe interim detection results during training.
- MAP (Mean Average Precision)
  A widely used metric for evaluating object detection models. mAP is the average of the Average Precision (AP) values across all classes, where AP is calculated as the area under the Precision-Recall curve for a given class. The mAP score provides a comprehensive overview of the model's ability to consistently detect objects across multiple categories.

To evaluate the performance of the detection model, four key parameters are used based on the model's classification results on the test data:

1. True Positives (TP): Cases where the model correctly classifies a positive object.
2. True Negatives (TN): Cases where the model correctly identifies a negative (non-target) object.
3. False Positives (FP): Errors where the model incorrectly classifies a negative object as positive.
4. False Negatives (FN): Errors where the model fails to detect a positive object and classifies it as negative.

These four parameters form the basis for calculating evaluation metrics such as precision, recall, and F1-score, which provide a comprehensive overview of the model's accuracy and reliability in object detection.

IV. RESULT

The model evaluation stage aims to measure how well the YOLOv8 model performs during training. In this study, the model was trained for 100 epochs, taking approximately one hour, using 241 validation images. The training outcomes include several key metrics:

- Precision (P): The model achieved a precision of 0.846, which indicates a high proportion of correctly predicted bounding boxes that actually contain a person. This value, being close to 1, reflects good model accuracy.
- Recall (R): With a recall of 0.858, the model successfully detected a large portion of actual people present in the images.
- mAP50: The model scored 0.951 at an 0.01 threshold of 0.5, demonstrating excellent performance in identifying objects at lower threshold levels.
- mAP50-95: The score dropped to 0.586 across stricter IoU thresholds (0.50 to 0.95), indicating reduced performance under more precise detection requirements.

These metrics illustrate that while the model performs exceptionally well under standard threshold conditions, its precision decreases as the required overlap for correct predictions increases.

4.1. Deeper Analysis of Performance Drop

A deeper analysis is necessary to identify the factors contributing to the model's performance drop, particularly the decline in mAP across stricter IoU thresholds (mAP@0.5:0.95). This decrease indicates that although the model performs well at a standard threshold, its localization precision diminishes when tighter bounding box overlaps are required.

One major cause of this performance decline is the occurrence of false positives, where the model incorrectly classifies background elements as humans. For example, objects such as leaves, tree branches, shadows, or other background features with human-like visual patterns were often misclassified, especially under challenging lighting conditions, low contrast, or dynamic backgrounds caused by wind. This issue emerges because the YOLOv8 model, despite its robust capabilities, sometimes learns features that also exist in non-human objects. Additionally, limitations in the training dataset—such as a lack of diversity in human poses, environmental conditions, and movement variations—exacerbate the risk of misclassification.

From a practical standpoint, this degradation in precision significantly affects the deployment of the model for public area surveillance. A high rate of false positives could trigger excessive false alarms, reducing the operational efficiency and trustworthiness of the surveillance system. In public safety scenarios, this could lead to confusion among security personnel and delayed responses to actual threats. Therefore, to enhance the model's practical effectiveness, further improvements are needed. These may include enriching the dataset with more varied real-world examples, adjusting the model architecture for better feature discrimination, and applying additional post-processing techniques to filter out background-induced false positives.

4.2 Evaluation of Confidence Scores

The training results in this study also utilized four types of graphs or curves as tools to compare the model's accuracy in relation to its confidence levels. These curves include the Precision-Confidence Curve, Recall-Confidence Curve, and the Precision-Recall Curve. These visualizations are used to analyze how variations in confidence thresholds affect the model's predictive performance.

Recall-Confidence Curve

As shown in Fig. 5, the x-axis represents the confidence threshold ranging from 0 to 1, while the y-axis shows precision. The curve demonstrates that as confidence increases, so does the precision, peaking at 1.00 when the confidence threshold reaches 0.956. This indicates that the model's predictions become more accurate at higher confidence levels, especially in detecting the "person" class.

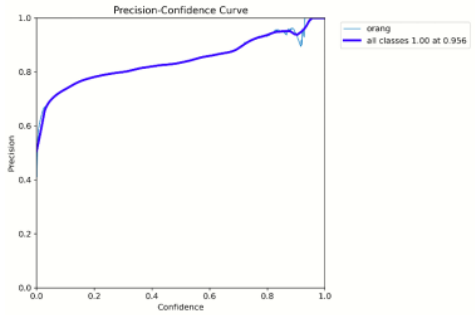


Fig. 5. Recall-Confidence Curve

Precision-Recall Curve

Fig. 6 illustrates the relationship between recall and confidence. Initially, the model maintains a high recall of 0.97 at a confidence level of 0.00. However, recall drops significantly as the confidence threshold increases. This trend suggests that while the model detects more objects at lower confidence, it becomes conservative at higher thresholds, resulting in missed detections.

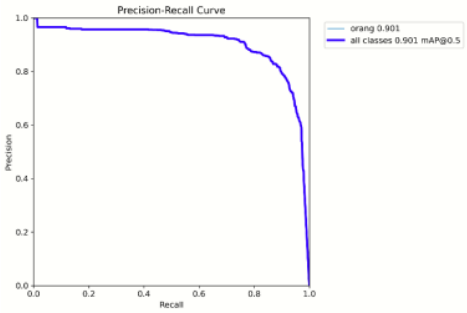


Fig. 6. Precision-Recall Curve

F1-Confidence Curve

Fig. 7 shows that F1-score, which harmonizes precision and recall, is stable at moderate confidence levels and begins to decline as confidence approaches 1. At a confidence level of 0.561, the model achieves an F1-score of 0.85, suggesting a well-balanced performance at this threshold. However, overconfidence may lead to reduced effectiveness due to increased false predictions.

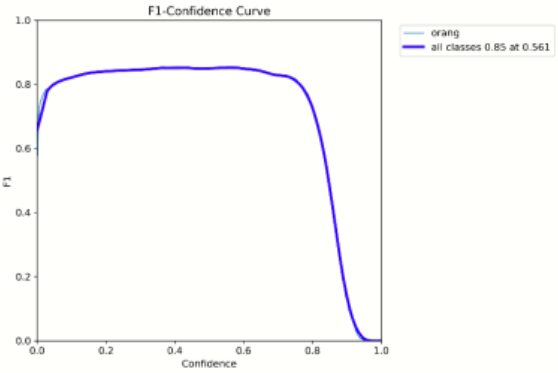


Fig. 7. F1-Confidence Curve

4.3 Confusion Matrix Evaluation

The confusion matrix analysis, presented in Fig. 8, provides a detailed breakdown of the model's classification results:

a. True Positives (TP): 0.90
b. True Negatives (TN): 0.00
c. False Positives (FP): 1.00
d. False Negatives (FN): 0.10

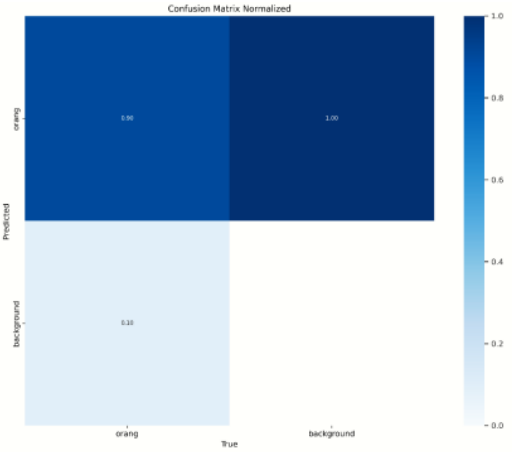


Fig. 8. Confusion Matrix Model

Based on these values, the following metrics are calculated:

a. Precision

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Precision = \frac{0.90}{0.90+1.00} = 0.47$$

b. Recall

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$Recall = \frac{0.90}{0.90+0.10} = 0.90$$

C. F1-score

$$F1score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (3)$$

$$F1score = 2 \times \frac{0.47 \times 0.90}{0.47+0.90} = 0.62$$

D. Accuracy

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

$$Accuracy = \frac{0.90+0}{0.90+1.00+0.10+0} = 0.45$$

Although the recall is high, the model's high false positive rate negatively affects precision, leading to a moderate F1-score. This result suggests that while the model is effective at detecting people, it struggles to distinguish them from background elements.

4.4. Evaluation Using Scatter Plot and Box Plot

Scatter plots are employed to visualize the distribution of bounding box coordinates and dimensions. These visual tools help in understanding the spatial and size characteristics of detected objects.

The (x, y) scatter plot in Fig. 9 shows that most bounding box centers are clustered between x = 0.1–0.4 and y = 0.2–0.6.

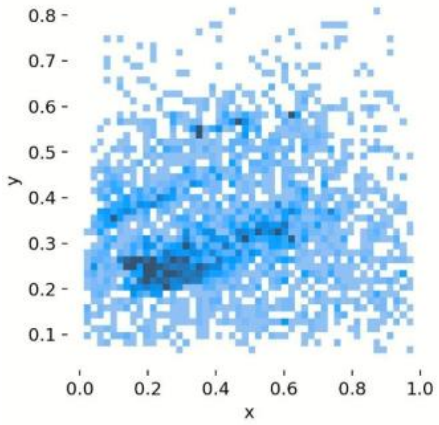


Fig. 9. Scatter Plot x,y

The (width, height) plot in Fig. 10 indicates that bounding boxes are generally taller than wide, with heights concentrated between 0.1–0.2 and widths between 0.0–0.1.

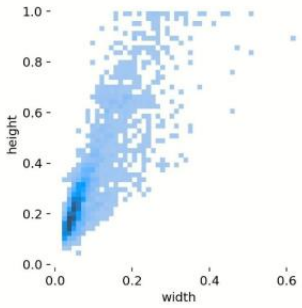


Fig. 10. Scatter plot (width, height)

Fig. 11 includes a bar chart showing fewer than 3500 bounding boxes labeled as "person" in the dataset, alongside an overlay of trained bounding boxes, demonstrating the model's coverage.

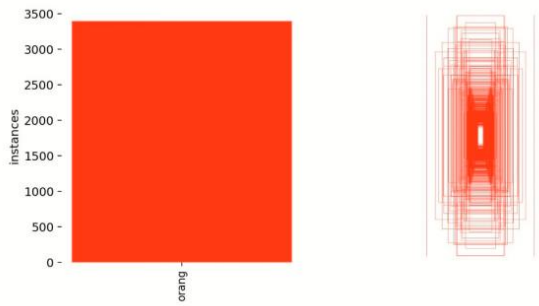


Fig. 11. Bar chart & Bounding box overlays

4.5 Training and Validation Metrics

Training and validation metrics serve as essential indicators of the model's learning progress and generalization ability. These metrics, which include loss functions, precision, recall, and mean average precision (mAP), provide insights into how well the model performs on both seen (training) and unseen (validation) data during each epoch of training. Fig. 12 provides an overall indication that the model is undergoing an effective learning process and progressively improving its performance during both training and validation phases.

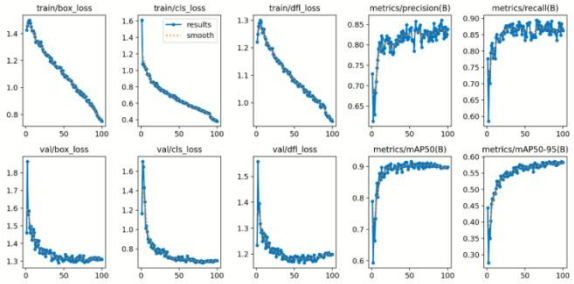


Fig. 12. Matrix & Data Validation

- The loss values (train/box_loss, train/cls_loss, train/dfl_loss) steadily decrease throughout training, indicating that the model is learning effectively in terms of localization, classification, and feature extraction.
- The validation losses (val/box_loss, val/cls_loss, val/dfl_loss) follow a similar downward trend, suggesting good generalization to unseen data.
- Metrics such as precision and recall consistently improve across epochs, showing progressive enhancement in detection performance.
- The increasing trends in mAP@0.5 and mAP@0.5:0.95 further confirm that the model's accuracy level continues to improve throughout the training and validation phases.

These graphical metrics collectively indicate that the model undergoes a positive learning process and progressively enhances its performance over time.

4.6 Train Batch

The concept of train batch refers to the process of training the model using grouped subsets of data, allowing for more stable and efficient learning. Train batching also helps minimize fluctuations in the object representations being learned, ensuring the model can generalize better and produce more consistent results. In this study, the training batch was divided into six parts derived from a total of 840 training images.

Fig. 13 presents the initial training batch, consisting of the first subset of training images introduced to the model at the beginning of the learning process. This batch plays a crucial role in setting the initial learning direction by introducing the model to the fundamental features of the target object.

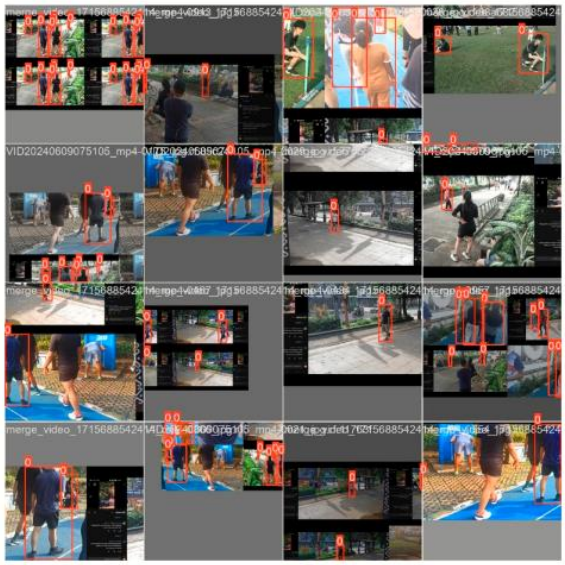


Fig. 13. Initial training batch

In contrast, Fig. 14 illustrates the final training batch, which comprises the last subset of images used during the concluding stage of training. This batch reinforces the model's learned patterns and helps stabilize its performance prior to final evaluation It helps reinforce the patterns and features the model has learned, ensuring the consistency and stability of detection performance by the end of the training cycle.

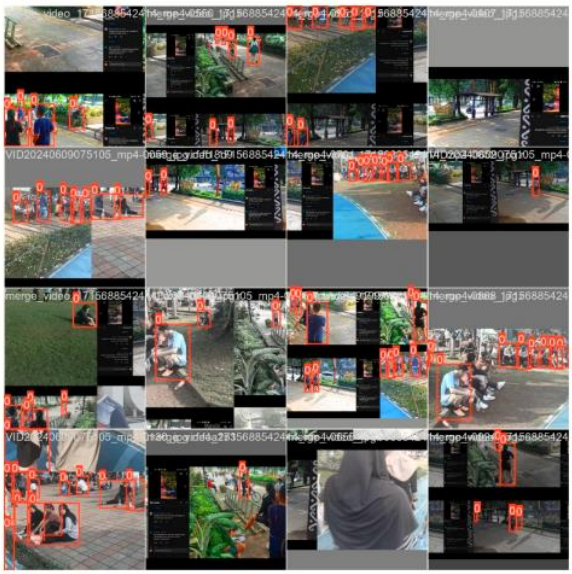


Fig. 14. Final training batch

4.7 Model Output Evaluation

The model's output was evaluated using external video samples from YouTube. Each video was analyzed for detection accuracy in various scenarios, such as crowd density, object distance, and occlusion.

- Accurately detects individuals who are visible and unobstructed.
- Struggles to detect individuals who are distant or partially blocked.
- Occasionally misclassifies non-human objects as "person" (e.g., leaves or fruit).

These findings confirm that while the model performs well under ideal conditions, its accuracy declines in complex environments, suggesting areas for further improvement.

V. CONCLUSIONS

This study successfully developed a YOLOv8-based pedestrian detection system at Bogor Botanical Garden by training it on 1,201 annotated images. Overall, the model demonstrated strong performance at moderate confidence thresholds, although its accuracy declined when stricter precision requirements were applied, particularly under more demanding IoU conditions.

The model, trained over 100 epochs with confidence thresholds ranging from 0.25 to 1.00, achieved a high precision of 0.846, a recall of 0.858, and a strong mAP@0.5 of 0.951. However, its mAP@0.5:0.95 dropped to 0.586, suggesting that while the model could detect people effectively under standard settings, it struggled with finer, more precise object localization.

Confidence-based analysis revealed that the model performed best at specific thresholds: precision peaked at a confidence of 0.956, recall reached 0.97 at a confidence of 0.00, and the highest F1-score of 0.85 was achieved at a confidence of 0.561. The Precision-Recall Curve also showed a high mAP of 0.901, indicating a good balance between precision and recall across different thresholds.

Further evaluation through the confusion matrix highlighted challenges in differentiating humans from similar-looking background elements. Although the model achieved a high recall (0.90), the precision dropped to 0.47 due to frequent false positives, resulting in a moderate F1-score of 0.62 and an overall accuracy of 0.45. These findings suggest that while the model was effective at detecting people, it sometimes misclassified objects such as leaves, shadows, or tree branches as human figures.

Scatter plot visualizations showed that most bounding box centers clustered between x = 0.1–0.4 and y = 0.2–0.6, consistent with the upright posture of standing humans. A width-to-height ratio of roughly 6:10 further supported that the model mainly detected vertical, human-like shapes. Analysis of training batches also confirmed that the model's learning process was stable and systematic throughout.

When tested on real-world videos, the model performed well in detecting clearly visible, unobstructed individuals but struggled with detecting people who were distant or partially blocked from view. Some false positives also occurred, especially when background objects resembled human features.

From a practical standpoint, the YOLOv8-based system shows real potential for integration into real-time CCTV surveillance platforms, particularly in public spaces like parks, squares, or transit areas. Thanks to its relatively lightweight design and efficient inference speed, the model can operate on moderately powered GPUs or even edge computing devices, making it a practical solution for real-world deployments. However, further fine-tuning and optimizations would be necessary to ensure reliable performance in dense or complex environments before full-scale adoption.

Future research is recommended to explore the use of keypoint detection for more precise recognition of human posture, replacing the current bounding box approach. Expanding the dataset with more diverse human movement patterns could further improve model accuracy. Additionally, optimizing the model for real-time detection and integrating it into CCTV-based surveillance systems would enhance its practical use in public safety applications

REFERENCES

[1] BBC News. (2022, October 30). Itaewon crush: At least 153 killed in South Korea Halloween crowd surge. https://www.bbc.com/news/world-asia-63440858

[2] Liu, Q., Zhang, Y., Zhao, X., & Wu, L. (2023). YOLOv8: Real-time object detection for complex scenes. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 1032–1041.

[3] Chen, L., Zhang, Y., & Zhao, H. (2023). Enhancing human behavior recognition using YOLOv8 in surveillance videos. Computer Vision and Image Understanding, 225, 103553.

[4] Wang, X., Chen, Y., & Li, T. (2024). Improved YOLOv8 for small object detection in urban surveillance. Sensors and Systems, 48(2), 204–219.

[5] Setiyadi, R., Santosa, P. I., & Nurhasanah, N. (2023). Improving object detection precision through architectural adjustment of YOLOv8 in dense crowd environments. Journal of Applied AI Research, 9(1), 89–98.

[6] Drantantiyas, D., Nugroho, Y., & Widodo, B. (2023). Evaluating YOLOv8 performance for occluded object detection in surveillance systems. Journal of Computer Science and Applications, 15(4), 321–329.

[7] Gehl, J. (2019). Cities for People. Washington, DC: Island Press.

[8] Muchlisin, Z. A. (2020). Desain kawasan pejalan kaki berbasis kenyamanan dan keamanan. Jurnal Ilmiah Teknik Sipil, 24(1), 17–25.

[9] Setyowati, A. (2017). Pedestrian friendly city: Kajian ruang pejalan kaki di kawasan pusat kota. Jurnal Teknik ITS, 6(2), A117–A122.

[10] Zhang, Y., Zhou, D., Chen, S., Gao, S., & Ma, Y. (2021). Crowd counting via scale-adaptive convolutional neural network. IEEE Transactions on Circuits and Systems for Video Technology, 30(1), 20–32.

[11] Ilyas, N., Mahmood, A., Mahmood, T., & Rho, S. (2020). A comprehensive survey of crowd counting and density estimation techniques. ACM Computing Surveys, 53(6), 1–37. https://doi.org/10.1145/3417982

[12] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. Cambridge, MA: MIT Press.

[13] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84–90. https://doi.org/10.1145/3065386

[14] Ashtari, S. (2022). Computer vision and its real-world applications. Journal of Artificial Intelligence Research, 78(3), 456–470.

ISSN 2085-4552

Optimizing CV Matching with Job Vacancies Using the Boyer-Moore Algorithm

Andrew Christofer Hadiwinata [1], Eunike Endariahna Surbakti [2]
[1,2] Informatics Department, Faculty of Engineering & Informatics,
Universitas Multimedia Nusantara Tangerang, Indonesia
[1]andrew.hadiwinata@student.umn.ac.id, [2]eunike.endariahna@umn.ac.id

Accepted 20 Mei 2025
Approved 28 Mei 2025

***Abstract*— According to a survey conducted by a public survey agency, there was a decline in the percentage from 37% to 20%, indicating that a gap still exists between the skills required by the job market and those possessed by job seekers. To address this issue, this study aims to assess the alignment between data from curriculum vitae (CV) and job vacancies. The Boyer-Moore algorithm is implemented through a web-based system. The system extracts text from a PDF file, which is then used in the application of the Boyer-Moore algorithm. The system also retrieves selected job vacancy data and generates keywords using YAKE (Yet Another Keyword Extractor). Before processing with the Boyer-Moore algorithm, all data undergoes pre-processing. The algorithm's output is either a "match found" or "no match found." The similarity score is determined by dividing the number of matching keywords by the total number of keywords. Additionally, the system recommends other job options, aiming to suggest alternative vacancies that may better match the CV. These recommendations are based on the highest percentage of keyword matches from all job vacancy data stored in the system, which are sorted accordingly. The Boyer-Moore algorithm was successfully implemented in the job vacancy system, and the system's performance evaluation, using 100 job vacancy data entries, yielded an average processing time of 2.84438 seconds.**

***Index Terms*— *Boyer-Moore, curriculum vitae, Similarity score*.**

I. INTRODUCTION

According to data from the Central Statistics Agency (BPS), 937,176 people registered or applied for jobs in Indonesia in 2022. Job seekers' educational backgrounds play a crucial role in the application process. Various job vacancies in Indonesia require different minimum education levels, ranging from high school to bachelor's degrees. However, the proportion of university-educated workers in the labor force is only around 9.92%. This discrepancy highlights the gap between required job skills and the qualifications possessed by job seekers. Notably, a significant number of workers (up to 63%) end up in jobs unrelated to their field of study. To address this, an effective system is needed one that helps job seekers identify relevant skills for their desired positions and assists recruiters in selecting candidates whose CVs align with job criteria. Unlike some other algorithms, Boyer-Moore doesn't require preprocessing, making it efficient for large-scale searches [1]. Boyer-Moore performs pattern matching by sliding the pattern over the text. It starts matching from the last character of the pattern, which allows for faster processing. The algorithm combines two heuristics: the Bad Character Heuristic and the Good Suffix Heuristic. When a mismatch occurs, the algorithm identifies the "bad character" (a character in the text that doesn't match the current pattern character). It then shifts the pattern to align with the last occurrence of the bad character in the pattern. This heuristic improves efficiency by skipping unnecessary comparisons [2]. The Boyer-Moore algorithm also considers the context of matching characters. If a mismatched character in the text occurs somewhere in the pattern, its index is used to skip over more characters, further reducing the number of comparisons. Boyer-Moore can be applied to strings of varying lengths and characters, making it suitable for CVs and job vacancy matching [3].

The research aims to implement the Boyer-Moore algorithm for assessing the suitability of CVs to specific job vacancies. It evaluates the efficiency of the algorithm based on computation time and focuses on English language patterns in PDF files. The goal is to assist HR professionals in streamlining the process and achieving more precise job candidate matches.

II. METHODOLOGY

A. Literature Studies

Literature studies aim to determine the development of knowledge, theories, concepts, methods, and current findings related to the problem that you want to research. Literature study includes searching, reading, analyzing, and evaluate relevant scientific sources about

algorithm implementation Boyer-Moore towards a system. There are references to pre- vious research regarding implementation of the Boyer-Moore algorithm taken from a written journal by Sara Nasr. In This research, found the best candidates using CV and process fuzzy or uncertain information with berth algorithm [4]. The Boyer-Moore algorithm is an efficient string search method. Discovered by Bob Boyer and J. Strother Moore, it has become a standard in string search literature [1] [6]. Key characteristics of the Boyer-Moore algorithm include right-to- left string matching. This approach allows the algorithm to skip further when encountering mismatches, avoiding unnec- essary character comparisons [6].

The *Boyer Moore* algorithm is a *string* search algorithm that matches characters from right to left, and uses two rules to shift the pattern to the right if a mismatch occurs, namely the *bad character* rule and the *good rule suffix* [1]. This algorithm requires two pre-search tables, namely the *bmBc* table and the *bmGs* table, which store the distance between each character or suffix and the end of the pattern. The mathematical formula for the *Boyer Moore* algorithm is

$$s = \max\{bmBc[\text{text}[i]] - m + 1 + j, bmGs[j]\} \quad (1)$$

B. Dataset

Data collection involves obtaining relevant data from both primary and secondary sources. It must be done in a valid, reliable, and ethical manner. In this research, a dataset named “Job Dataset” from Kaggle is used for job vacancy data [5]. The relevant fields include company name, city, country, job title, role, salary, description, skills, and requirements.

C. Requirements

The system requirements for implementing the Boyer- Moore algorithm include allowing users to view all job postings. Users can scan their CV’s against specific job criteria and match their CV data with job descriptions and skills requirements. The matching process involves extracting keywords using YAKE (Yet Another Keyword Extractor) from job descriptions and skills requirements, followed by applying the Boyer-Moore algorithm for accurate matching. The system should display the match percentage for the selected job and provide alternative options with the top 5 match percentages.

D. Planning

Planning refers to the process of creating a plan or sketch for the functions, features, and technologies to be used in a system under development. It involves creativity, innovation, and alignment with user specifications and needs. The design phase encompasses system blueprints, user and system requirements, sitemaps, flowcharts, and low-fidelity prototypes. Here is the flowchart that defines the web-based system process.

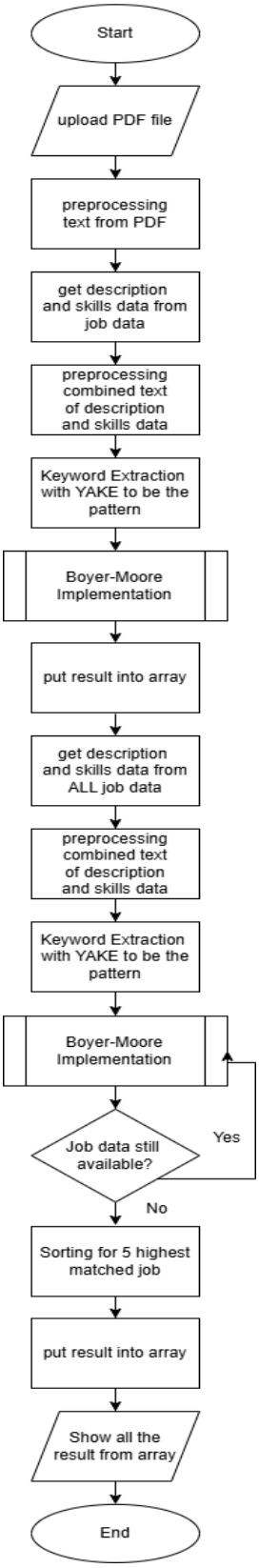


Fig. 1. System Flowchart

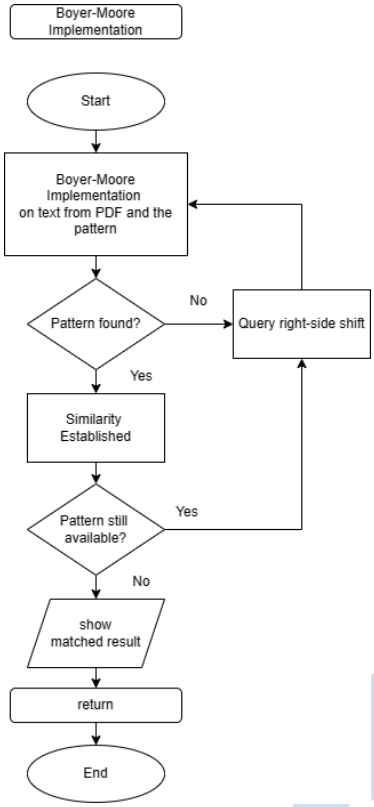


Fig. 2. Boyer-Moore Implementation Flowchart

The system design at Figure 1 System flowchart and Figure 2 Booyer Moore Implmentation, outlines the process for implementing the Boyer-Moore algorithm. It begins by requesting user input in the form of a PDF file. The text within the PDF is then extracted and tokenized, followed by pre- processing (including removing non-alphanumeric characters and converting to uppercase). Next, the system combines job description data with skills data based on the selected job. The resulting combined text is tokenized and preprocessed as well.

For string matching using the Boyer-Moore algorithm, both the text and pattern are needed. The text comes from the extracted and preprocessed data from the PDF, while the pattern is derived from the combined job description and skills text. To obtain the pattern, keyword extraction using YAKE is performed [8]. The Boyer-Moore string matching process begins, adding match percentages when a pattern is found. If no match occurs, a query right-side shift is applied. This process continues until the entire pattern is exhausted. After Boyer-Moore completes, the match percentage is displayed. For example, if 20 keywords were extracted, but only 15 matched, the system shows a 75% match. These percentages are stored in an array.

In the recommendation system, sorting is crucial to display the most relevant results to users. The system proceeds to obtain the top 5 recommendations. Sorting involves taking the match percentages generated by Boyer-Moore and arranging them from highest to lowest. The same process is applied to all job postings, resulting in 5 additional recommendations. These top 5 matches are also stored in an array. Once all steps are completed, the system displays the array results.

III. RESULT

A. Implementation

The implementation of the Boyer-Moore algorithm within a Python-based web system requires the Flask library. Flask handles user input from the website, processes data using the algorithm, and returns the results to display on the website. The system follows the flowchart design outlined in Figure 1, involving steps such as selecting job postings, uploading documents, processing data, and displaying results along with recommendations.

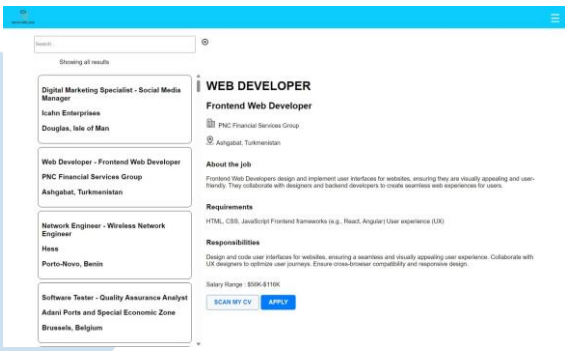


Fig. 3. Job Listings Page

Figure 2 Job Lisitings Page is the job listing page allows users to view all available job postings. It displays job titles, roles, and locations. When a specific job is selected, detailed information including job descriptions and skills requirements is shown on the right side of the page. Below this information, a ”SCAN MY CV” button redirects users to the document upload page.

Fig. 4. Upload Document Page

The “Upload Document” page, as depicted in Figure 4, allows users to upload PDF files related to job applications. It displays job titles, roles, and locations. After selecting a file, users can press the “Scan Now” button to initiate the process. To extract text from a PDF, libraries like PyPDF2 are com- monly used [7]. Preprocessing involves tokenization (splitting text into

individual words) and removing stop words (common words with little meaning). The extracted text is combined and then processed further. YAKE (Yet Another Keyword Extractor) is used for keyword extraction [8].

The Boyer-Moore algorithm, an efficient string search method, doesn't require extensive preprocessing, making it suitable for large-scale searches. It compares text and pattern from right to left. Two key heuristics are employed: the bad character rule (adjusting the pattern based on the last mismatched character) and the good suffix rule (shifting the pattern if the mismatched character doesn't appear elsewhere in the pattern). By combining these heuristics, Boyer-Moore achieves efficient pattern matching.

```
class BoyerMooreSearch:
    def __init__(self, text, pattern):
        self.text = text
        self.pattern = pattern
        self.m = len(pattern)
        self.n = len(text)
        self.skip = []
        for i in range(256): self.skip.append(-1)
        for i in range(self.m): self.skip[ord(pattern[i
])] = i

    def bad_character_heuristic(self):
        i = 0
        while i <= self.n - self.m:
            j = self.m - 1
            while j >= 0 and self.pattern[j] == self.
text[i + j]:
                j -= 1
            if j < 0:
                return [i]
            else:
                i += max(1, j - self.skip[ord(self.text
[i + j])])
        return []

    bm_search = BoyerMooreSearch(text, word)
    positions = bm_search.bad_character_heuristic()
```

Fig. 5. Boyer-Moore Algorithm

BoyerMooreSearch(text, word) : This is the creation of a BoyerMooreSearch object with the text and words (pattern) you want to search for. bad_character_ heuristic(): This func- tion is an implementation of the bad character heuristic in the Boyer-Moore algorithm. This function returns the position where the word (pattern) is found in the text. _init_(self, text, pattern): This is the class constructor. This function takes two arguments, text and pattern, which are respectively the text on which the search is performed and the pattern to search for. This function also initializes some variables used in the algorithm. bad_character_heuristic(self): This is a method that implements the bad character heuristic of the Boyer-Moore algorithm. This heuristic speeds up the search by skipping characters that do not match in the pattern. The first loop (while i ¡= self.n - self.m:) moves the search window through the text. The second loop (while j ¿= 0 and self.pattern[j] == self.text[i + j]:) compares the characters in the pattern and text from right to left. If a pattern is found (if j ¡ 0:), the starting index of the pattern in the text is returned. If a non-matching character is found, the search window advances a maximum distance between 1 and j - self.skip[ord(self.text[i + j])].

The results of string matching using the Boyer-Moore algorithm display job titles, positions, company names, lo- cations, and match percentages. After completing the Boy- erMooreSearch process, the results are available only for the selected job. To obtain job recommendations with the highest match percentages, testing is performed across all job data stored in the database. The process is similar to before, but without using a unique ID since the query aims to select all data. Iterating through all data, job descriptions and skills requirements are combined and pre-processed. Keyword ex- traction is then applied, followed by matching using the Boyer- Moore algorithm, resulting in the same format. Additional job recommendations are displayed to provide alternative options that may have higher match percentages than the selected job. Sorting with a Sortation Algorithm which is Timsorts determines the highest match percentage recommendations [9] [10].

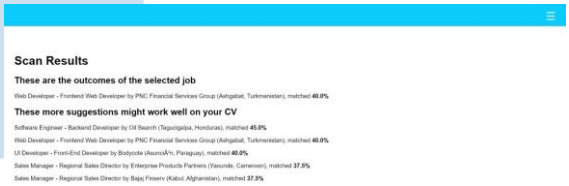


Fig. 6. Result Page

Once the top recommendations are obtained, the results are stored in an array and sent to the page depicted in Figure 6, the "Results Page," using the render template function in Flask. The "Results Page" displays job titles, roles, locations, and match percentages. The top result corresponds to the selected job, while the lower section shows additional recommenda- tions with the highest match percentages.

B. Testing

The testing process for the Boyer-Moore algorithm applied to CV data matching with job vacancies. The testing process involves several steps:

1) File Upload: Begin by uploading a simple CV in PDF format.
2) Processing: Extract the text from the file and convert it into string tokens.
3) Text Pre-processing: Perform pre-processing on the ex- tracted text.
4) Select Job Vacancy Data: Choose a specific job vacancy to match against.
5) Combine Job Description and Skills

Requirements: Pre- process the text from the job description and skills requirements, combining them into a single text.

6) Keyword Extraction with YAKE: Use the YAKE algo- rithm to extract keywords. Set the parameter 'n' to 1 to ensure that the generated keywords consist of single words.
7) String Matching with Boyer-Moore: Apply the Boyer- Moore algorithm to search for the extracted keywords. The algorithm provides accurate outputs for both "match found" and "no match found" cases.
8) Calculate Matching Score: Based on manual testing, there was an 8 out of 20 keyword match, resulting in a 40% match percentage. Additionally, processing one CV against one job vacancy took less than 1 second.

Overall, the system is categorized as fast in displaying results

C. *Evaluation*

Evaluating an algorithm multiple times is crucial to obtain accurate results. Various factors, such as operating system conditions and processor load, can affect execution time. Therefore, conducting multiple tests ensures consistent and reliable outcomes. While there is no strict rule about the exact number of evaluations, averaging results over multiple runs reduces variability and provides a more accurate performance assessment.

The primary reason for measuring computational time dur- ing algorithm evaluation, especially for the efficient Boyer- Moore string search algorithm, is to assess its reliability and speed in real-world applications. In unsupervised learning contexts, where algorithms work with unlabeled data, speed and reliability are critical. Algorithms must identify patterns or structures in data without external assistance.

Computational time measurement is crucial for systems dealing with large or real-time data processing. Reliable and fast algorithms enable accurate recommendations or results within short timeframes. This is valuable in practical applica- tions like job matching, text analysis, and recommendation systems. Therefore, evaluating the Boyer-Moore algorithm considers both accuracy and speed.

The following are the table of average evaluation results.

TABLE I
AVERAGE EVALUATION TIME RESULT

Data	CV AndrewChristofer (Time Taken)	CV Lydia (Time Taken)
100	3.05692 s	2.84438 s
200	5.78710 s	5.36939 s
500	15.19051 s	12.92706 s
1000	28.64033 s	26.95702 s

The increase in processing from 100 data to 200 data nearly doubles the required time. From the data, we can infer that processing 100 data takes approximately 3 seconds. Below is the graph generated based on the average evaluation results from Table 1.

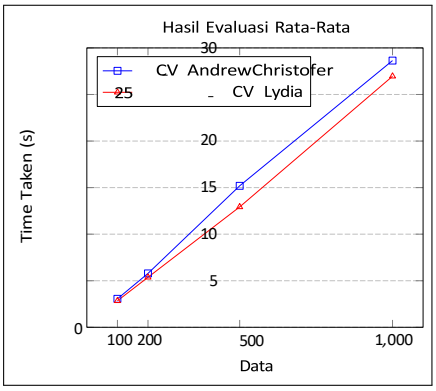


Fig. 7. Average Evaluation Time Result Graph

The graph illustrates the average evaluation results for these two CVs. The x-axis represents the processed data volume, while the y-axis shows the time required for processing. As expected, processing time increases with larger data volumes a common occurrence in computing.

For CV AndrewChristofer (indicated by the blue line), the processing time was approximately:

1) 100 data: 3.06 seconds
2) 200 data: 5.79 seconds
3) 500 data: 15.19 seconds
4) 1000 data: 28.64 seconds

Similarly, for CV Lydia (indicated by the red line), the processing time was approximately:

1) 100 data: 2.84 seconds
2) 200 data: 5.37 seconds
3) 500 data: 12.93 seconds
4) 1000 data: 26.96 seconds

The difference in processing time between the two CVs diminishes as the data volume increases. This suggests that both CVs perform similarly for large-scale data processing. In summary, while the number of words and characters does not significantly impact computational time, the data volume does affect processing time

IV. CONCLUSIONS

Implementation of the Boyer-Moore algorithm within a website-based system for checking the compatibility of cur- riculum vitae (CV) with job vacancies is successful. The implementation process involved several key steps. Literature Study, prior to implementation, thorough literature research was conducted to ensure that the chosen approach addressed the problem statement effectively. Data Collection and Anal- ysis, a dataset of job vacancies served as a

reference for eval- uating CV compatibility. Additionally, requirements analysis and system design formed the foundation for the website-based system. Algorithm Integration the Boyer-Moore algorithm was integrated into the system, aiming to assess the suitability between CV documents and job vacancy data. Additional algorithms, such as YAKE for keyword extraction and Timsort for sorting, were also utilized. Manual Testing and Evaluation, manual testing verified whether the system's processed data matched the expected output. Finally, performance evaluation measured the time required to process a PDF CV against job vacancy data and provide recommendations based on the highest match percentage.

The evaluation revealed that while word count and character length had minimal impact on computational time, the data volume significantly affected processing time.

REFERENCES

[1] R. S. Boyer and J. S. Moore, ""A Fast String Searching Algo- rithm","Communications of the ACM, vol. 20, no. 10, pp. 762–772, 10 1977

[2] W. Rytter, ""Correctness of the Boyer-Moore Algorithm"," Information Processing Letters, vol. 9, no. 5, pp. 232–234, 12 1979

[3] OpenCV. (2024, 2) "Feature Matching". [Online; accessed 17-Feb-2024].

[4] S. Nasr and O. German, "Resume searching to decide best candidate based on relief method," Open Science Journal, vol. 5, no. 2, 2020

[5] R. S. RANA, "A Comprehensive Job Dataset for Data Science, Research, and Analysis." [Online]. Available: https://www.kaggle.com/datasets/ravindrasinghrana/job-description- dataset

[6] Fitriyah, "The Implementation of Boyer-Moore Algorithm in WEB Based Computer and Informatic Terms Dictionary," in 2020 4th Inter- national Conference on Vocational Education and Training (ICOVET), 9 2020

[7] T. Aggarwal, "Pypdf2: A comprehensive guide to mastering pdf manip- ulation with python," 2023

[8] A. M. J. A. J. Ricardo Campos, Ga el Dias, "YAKE: Yet Another Keyword Extractor"

[9] S. Valdarrama, "Sorting Algorithms in Python"

[10] GeeksforGeeks, "TimSort - Data Structures and Algorithms Tutorials."

[11] S. I. Hakak, A. Kamsin, S. Palaiahnakote, and G. A. Gilkar, "Exact String Matching Algorithms: Survey, Issues, and Future Research Di- rections," ResearchGate, 2019

[12] A. K. Singh, A. K. Singh, and A. K. Singh, ""A String Matching Algorithm for Job Searching and Skill Analysis"," International Journal of Computer Applications, vol. 117, no. 16, pp. 10–14, 5 2015

[13] R. S. Patil and S. S. Sherekar, ""A Novel Approach for Skill Matching in Job Recruitment System Using String Matching Algorithm"," Interna- tional Journal of Computer Science and Information Technologies, vol. 5, no. 3, pp. 3627–3630, 2014

[14] A. K. Singh, A. K. Singh, and A. K. Singh, ""A Novel Approach for Word Search Puzzle Game Using String Matching Algorithm"," International Journal of Computer Applications, vol. 120, no. 19, pp. 1–4, 6 2015.

[15] OpenCV. (2024, 2) "Feature Matching".

[16] CodeCrucks. (2024, 2) "String Matching Algorithms - CodeCrucks"

[17] D. Gusfield, Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology. Cambridge, UK: Cambridge Uni- versity Press, 1997.

[18] A. P. A. M. J. C. N. A. J. Ricardo Campos, Vitor Mangaravite, "Unsupervised Multiword Extraction for Keyword Generation."

ISSN 2085-4552

SI-MAMAS : Design and Development of a Mobile-Based Mosque Management Information System

Muhamad Alda
Department of Information System, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia
muhamadalda@uinsu.ac.id

Accepted 21 Mei 2025
Approved 28 Mei 2025

***Abstract*— The Nurul Huda Mosque is one of the mosques in the city of Medan. So far, the Nurul Huda Mosque has not provided a facility or media that can be used to manage data, such as management data, financial data, activity data, and inventory data. The data management process is still carried out conventionally on a book, so several problems still occur, especially regarding data completeness, data damage, and loss of data that has been processed. The aim of this research is to build a mobile-based information system that can be used to manage mosque data. In building the information system, the author used the Rapid Application Development (RAD) method, which consists of planning stages, design workshops, and implementation. Meanwhile, the tool used by the author in creating information system designs visually is UML (Unified Modeling Language). The mobile-based information system was built using the Kodular framework and the Airtable database. With this information system, it is hoped that it can help the Nurul Huda Mosque process data easily and quickly via an Android smartphone.**

***Index Terms*— *Android; Management; Mosque; Mobile; Information Systems*.**

I. INTRODUCTION

The modern era we are currently living in has witnessed very rapid developments in several science and technology sectors. This progress directs how each person, community, group, or organization can deal with it. Effective management, or governance, is one way to overcome this. Effective organizational management can produce good and high-quality results. Therefore, this management is very important for any organization [1].

A mosque is a physical location and representation of Islam. Mosques play a very important role in the service of the people as a gathering place for religious activities, a place for studies of the Koran and Al-Hadith, and a cultural center for Muslims [2].

Poor management will prevent a mosque from operating at its highest potential, regardless of its complexity or simplicity. Because productive and effective administration is a must for every organization or institution, As a result, in a mosque, management must be the main focus. Increasing awareness of the purpose of all mosque management activities, namely, developing human resources for organizational success, is necessary to achieve this situation [3].

Nurul Huda Mosque is one of the mosques in Medan City, located at Jl. Marelan Raya, Tanah Enam Hundred, District. Medan Marelan, Medan City, North Sumatra 20244. So far, the Nurul Huda Mosque has not provided a facility or media that can be used to manage data, such as management data, financial data, activity data, and inventory data. The data management process is still carried out conventionally on a book, so several problems still occur, especially regarding data completeness, data damage, and loss of data that has been processed. Aside from that, the Nurul Huda Mosque also does not provide technology-based facilities for conveying information to the public. The process of conveying information still uses mosque loudspeakers, so that the process of conveying information and receiving information by the public becomes ineffective.

Similar research conducted by Budhy et al. in 2021 discussed the steps involved in developing a website-based mosque management information system using the PHP programming language and using the waterfall development method at the Baitul Ikhwan Mosque. Now it is easier for the public to find information about mosque management thanks to this website, and it also makes it easier for mosque managers to summarize data by printing it in softcopy in PDF format because it will be easier to search for files in the system and print them [4].

Similar research was conducted by Yanni Suherman and Erien Nada Azandra in 2021, discussing that managers of the Sungai Limau Grand Mosque often have difficulty managing data due to the large number of mosque operations. A waterfall system development approach is used to create a system that will really help administrators process data related to mosque management. This system utilizes a web-based application program for program planning and report preparation [5].

Similar research was conducted by Wagino in 2020 discussing the Sabilal Muhtadin Grand Mosque, Banjarmasin, which had created a Management Information System to computerize the management of the Banjarmasin Sabilal Mosque management body, which is currently still manual in its management. This information system was created with the help of Microsoft Access 2013 and Borland Delphi 7.0 software. The management of the Sabilal Muhtadin Banjarmasin Mosque considers this information system to be very useful, for example, controlling the use of the computerized multi-purpose hall and archiving data on mosque management, cash in and cash out data, data on lecturers, data on kiai and Friday prayer speakers, as well as data on the imam who leads obligatory prayer [6].

The authors' research differs from previous research in several important ways. The author conducted research aimed at building a mobile-based information system that can be utilized by Nurul Huda Mosque administrators in managing mosque data, including financial data and inventory data, and can also be used as a medium for conveying information to the public.

II. METHODOLOGY

A. System Development Method

Rapid Application Development (RAD) is the information system development approach used in this research. The three key phases of this methodology are planning, RAD design, and implementation. Speed in the development process is also prioritized with the RAD development method. where the development process can be completed quickly [7]. The RAD model can be seen in Figure 1.

Fig.1. RAD Method Stages

The stages contained in RAD can be explained as follows [8]:

1. Analysis of problems and identification of needs for the information system to be built. Actions taken include:
   - Identifying what happened at the Nurul Huda Mosque related to the process of processing financial data, inventory, and conveying information.
   - Identify functional requirements and information needs.
   - Presents alternatives to the suggested system. This section lists any requirements for software, hardware, and human resources.
   - System selection and continuity. choosing one of several alternative system solutions available.
   - Object modeling. In this section, we will model the system using objects and related classes.
2. RAD Design Workshop: At this point, the author explains about creating an information system. Modeling the information system to be created is the current process step. UML (Unified Modeling Language), which includes use case diagrams, sequence diagrams, activity diagrams, and class diagrams, is used to model this system. In this step, the new system design is written and described, and the following tasks are completed:
   - Combines diagrams with features that allow the creation of models, outputs, processes, and transactions using specific symbols.
   - Designing database and table requirements
   - Display interface class design.
3. Implementation, which involves converting all previously completed work into an information system that utilizes the Kodular framework and Airtable database.

B. Method of Collecting Data

Data collection procedures are the fundamental methods used in a study to gather data, which is the primary objective of the investigation. The author employs methodologies for data collection in conjunction with the system development approach. In the absence of familiarity with the methodologies, researchers will be unable to gather data that adheres to pre-established criteria [9].

The author collected the necessary data for this research activity using the following method:

1. Observation

Data collection occurs through direct observation and recording of research objects. At this stage, the researcher recorded matters relating to the process of processing financial data, inventory, and information at the Nurul Huda Mosque.

2. Interview

Gathering information through direct conversations with the chairman of the Nurul Huda Mosque management and the community. Here, researchers ask questions related to the process of processing financial data, inventory, and conveying information.

3. Literature review

To search for ideas or concepts that might serve as a basis for a theory or research framework, to identify the best research procedures, and to contrast current theory with empirical data.

III. RESULTS AND DISCUSSION

A. Analysis of Problem and Identification

1. Analysis of Problems

Before carrying out the design stage, the author first analyzes the problems in the running system. The analysis carried out by the author focuses on the financial processing process, the mosque inventory processing process, and the process of conveying information to the public. Based on the results of the author's observations at the Nurul Huda Mosque and the results of the author's interviews with the chairman, treasurer, facilities infrastructure department and the public, the financial processing and inventory process of the mosque is still carried out manually, namely recorded in books, while the process of conveying information to the public is still carried out using loudspeakers or mosque toa. Based on this process, there are still several problems or obstacles that occur, including:

- The mosque still processes its finances and inventory by hand. So it requires a lot of time, money, and energy.
- Difficulty searching for data.
- Data damage and loss occur.
- Loudspeakers or mosque toa do not provide clear communication.
- The scope of information delivery is limited.

2. Identify System Requirements

After conducting a thorough analysis of the current system, the author suggests a solution to address this issue. The proposal involves creating an Android-based mosquito management information system. This system will aid Nurul Huda Mosque administrators in managing financial data, inventory data, and other related processes. The solution will expedite the efficient distribution of information to the general population using Android smartphones. The author employed the following strategy to gather the necessary data for this research endeavor:

.

B. RAD Design Workshop

The information system design process that will be built uses the Unified Modeling Language (UML) model, which consists of use case diagrams, sequence diagrams, activity diagrams, and class diagrams.

1. Use Case Diagram

A use-case diagram is a graphical representation that depicts the standard interactions between system users and the system itself by narrating the system's usage story. A use-case diagram comprises an actor and the interactions they execute. The system's actors can include humans, hardware, other systems, or other entities that interact with the system. The author proposes a solution to address this issue based on the findings of the ongoing system analysis. The proposal involves creating an Android-based mosquito management information system. This system will aid Nurul Huda Mosque administrators in managing financial data, inventory data, and other related tasks. The solution will expedite the efficient distribution of information to the general population using Android smartphones. Meanwhile, the author used the following strategy to collect the necessary data for this study [10].

The Mosque Management information system was built on a multiuser basis so that it can be accessed by different users, namely the head of the management. Treasurer, facilities and infrastructure sector, public relations, and community sector. The use case diagram of the mosque management information system to be built can be seen in Figure 2.

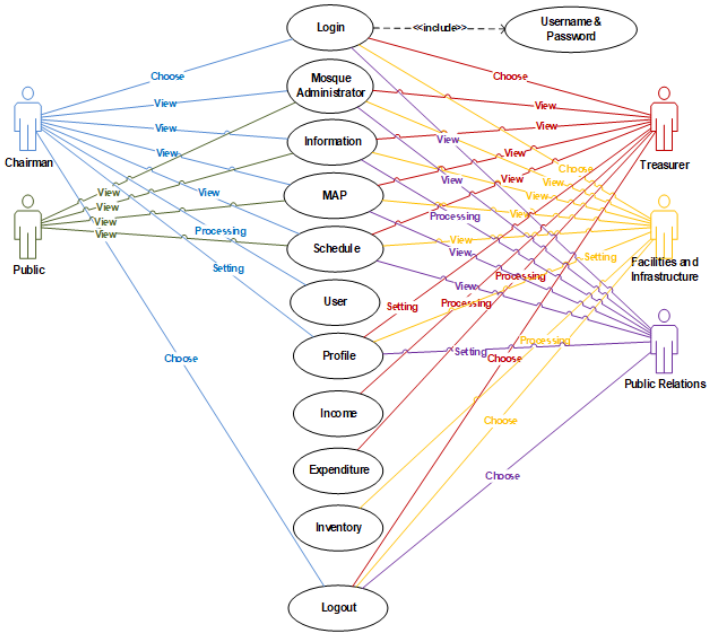


Fig.2. Use Case Diagram

Figure 2 presents the use case diagram for the upcoming information system. Use case diagrams depict the actions that actors can perform within the information system under development. The mobile-based mosque management information system is accessible to the chairman of the management, treasurer, and the facilities and infrastructure department. Section on Humanity and Society. Prior to accessing the information system, every user is required to log in by providing a user name and password and selecting pre-established access privileges. The chairman holds the authority to manage user data, the treasurer oversees income and expenditure data, the facilities and infrastructure department manages mosque inventory data, the public relations department manages mosque activity schedules and disseminates information to the public, and the public has the ability to access and view the managed information. The public relations department of the mosque issues these schedules.

2. Sequence Diagram

Sequence diagrams depict the exchanges and communications that occur among items. Sequence diagrams provide a detailed description of the actions and interactions that occur in a particular event. A use-case diagram illustrates a set of sample objects and the messages that are sent between them [11]. Illustrates the sequential or chronological interactions among items within the system, such as users, displays, and others, through the exchange of messages [12].

Sequence diagrams are divided into chairman sequence diagrams, treasurer sequence diagrams, infrastructure sector sequence diagrams, and public relations sequence diagrams. The treasurer's sequence diagram This sequence diagram illustrates the interaction between the treasurer and the information system in order to acquire the necessary information. Figure 3 displays the Treasurer's Sequence Diagram. The sequence diagram of the mosque management information system to be built can be seen in Figure 3

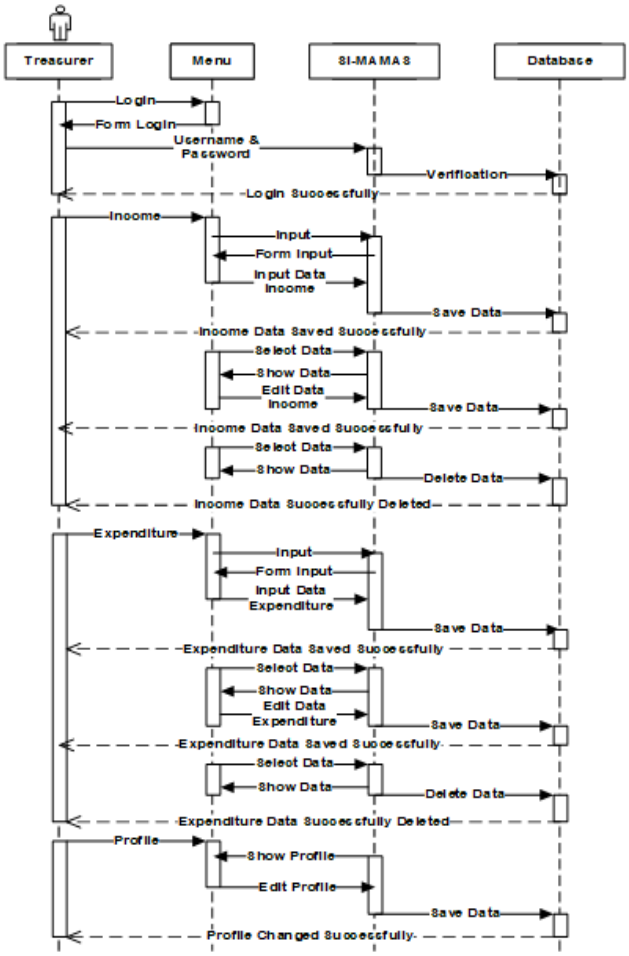


Fig.3. Use Case Diagram

Figure 3 depicts the sequence diagram of the information system under construction. The sequence diagram for the forthcoming mobile-based mosque management information system begins with the treasurer's authentication into the system. When the username and password are entered during the login procedure, the system will verify the credentials against the database. Next, the treasurer will navigate to their designated page. When the treasurer selects the income or expenditure menu, they can input, edit, or

delete the income or spending data. The treasurer has the ability to enter revenue or spending data into the information system using the provided form, and then store the data in the database. Within the edit feature, the treasurer has the ability to choose either revenue or spending data and modify the selected data before saving it. In the delete option, the treasurer has the ability to select and erase specific income or spending data from the database. The profile menu allows the treasurer to access and manage their profile data.

Activity diagrams depict the sequence of activities within a planned system, beginning with its initiation and encompassing the decision-making process until its completion. Activity diagrams can also represent concurrent processes that take place during several executions. [13]. The activity diagram of the mosque management information system to be built can be seen in Figure 4.

3. Activity Diagram

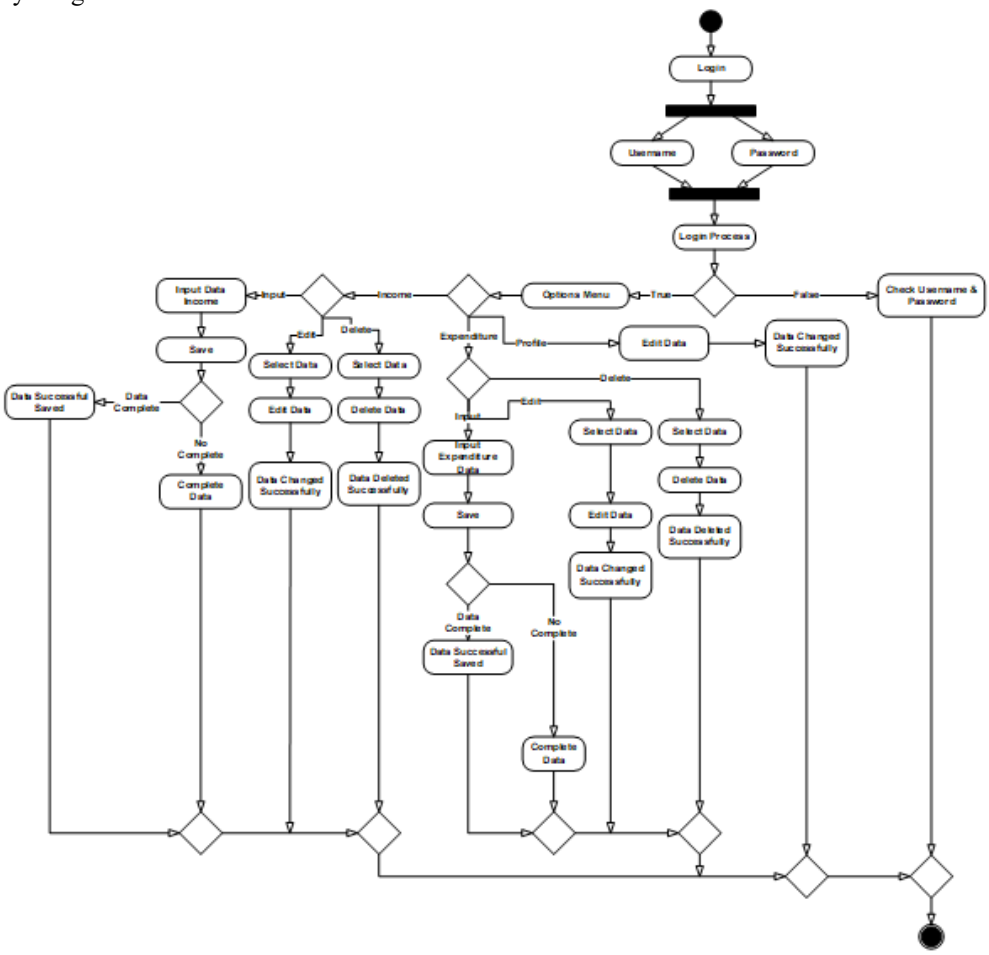


Fig.4. Activity Diagram

Figure 4 illustrates the activity diagram. The activity diagram for the forthcoming mobile-based mosque management information system begins with the treasurer's login to the system, which is completed by inputting the designated username and password. Subsequently, the information system authenticates the username and password provided by the treasurer. Provided that the username and password are accurate, the treasurer is able to access the choices menu page and choose from the available menu items. If the login and password provided are incorrect, the treasurer will be unable to access the information system and will need to verify the authenticity of the submitted username and password. Within this information system, the treasurer has the ability to choose from multiple menu selections. The treasurer will receive input choices to update or delete income or spending data in the information system upon selecting the revenue or expenditure menu. The input option in the information system verifies the integrity of the revenue or spending data entered by the treasurer. The database can process and store comprehensive data. However, we cannot process incomplete revenue or spending data. In such cases, the treasurer must first fill in the data. Previously. Within the edit and delete choices, the information system will securely store the modifications made by the treasurer to the revenue or spending data. After selecting the profile menu, the treasurer can access and manage their profile data.

4. Class Diagram

A class diagram is used to visualize the system structure by defining the classes that build it. Classes have properties, methods, and operations that are called [14].

Classes or tables in the mosque management information system consist of user tables, income tables, expenditure tables, inventory tables, and information tables. The class diagram can be seen in Figure 5.

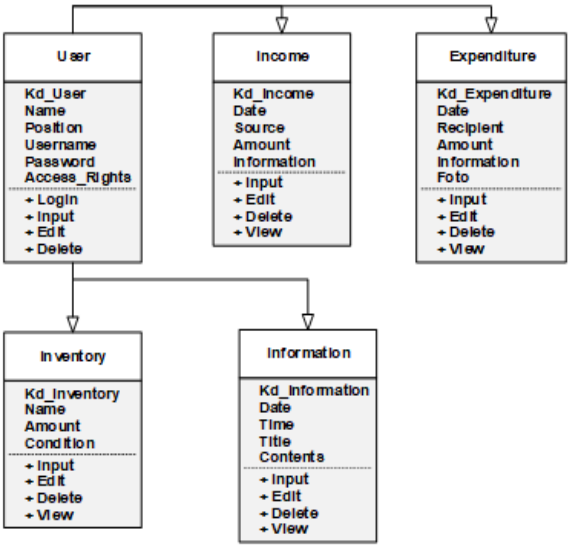


Fig.5. Class Diagram

Figure 5 illustrates the correlation between each table in the mobile-based mosque management information system. The structure of the user table includes the following columns: kd_user, name, position, username, password, and access rights. Five columns make up the income table: kd_income, date, source, amount, and information. The structure of the expenditure table includes the following columns: kd_expenditure, date, recipient, amount, and photo. The structure of the inventory database includes the following columns: kd_inventory, name, amount, and condition. The structure of the information table has the following columns: kd_information, date, time, title, and contents. The chairman has the authority to handle user data; the treasurer is responsible for managing income and spending data; the facilities and infrastructure department is in charge of processing inventory data; and the public relations department is responsible for managing information and schedules related to mosque activities.

5. Entity Relationship Diagram (ERD)

Entity Relationship Diagram (ERD) for modeling file structure is a software system object which is the elements of the designed software system, one of these objects is data. Data are raw facts, which must be documented by the system to create information [15]. An ERD image of the mosque management information system to be built can be seen in the picture 6.

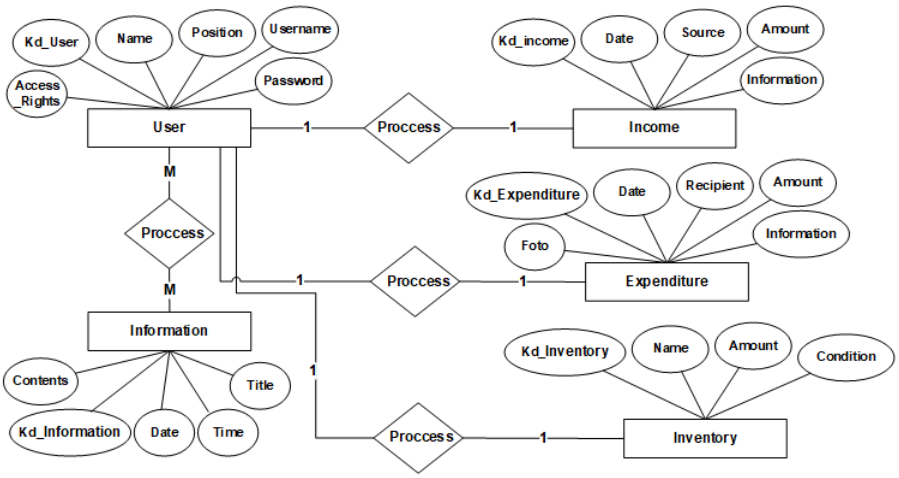


Fig.6. Entity Relationship Diagram (ERD)

Figure 6 is an entity relationship diagram (ERD) model which describes the structure of the management information system to be built. Management information systems consist of four main elements, namely users, income, expenditure, inventory, and information. Each element has its own attributes and interacts with one another. There are five different types of users, namely chairman, treasurer, facilities and infrastructure, public relations and public. Each can carry out different data processing processes. The treasurer can process income and expenditure data, the facilities and infrastructure sector can process inventory data, the public relations sector can process information and schedules. The chairman can monitor and see all data that has been processed by each section and all users can see information that has been processed by the public relations sector.

D. Implementation

1. Impelementation Information System

The page display of the information system built consists of the login page display, user data page display, inventory data page display, income data page display, expenditure data page display, financial report page display, and inventory report page display.

- Login Page Display

  Prior to accessing the information system, the user is required to log in by providing the designated username and password. Figure 7 depicts the visual representation of the page layout.

.

Fig.7. Login Page Display

- User Data Page Display

  On this page, the chairman can process user data. Figure 8 illustrates the page display.

Fig.8. User Data Page Display

- Inventory Data Page Display

  On this page, the infrastructure department can process mosque inventory data. Figure 9 illustrates the page display.

Fig.8. Inventory Data Page Display

- Display the Income Data Page
  On this page, the treasurer can carry out the processing of cash inflows. The page display can be seen in Figure 10.

Fig.10. Display of the Income Data Page

- Expenditure Data Page Display\
  On this page, the treasurer can carry out the processing of money disbursement. Figure 11 illustrates the page display.

Fig.11. Expenditure Data Page Display

- Financial Report Page Display
  Users can view processed financial reports by month on this page. Income and expenditure reports. Figure 12 illustrates the page display.

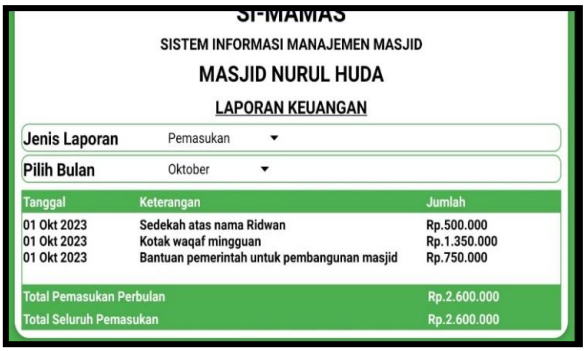


Fig.12. Financial Report Page Display

2. Information System Testing

We conduct information system testing to assess the program's functionality and identify any errors or issues. We tested this research using the black box method. Black box testing is a type of testing that focuses specifically on an information system's requirements [16], including user interface, performance, input, processing, and output. It also examines how information system features are used in the user's environment. This type of testing is followed by validation testing [17]. The black box testing results are shown in Table I.

TABLE I. TABLE STYLES

Testing Module	Testing Procedures	Input	Output	Conclusion
Login Page	-Run the Information System -Select the Login Menu -Enter the username "hanafi" and password "123456 -Select “Chairman” access rights -Click Login	Username "hanafi" and password "123456"	The chairman can enter the information system and select the available menu	Valid
User Data Page	-Run the Information System -Login -Select the user data menu -Enter complete user data -Click input	Complete user data	User data added successfully	Valid
Inventory Data Page	-Run the Information System -Login -Select the inventory data menu -Enter complete inventory data -Click input	Complete inventory data	Inventory data added successfully	Valid
Income Data Page	-Run the Information System -Login -Select the input data menu -Enter complete income data -Click input	Complete income data	Income data added successfully	Valid
Expenditure Data Page	-Run the Information System -Login -Select the expenditure data menu -Enter complete expenditure data -Click input	Complete expenditure data	Expenditure data added successfully	Valid
Information Page	-Run the Information System -Login -Select the information menu -Enter complete information -Click input	Complete information	Information added successfully	Valid

Furthermore, the information system undergoes testing on Android smartphones to evaluate its appearance, functionality, and efficiency when installed on various Android smartphone models. The Android device classification in this test is based on the user's operating system and brand of Android smartphone. It includes ten devices: three running Android version 10, three running Android version 11, three running Android version 12, and three running Android version 13. The outcomes of the testing conducted for the implementation of the information system are presented in Table II.

TABLE II. TABLE OF TESTING THE INFORMATION SYSTEM ANDROID DEVICE

Devices	Android Version	Information System Display	Information System Process	Information System Performance
Xiaomi Mi A2	Version 10	Successful	Successful	Successful
Realme X2 Pro	Version 10	Successful	Successful	Successful
vivo V15 Pro	Version 10	Successful	Successful	Successful
Samsung Galaxy S10	Version 11	Successful	Successful	Successful
Realme 6 Pro	Version 11	Successful	Successful	Successful
Vivo V17 Pro	Version 11	Successful	Successful	Successful
Xiaomi Mi 11X Pro	Version 12	Successful	Successful	Successful
Oppo Find X3 Pro 5G	Version 12	Successful	Successful	Successful
Realmi 8i	Version 12	Successful	Successful	Successful
Xiaomi 12S Pro	Version 13	Successful	Successful	Successful
Vivo X80 Pro	Version 13	Successful	Successful	Successful
Realmi 9i	Version 13	Successful	Successful	Successful

IV. CONCLUSIONS

After conducting this research, the author draws conclusions from the research that has been carried out regarding the design and development of a mobile-based Nurul Huda Mosque management information system, namely: The information system built can help the chairman of the Nurul Huda Mosque management see data reports that have been processed by the mosque management, such as financial reports and mosque inventory reports, the information system built can help the treasurer of the Nurul Huda Mosque to process financial data, consisting of income and expenditure, the information system built can help the facilities and infrastructure sector of the Nurul Huda Mosque to process mosque inventory data, the information system built can help the public relations sector of the Nurul Huda Mosque to convey information related to mosque activities or other information to the public, the developed information system facilitates public access to information disseminated by the administrators of Nurul Huda Mosque. An Android smartphone can access this technologically advanced data processing system, which replaces the mosque's manual data processing system. We specifically designed the information system to meet the specifications of Android smartphones commonly used by most individuals, ensuring a seamless installation and utilization experience. Android smartphones can access this technologically advanced data processing system, which replaces the mosque's manual data processing system. We specifically designed the information system to meet the specifications of Android smartphones commonly used by most individuals, ensuring a smooth installation and utilization experience. Based on the results of testing the use of a mobile-based mosque management information system that has been built by users including the chairman, treasurer, facilities and infrastructure, public relation and the public. Conclusions that can be drawn from before and after using the management information system by the user. The conclusion results can be seen in table III.

TABLE III. BEFORE AND AFTER USING THE APPLICATION

Indicator	Before using the app	After using the app
Financial data processing time	The processing time for financial data is longer because it is done manually by writing it in a book.	Financial data processing time is faster because it is done automatically by the information system.
Inventory data processing time	The inventory data processing time is longer because it is done manually by checking the inventory data one by one and then writing it in a book.	Inventory data processing time is faster because it is done automatically by the information system.
Time and scope of the process of conveying information to the public	The process of conveying information is longer, and the scope is limited.	The process of conveying information to the public becomes faster and the scope of information is wider because it uses an online-based information system
Time required for the creation of a report.	The report creation process takes longer because it is done manually by viewing and writing data one by one	The report creation process is faster because it is done automatically by the information system

REFERENCES

[1] S. Mannuhung, A. M. Tenrigau, and D. D., "Manajemen Pengelolaan Masjid dan Remaja Masjid Di Kota Palopo," *To Maega J. Pengabdi. Masy.*, vol. 1, no. 1, pp. 14–21, 2018.

[2] Elta Andea, Yuliantoro, and A. Fikri, "Masjid Jami' Masjid Bersejarah Di Kecamatan Pangean Kabupaten Kuantan Singingi," *Innov. J. Soc. Sci. Res.*, vol. 1, no. 2, pp. 423–427, 2021.

[3] M. R. Taufik Hidayat, "Optimalisasi Fungsi Manajemen Dalam Meningkatkan Peran Dan Fungsi Masjid," *Tadbir J. Manaj. Dakwah*, vol. 5, no. 3, pp. 285–304, 2020, doi: 10.15575/tadbir.v5i3.2106.

[4] E. Budhy, R. Dewi, and H. F. Negara, "Sistem Informasi Manajemen Masjid Berbasis Website (Studi Kasus : Masjid Baitul Ikhwan)," *Tekinfo J. Bid. Tek. Ind. dan Tek. Inform.*, vol. 22, no. 2, pp. 15–23, 2021, [Online]. Available: https://journals.upi-yai.ac.id/index.php/TEKINFO/article/view/1740/1453.

[5] Y. Suherman and E. N. Azandra, "Sistem Informasi Manajemen Masjid Berbasis Web," *J. J-Click*, vol. 8, no. 1, pp. 32–41, 2021.

[6] W. Wagino, "Sistem Informasi Manajemen Pengelolaan Masjid Raya Sabilal Muhtadin Banjarmasin," *Technologia*, vol. 11, no. 4, pp. 248–251, 2020, [Online]. Available: https://ojs.uniska-bjm.ac.id/index.php/JIT/article/view/4775/2930.

[7] D. Setiawan Putra and A. Fauzijah, "Perancangan Aplikasi Presensi Dosen Realtime Dengan Metode Rapid Application Development (RAD) Menggunakan Fingerprint Berbasis Web," *J. Inform. J. Pengemb. IT*, vol. 3, no. 2, pp. 167–171, 2018, doi: 10.30591/jpit.v3i2.836.

[8] A. Fauzi and E. Harli, "Peningkatan Kualitas Pelayanan Melalui CRM dengan Metode RAD," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 1, pp. 76–81, 2017, doi: 10.29207/resti.v1i1.16.

[9] A. A. Effendy and D. Sunarsi, "Persepsi Mahasiswa Terhadap Kemampuan Dalam Mendirikan UMKM Dan Efektivitas Promosi Melalui Online Di Kota Tangerang Selatan," *J. Ilm. MEA (Manajemen, Ekon. dan Akuntansi)*, vol. 4, no. 3, pp. 702–714, 2020, [Online]. Available: http://journal.stiemb.ac.id/index.php/mea/article/view/571/248.

[10] T. B. Kurniawan, "Perancangan Sistem Aplikasi Pemesanan Makanan dan Minuman Pada Cafetaria NO Caffe di TAnjung Balai Karimun Menggunakan Bahasa Pemrograman PHP dan My.SQL," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2020.

[11] N. Rizkita, E. Rosely, and H. Nugroho, "Aplikasi Pendaftaran dan Transaksi Pasien Klinik Hewan di Bandung Berbasis Web," in *e-Proceeding of Applied Science*, 2018, vol. 4, no. 3, pp. 1512–1520.

[12] S. Wahyudi, "PENGEMBANGAN SISTEM INFORMASI KLINIK BERBASIS WEB (Studi Kasus : Klinik Surya Medika Pasir Pengaraian)," *Riau J. ofComputer Sci.*, vol. 06, no. 01, pp. 50–57, 2020, [Online]. Available: http://e-journal.upp.ac.id/index.php/RJOCS/article/view/1979.

[13] Mohammad Suhatsyah, Runny Agustianti, and Syarifuddin, "Perancangan Sistem Penjualan Pada Pt Digital Mandiri Jaya Di Tanjung Balai Karimun Dengan Menggunakan Php Dan Mysql," *Tikar*, vol. 2, no. 1, pp. 59–71, 2021.

[14] S. Julianto and S. Setiawan, "Perancangan Sistem Informasi Pemesanan Tiket Bus Pada Po. Handoyo Berbasis Online," *J. Intra-Tech*, vol. 3, no. 2, pp. 11–25, 2019, [Online]. Available: https://journal.amikmahaputra.ac.id/index.php/JIT/article/view/56/48.

[15] S. Munir, D. P. M. Santoso, and R. Rivandi Arfans, "Perancangan Aplikasi Absensi Pegawai Pada Pt Avia Jaya Indah," *PROSISKO J. Pengemb. Ris. dan Obs. Sist. Komput.*, vol. 10, no. 1, pp. 63–70, 2023, doi: 10.30656/prosisko.v10i1.5566.

[16] R. Cahyadi, A. Damayanti, and D. Aryadani, "Recurrent Neural Network (RNN) dengan Long Short Term Memory (LSTM) untuk Analisis Sentimen data instagram," *J. Inform. dan Komput.*, vol. 5, no. 1, pp. 1–9, 2020, [Online]. Available: https://ejournal.akakom.ac.id/index.php/jiko/article/view/407.

[17] R. A. Purba, "Application design to help predict market demand using the waterfall method," *Matrix J. Manaj. Teknol. dan Inform.*, vol. 11, no. 3, pp. 140–149, 2021, doi: 10.31940/matrix.v11i3.140-149.

ISSN 2085-4552

Using Convolutional Neural Network and Saliency Maps for Cirebon Batik Recognition

Yoga Aditiya [1], Marlinda Vasty Overbeek[2], Suwito Pomalingo[3]
[1,2,3]Informatics Study Program, Universitas Multimedia Nusantara, Tangerang, Indonesia
[1]yoga.aditiya@student.umn.ac.id, [2]marlinda.vasty@umn.ac.id, [3]suwito.pomalingo@umn.ac.id

Accepted 20 May 2025
Approved 28 May 2025

***Abstract*— Cirebon Batik is one of Indonesia's cultural heritages that has its own unique patterns and motifs, reflecting the cultural richness and history of its region of origin. This study aims to address the challenges in classifying the complex motifs of Cirebon Batik by implementing Convolutional Neural Network (CNN) and Saliency Map methods. The three main motifs used are Mega Mendung, Singa Barong, and Keratonan. The dataset was obtained from various online sources and processed using image augmentation techniques. CNN is used to recognize complex visual patterns, while Saliency Map highlights important areas in the image that influence the model's decision. The results show that the developed CNN model achieved an accuracy of 82%, precision of 83%, recall of 82%, and F1-score of 82%. The use of Saliency Map provides better interpretability and enhances the understanding of the classification process.**

Index Terms*— Batik Cirebon, classification, CNN, image recognition, Saliency Map.

I. INTRODUCTION

Cirebon Batik is a vital part of Indonesia's cultural her- itage, renowned for its intricate motifs like Mega Mendung, Singa Barong, Paksinaga Liman, Keratonan, and Pratan Keris. Preserving these unique designs has become increasingly im- portant amid globalization, which often dilutes local identities [1]. The digitization of Cirebon Batik motifs presents an opportunity for effective preservation and wider international recognition, though the complexity and diversity of these patterns present challenges in classification and recognition efforts [2]. Manual classification is inefficient and requires expertise, making it impractical for large-scale application [3]. However, the diversity and complexity of Cirebon Batik motifs, which are rich in detail and philosophy, pose their own challenges in classification and recognition efforts. There- fore, an innovative approach is needed to ensure that the preservation of these batik motifs can be done accurately and efficiently, which will ultimately support the sustainability of this culture in the modern era.

Artificial intelligence (AI), particularly deep learning, has shown promise in addressing these challenges. Convolutional Neural Networks (CNN) can extract complex visual patterns and are well-suited for batik motif classification due to their ability to handle intricate designs. Prior studies, such as those by Wulandari et al. and Azzalini et al., demonstrate CNN's effectiveness in textile and motif recognition, suggesting its relevance for Cirebon Batik [4][5]. The addition of Saliency Maps, which highlight important image areas, provides further insights into model decisions, improving interpretability and aiding in motif preservation efforts [6][8].

This research explores the use of CNN combined with Saliency Maps to enhance the classification accuracy of Cire- bon Batik motifs while promoting cultural preservation. By increasing transparency and interpretability, this approach can support the development of a more informative database of batik patterns and facilitate their digital documentation. The integration of advanced AI technologies aims to contribute not only to cultural preservation but also to advancing pattern recognition methods in the context of Indonesia's rich artistic heritage [9]–[11].

II. METHODOLOGY

Figure 1 presents the entire process of program development from start to finish for the Cirebon batik classification task. It begins with the preparation of the Cirebon batik dataset, which is split into training and validation data. The first stage involves training a Convolutional Neural Network (CNN) with three layers using the training data, followed by validation to evaluate the model's performance. The saliency map is then generated based on the trained model to highlight key features that influence the classification. After obtaining the saliency maps, they are applied to the Cirebon batik dataset, creating a new version of the dataset layered with additional visual insights. This enhanced dataset is again split into training and validation sets and reintroduced into the CNN model for further training and validation. The final step is

an evaluation report that summarizes the performance and accuracy of the model after this entire process. This framework ensures that the model not only improves in accuracy but also enhances interpretability by incorporating saliency maps into the training workflow.

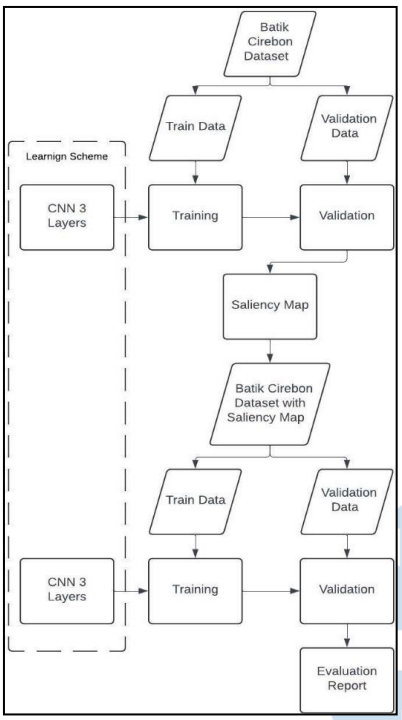


Fig. 1. Framework Full Program

A. Data Gathering

In this research, the dataset used is images of Cirebon batik which consists of three batik motifs. This dataset is obtained from various sources on the internet. The three batik motifs are Mega Mendung motif, Singa Barong motif, Keratonan motif.

B. Pre-processing

At this stage, data will be uploaded to the system that is grouped based on batik motifs. This data will be added based on the augmentation results using "ImageDataGenerator". With "imageDataGenerator" the uploaded data will be rescalled to change the pixel value from the 0-255 range to the 0-1 range. Figure 2 shows the flow of pre-processing.

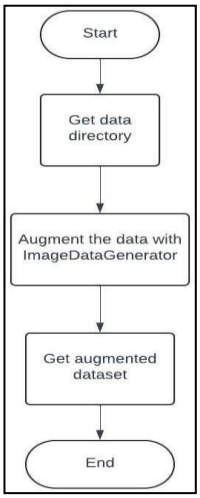


Fig. 2. Flowchart pre-processing

C. Split Dataset

At this stage, we used a split of 75% for training and 25% for validation. A separate test set was not used. We have clarified the revised manuscript under the split dataset detection. Figure 3 shows the flow of the split dataset.

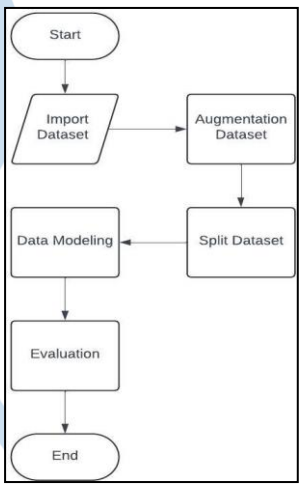


Fig. 3. Flowchart split dataset

D. Model Development

Convolutional Neural Network (CNN) is a type of artificial neural network designed to process grid-like data, such as images. CNN is claimed to be the best model for solving problems in object recognition [12], [13]. CNN is developed based on the Multilayer Perceptron (MLP) which is designed to process two-dimensional data. CNNs consist of several types of main layers, namely Convolutional layer, the Pooling layer, and the Fully Connected layer. [14][18]

At this stage, we will build the model gradually from layer to layer. Three layers are built using ReLu activation with 32, 64, and 128 filters each with a kernel size of 3x3. From each convolution layer, the 'max pooling' feature will be used to reduce the number of

parameters generated. After three layer runs, a two-dimensional matrix is converted into a one-dimensional vector that will be input into the dense layer. Before entering the dense layer, the dropout layer is added with a dropout rate of 20% (rate = 0.2). This dropout serves as one of the effective regularization methods to prevent overfitting. In the dropout process, a random number of neurons will be disabled during the training process, so that the model is less dependent on certain neurons and is able to produce a more generalized model. The dropout is applied before the first dense layer, right after the flattening process, to maximize the regularization effect on the dense layers. There are two dense layers used, the first layer uses 64 neurons with ReLu activation. The second layer uses three neurons corresponding to the number of classes to be classified with Softmax activation. Figure 4 shows the flow of model building.

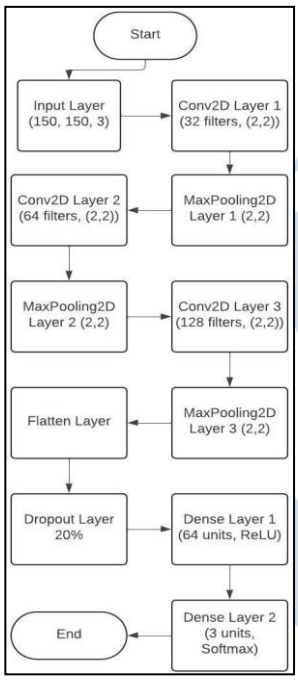


Fig. 4. Flowchart model

E. Model Fit and Callbacks

At this stage, the model that has been built will be tested based on the validation data against the training data. The initial repetition is determined as 30 repetitions with 2 call- backs provided. The 'red_lr' callback is used to reduce the learning rate of the model by three times. my_callback is used to prevent overfit by stopping the repetition when it has reached an accuracy value of 90%. Figure 5 shows the flow of the model and callbacks.

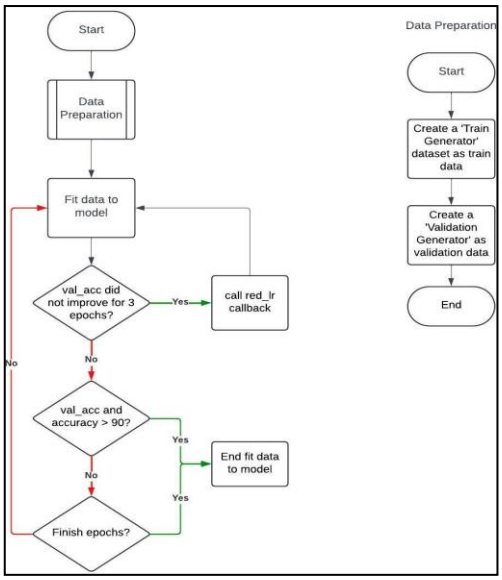


Fig. 5. Flowchart model fit and callbacks

F. Saliency Map

Saliency map is an important concept in pattern recognition and computer vision used to identify the most significant areas in an image that influence the decision of a model. The basic theory of saliency maps is rooted in how the human visual system works, where our brain automatically focuses on parts that stand out or attract attention in a scene. In the context of machine learning and artificial neural networks, saliency maps serve to provide a visual interpretation of the model's decision- making process, by highlighting the pixels in the image that most influence the classification or detection results. Some of the main methods in saliency map generation are gradient, Class Activation Maps (CAM), and Grad-CAM.

At this stage, parts of the batik that most affect the results of the model prediction will be displayed. The part will be displayed with a different color placed above the original image. Figure 6 shows the flow of the saliency map.

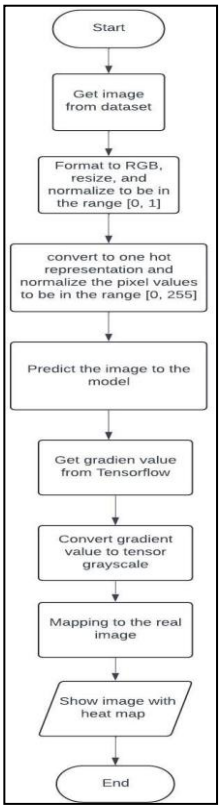


Fig. 6. Flowchart saliency map

G. Modeling with Saliency Map Data

The Cirebon batik dataset, enhanced with saliency maps, will be reused as a new dataset to be fed into the pre- viously constructed model. This dataset, now layered with visual insights from the saliency maps, will undergo the same 75:25 split, where 75% is allocated for training and 25% for validation. The training process will follow the steps outlined in Section D, using the same CNN architecture with convo- lutional, pooling, and fully connected layers. The model will be trained similarly as described in Section E. This approach aims to refine the model's ability to recognize key features of the batik motifs, leveraging the saliency maps to focus on the most relevant patterns, and ensuring both improved accuracy and interpretability in the classification results.

H. Evaluation

Evaluation is done to determine the success rate of the model that has been built. The model will be tested by comparing the training data with the testing data based on several parameters, namely, accuracy, precision, recall, and F1 value calculated based on the confusion matrix. At this stage, an evaluation is carried out using a confusion the actual class. In addition, the accuracy, precision, recall and F1 Score values. Figure 7 also shown flow of the evaluation result matrix to determine the class distribution of the classification.

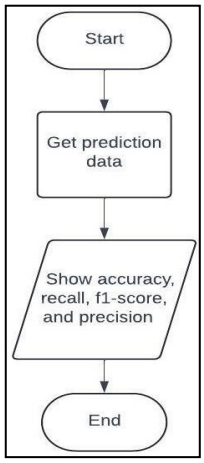


Fig. 7. Flowchart evaluation

III. Result and Discussion

A. Training Model

The CNN model was trained using Cirebon Batik motif data stored in a specific directory. First, the data was generated and resized to 150x150 pixels. The training data is taken from the training data subset, while the validation data is taken from the validation subset. After the data generator was created, the model was trained using the fit function for 100 epochs. Table I shows the last 10 repetitions performed in model in the last repetition is 91.21% and the validation accuracy of the model in the last repetition is 84.51%. training. It can be seen that the training accuracy of the model.

TABLE I
CNN Model Training Results for Epochs 90 to 100

Epoch	Accuracy	Loss	Val Accuracy	Val Loss
90	0.9680	0.1821	0.8592	0.3755
91	0.9842	0.1397	0.8873	0.3509
92	0.9714	0.1206	0.8732	0.3546
93	0.9757	0.1286	0.8521	0.3903
94	0.9657	0.1695	0.8732	0.3917
95	0.9768	0.1465	0.8310	0.3947
96	0.9479	0.1694	0.7887	0.5655
97	0.9396	0.1933	0.8310	0.3618
98	0.9692	0.1180	0.8521	0.4575
99	0.8961	0.2813	0.8099	0.5629
100	0.9121	0.2514	0.8451	0.4232

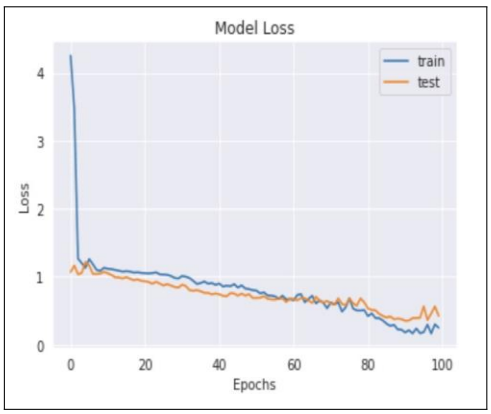


Fig. 8 Loss Graph

In Figure 8, a graph of the loss value of the trained model is shown. The value of loss in the training data has decreased significantly from each epoch, starting from 7 and continuing to decrease until it reaches about 0.02. Meanwhile, the loss value on the validation data also shows a consistent decrease, starting from around 1 until it stabilizes at 0.68. This graph indicates that the model is getting better at minimizing prediction errors in both training and validation data as the number of epochs increases.

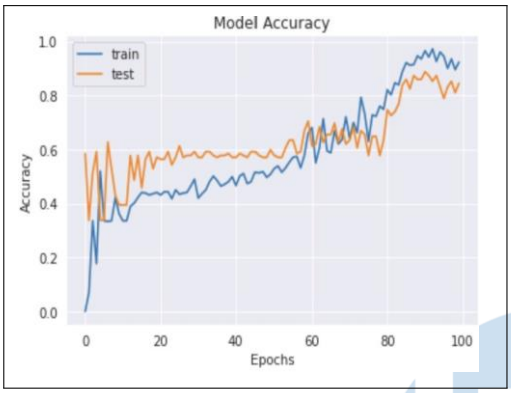


Fig. 9. Accuracy graph

In Figure 9, a graph of the model accuracy value during the training process is shown. The model accuracy on the training data shows a consistent increase as epochs increase, starting from 0.2 until it reaches around 0.98, indicating that the model is able to learn well on the training data. Accuracy on validation data also increased gradually, starting from around 0.4 until it stabilized at around 0.8. However, fluctuations were seen at some points, indicating that the model had some difficulty in maintaining a stable accuracy on the validation data. Overall, however, the model performed well on both training and validation data.

B. Prediction Result

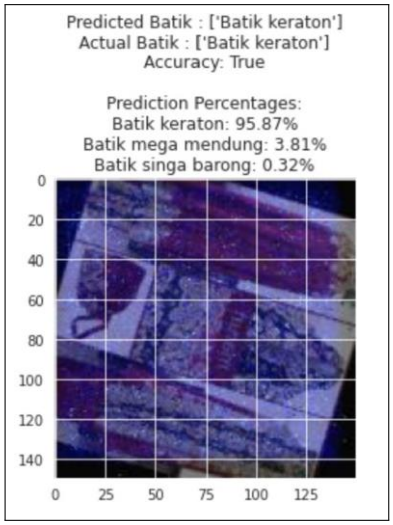


Fig. 10. Prediction result of model

Figure 10 shows an example of predicted data based on the trained model. The model's prediction is "Batik keraton", which also matches the actual data, indicating that the prediction is accurate. The accuracy of the prediction is supported by the model's confidence percentage of the classification, where "Batik keraton" has the highest probability of 85.87%. In addition, the model also gives a small percentage probability for two other motifs, namely "Batik mega mendung" at 3.81% and "Batik singa barong" at 0.32%

C. Saliency Result

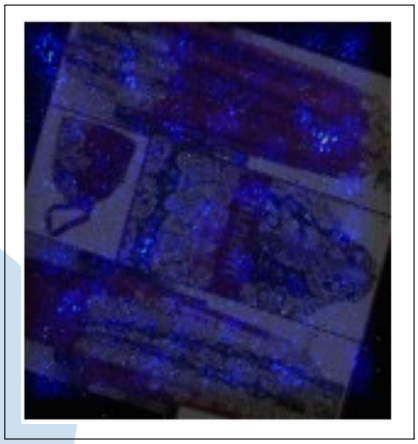


Fig. 11. Saliency result

Figure 11 shows the visualization of batik that has been highlighted by the Saliency Map so that it can be seen which part of the image has the most influence on the model. The part covered in blue indicates the part that is considered important by the model, the brighter the blue color in a part of the image indicates that the part is more important to the model.

D. Evaluation Result

Table II shows an accuracy value of 87%, a weighted precision value of 89%, a weighted recall value of 88%, and a weighted f1-score of 87%.

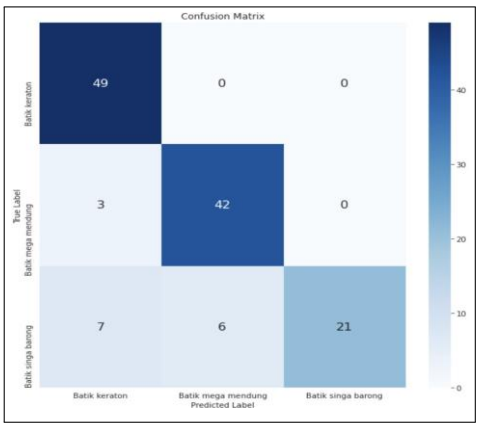


Fig. 12. Confusion matrix

TABLE II
EVALUATION RESULT

Metric	Value
Accuracy	0.87
Precision	0.89
Recall	0.88
F1-Score	0.87

Figure 12 shows the result of the prediction distribution performed by the model. The True Labels (y-axis) represents the original categories of the Batik motifs. The Predicted Labels (x-axis) represents the predicted categories from the model. In True Label batik Keraton, there are 49 samples that correctly classified as "Batik Keraton" and there are no samples that misclassified. In True Label batik Mega Mendung there are 42 samples that correctly classified as "Batik Mega Mendung" and there are 3 samples that misclassified as "Batik Keraton". In True Label batik Singa Barong, there are 21 samples that correctly classified as "Batik Singa Barong" and there are 7 samples that misclassified as "Batik Keraton" with 6 samples that missclassified as "Batik Mega Mendung".

IV. CONCLUSIONS

Convolutional Neural Network (CNN) and Saliency Map algorithms have been successfully applied in the classification of Cirebon Batik motifs with satisfactory results. The build model shows a good performance with 87% accuracy, 89% precision, 88% recall, and 87% f1-score, reflecting the balance between accuracy and the model's ability to detect the correct motif. In addition, the Saliency Map successfully highlights the parts of the image that have the most influence on the classification decision, providing deeper insight into how the model recognizes and distinguishes batik motifs.

ACKNOWLEDGMENT

The authors would like to thank to Universitas Multimedia Nusantara for funding supporting this research, cultural actors in Batik Pesisir Cirebon in Cirebon Batik UKM, and Universitas Muhammadiyah Cirebon as partner in this research.

REFERENCES

[1]. R. Susanti and D. Purnomo, "Batik sebagai warisan budaya tak benda: Tantangan dan peluang pelestarian di era digital," Jurnal Kebudayaan Indonesia, vol. 12, no. 2, pp. 55–67, 2019.
[2]. Rahman and H. Wicaksono, "Digitalisasi budaya: Strategi pelestarian motif batik melalui teknologi informasi," Jurnal Teknologi dan Infor- masi, vol. 18, no. 1, pp. 43–59, 2020.
[3]. Kompas, "Batik cirebon: Mempertahankan tradisi di tengah modernisasi," 2022, https://www.kompas.com/batik-cirebon- mempertahankan-tradisi.
[4]. R. Wulandari et al., "Batik image classification using convolutional neural network," International Journal of Electrical and Computer Engineering, vol. 10, no. 1, pp. 217–225, 2020.
[5]. M. Azzalini et al., "Textile pattern recognition using convolutional neural networks: A case study on traditional textiles," Journal of Computational Design and Engineering, vol. 8, no. 2, pp. 207–215, 2021.
[6]. K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep inside convolutional networks: Visualising image classification models and saliency maps," arXiv preprint arXiv:1312.6034, 2013. [Online]. Available: https://arxiv.org/abs/1312.6034
[7]. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolu- tional networks," in European Conference on Computer Vision (ECCV), 2014, retrieved from ECCV.
[8]. Kompas.com, "Digitalisasi batik cirebon dorong pelestarian budaya dan industri kreatif," 2023, https://www.kompas.com.
[9]. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, retrieved from CVPR.
[10]. Sutaryo, "Batikan: Pelestarian batik cirebon dengan teknologi mod- ern," 2023, retrieved from https://www.sutaryo.co.id/batikan-pelestarian- batik-cirebon.
[11]. M. Hadi, "Penerapan teknologi digital dalam pelestarian budaya batik," Jurnal Kebudayaan Indonesia, vol. 17, no. 2, pp. 145–160, 2022.
[12]. Y. Gultom, A. M. Arymurthy, and R. J. Masikome, "Batik classifica- tion using deep convolutional network transfer learning," Jurnal Ilmu Komputer Dan Informasi, vol. 11, no. 2, p. 59, 2018.
[13]. S. F. Tumewu, D. H. Setiabud, and I. Sugiarto, "Klasifikasi motif batik menggunakan metode deep convolutional neural network dengan data augmentation," Jurnal Infra, vol. 8, no. 2, pp. 189–194, 2020.
[14]. R. Refianti, A. B. Mutiara, and R. P. Priyandini, "Classification of melanoma skin cancer using convolutional neural network," Interna- tional Journal of Advanced Computer Science and Applications, vol. 10, no. 3, pp. 409–417, 2019.
[15]. M. Alexander, "Sistem pengenalan wajah untuk pre- sensi dengan algoritma convolutional neural network," https://kc.umn.ac.id/id/eprint/13923, 2020, accessed: 2024-08-09.
[16]. T. A. Hartanto and S. a. Hansun, "Comparative analysis of pre-trained cnn models on retinal diseases classification," international journal of industrial Engineering Production Research, vol. 35, no. 3, 2024. [Online]. Available: http://ijiepr.iust.ac.ir/article-1-1935-en.html
[17]. C. Young, "Applicability of various pre-trained deep convolutional neural networks for pneumonia classification based on x-ray images," International Journal of Advanced Trends in Computer Science and Engineering, vol. 9, 2020.
[18]. Archilles, A., & Wicaksana, A. (2020). Vision: a web service for face recognition using convolutional network. TELKOMNIKA (Telecommunication Computing Electronics and Control), 18(3), 1389-1.

ISSN 2085-4552

Application of Fuzzy AHP-TOPSİS Hybrid Method in Facility Location Selection for Software Systems

K.R. Aliyeva
Azerbaijan State Oil and Industry University, 20 Azadlig Ave., AZ1010, Baku, Azerbaijan
kamala.aliyeva@asoiu.edu.az

Accepted 09 January 2025
Approved 09 June 2025

***Abstract*— Facility location is an integral part of the strategic planning process of almost every organization. Selecting the right location for software systems facilities involves considering various factors to ensure optimal performance, reliability, and cost-effectiveness. For business success, and competitive advantage there are some critical factors that very highly affect facility location. They are proximity to customers, infrastructure, labor quality, total cost, suppliers, etc. The criteria for selecting a facility location may be vaguely defined or open to interpretation. External factors such as economic conditions, political stability, and environmental risks may introduce vagueness and unpredictability into facility location decisions. In this paper we apply fuzzy AHP-TOPSIS hybrid method for facility location in software systems. Fuzzy AHP (Analytic Hierarchy Process) and fuzzy TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) are both decision-making methods commonly used in facility location selection. Fuzzy AHP is particularly useful in situations where decision criteria are subjective and uncertain, providing a more robust framework for making well-informed decisions. Fuzzy TOPSIS is useful in complex decision-making where uncertainty and subjectivity play a significant role, offering a flexible and comprehensive approach to evaluating and ranking alternatives. In the first part of the facility location selection process, we use fuzzy AHP method for determining weights of criteria that are important in selection process. Then by using fuzzy TOPSIS we rank alternatives and select appropriate location for facility. Selection of the best location for software systems provided according to three attributes and three alternatives (A, B, C). Relative closeness of each alternative to the ideal solution represents that alternative C is the best alternative.**

Index Terms*— *Fuzzy numbers, facility location selection, software systems, ideal solution, fuzzy AHP-TOPSIS.

I. INTRODUCTION

Facility location typically refers to the problem of selecting the location of facilities (for example, warehouses, factories, offices, production centers, stores) based on various criteria, such as minimizing transportation costs or maximizing customer service [1]. Facility location in the context of software systems often refers to the strategic placement of data centers, servers, or other infrastructure to optimize performance, reliability, and cost-effectiveness. Key considerations for facility location in software systems are proximity to users, redundancy and disaster recovery, cost optimization, regulatory compliance, network connectivity, security, environmental considerations etc. Placing facilities closer to many users can reduce latency and improve response times. This is particularly important for real-time applications like online gaming or video streaming. Distributing facilities across different geographic locations helps ensure redundancy and disaster recovery. In the event of a natural disaster or network outage in one location, services can be quickly restored from another location. Considerations such as the cost of real estate, electricity, cooling, and labor vary by location. Choosing locations with lower operational costs can result in significant savings over time. Compliance with data protection regulations may require storing data in specific geographic regions. It's essential to choose locations that comply with relevant laws and regulations. Facilities should be in areas with robust network connectivity to ensure high-speed and reliable connections to the internet backbone and other networks. Choose locations that can accommodate future growth and scalability needs. This includes factors such as available space for expansion and access to skilled labor. Security considerations, including physical security and access controls, are crucial for protecting data and infrastructure. Facilities should be in areas with low crime rates and have appropriate security measures in place. Choosing locations with access to renewable energy sources or implementing energy-efficient technologies can help reduce environmental impact. Facility location decisions for software systems require careful consideration of technical, regulatory, financial, and environmental factors to ensure optimal performance, reliability, and

compliance. Some attributes are so important in selection of location that they control all decision process. These are capacity, successful labor climate, distances, accessibility, service, proximity to suppliers and resources. Capacity considerations are an important consideration in facility location decisions that directly impact operational efficiency and customer satisfaction. Successful labor climate plays a critical role in facility location decisions by influencing workforce availability, costs, stability, productivity, regulatory compliance, and community relations. Businesses that prioritize a favorable labor climate are better positioned to establish sustainable and successful operations in their chosen locations. Labor climate is a criterion of wages, training needs, regards to work, labor performance, and union strength. Proximity to customers or locating near customers is important when customers need technical support, products are voluminous, and shipping rates are high [2]. Proximity to suppliers and resources is a critical consideration in facility location decisions, impacting transportation costs, supply chain efficiency, quality control, collaboration, risk mitigation, and access to specialized skills. By strategically locating facilities close to suppliers and key resources, businesses can gain competitive advantages, improve operational performance, and enhance overall business resilience. These need permanent coordination and negotiation, which can become heavier as distance increases [3]. In selection process is especially important availability resources and minimal costs. Different facility location factors are determined in place choosing, including investment cost, availability of high-quality labor, shipping, infrastructure. and facility location thus obviously involves multiple criteria.

Uncertainty is a significant factor in facility location decisions due to several reasons. Market dynamics, economic conditions, consumer preferences, and market demand can fluctuate, leading to uncertainty in forecasting future sales volumes and distribution patterns. This uncertainty makes it challenging to determine the optimal location for facilities to meet evolving market demands effectively. It offers a way to model and reason about uncertain or ambiguous situations by capturing the inherent fuzziness in human reasoning and natural language [4]. Changes in regulations, zoning laws, and government policies can impact facility location decisions. Uncertainty regarding future regulatory requirements or restrictions may influence the suitability of certain locations and affect long-term investment decisions. The conditional method of approaching facility location problems like cost volume analysis that is s a managerial accounting technique used to examine the relationship between costs, volume, and profit within a business, factor rating method that is a decision-making tool utilized in place choosing or facility location analysis, and center of gravity method that is a quantitative methodology used in facility location analysis to define the optimal place for a facility, such as a warehouse, distribution center, or manufacturing plant are generally less effective at combating imprecision or vagueness in linguistic judgments [5]. In real life, data for assessing the accessibility of object locations for various subjective factors and factor weights are represented in linguistic terms [6]. For effectively resolve the vagueness that often arises in accessible information and eliminate the essential vagueness of human thinking and preferences, fuzzy set theory has been employed to identify uncertain multi-attribute decision-making problems [7]. Thus, the fuzzy AHP -TOPSIS hybrid method offered in this article for facility location selection problem in software system, where the ratings of various alternative locations under different subjective attributes and the weights of all attributes are introduced by fuzzy numbers [8]. The basis for decision-making is that the main decisions are made based on the results of AHP [9]. The decision-makers require to estimate alternatives always comprise vagueness. For modeling such vagueness in the facility location selection, fuzzy sets can be integrated with binary comparisons, for example using the AHP extension [10]. The fuzzy AHP technology permits a more precise specification of the decision-making process [11]. The fuzzy TOPSIS technology is a highly employed technique in decision-making for prioritization alternatives [12].

In this article, the fuzzy AHP-TOPSIS hybrid technique used for defining weights of importance of attributes, ranking alternatives, and determining the best location for software system [13-15]. Basic goal in facility location for software systems is to create a strategic infrastructure footprint that optimizes performance, reliability, and cost-effectiveness while ensuring compliance with regulations and minimizing environmental impact. This article is structured as follows. Section 2 represents the main steps of the fuzzy AHP-TOPSIS hybrid technique that is employed in facility location problem. Section 3 offered hybrid methodology with fuzzy numbers for the facility location problem. Section 4 represents discussion and Section 5 presents conclusions of this research.

II. METHODOLOGY

Fuzzy AHP-TOPSIS methodology is one of the largely used techniuqes of multi-attribute decision making and this hybrid decision-making method combines the Analytic Hierarchy Process methodology with the Technique for Order of Preference by Similarity to Ideal Solution method, while also incorporating fuzzy set theory. This method is especially useful when dealing with decision problems involving multiple criteria or attributes that are subjective, imprecise, or uncertain in nature. Basic functions of fuzzy AHP-TOPSIS hybrid methodology are hierarchy formation, fuzzy pairwise comparison, aggregation of weights, fuzzy normalization, determining fuzzy similarity to ideal solution. Hierarchy formation is identifying the decision hierarchy, which consists of the main objective, attributes, and alternatives. Break down the main objective into multiple criteria and further decompose each criterion into sub-criteria if necessary. Fuzzy AHP

use fuzzy pairwise comparison matrices to assess the comparative significancy or weights of attributes and sub-attributes. Decision-makers assign linguistic terms such as "weak", "strong", "very strong", "extremly" or fuzzy numbers to express the pairwise comparisons between criteria based on their perceived importance or preference. Aggregation of weights function is aggregating the fuzzy pairwise comparison matrices to calculate the overall weights of attributes and sub-attributes. Various aggregation methods, such as fuzzy geometric mean, fuzzy arithmetic mean, or fuzzy weighted average, can be used to compute the aggregated weights. Fuzzy normalization normalize the decision matrix for each criterion to convert linguistic assessments or fuzzy numbers into crisp values. Fuzzy normalization techniques, such as triangular or trapezoidal fuzzy numbers, can be employed to handle vagueness and imprecision in the data. Fuzzy TOPSIS technique determine fuzzy positive ideal solution and fuzzy negative ideal solution is determined [13]. Then, computed the fuzzy closeness coefficient or similarity score for each alternative relative to the fuzzy positive ideal solution and fuzzy negative ideal solution using fuzzy TOPSIS. Alternatives with higher similarity scores to the fuzzy positive ideal solution and lower similarity scores to the fuzzy negative ideal solution are considered more preferable. Ranking and sensitivity analysis is odering the alternatives on base of their fuzzy closeness coefficients to identify the most preferred alternative(s). Performance sensitivity analysis to assess the robustness of the rankings to changes in the criteria weights or input data and evaluate the stability of the decision. Fuzzy AHP-TOPSIS provides a structured and systematic approach to decision-making in complex and uncertain environments, allowing decision-makers to incorporate subjective judgments, imprecise data, and uncertainty into the decision process. By integrating fuzzy set theory with AHP and TOPSIS, this method enables more comprehensive and nuanced decision analysis, leading to more informed and robust decisions. Some of the superiorities of this hybrid technique are determining weights of criterias by calculating consistency ratio, rationality, comprehensibility, well computational efficiency and the ability to measure the relative performance of each alternative in a simple mathematical form. The basic steps in multi-attribute decision-making technique fuzzy AHP-TOPSIS are the following:

Step 1. Creating a fuzzy comparison matrix. The scale of linguistic terms is determined. The scale used is the triangular fuzzy numbers scale from one to nine and determined by the membership function that represented in Table I.

TABLE I. SCALE OF INTEREST

Scale of interest	Linguistic term	Member function
1	Equally important	(1,1,1)
3	Weakly important	(1,3,5)
5	Strongly more important	(3,5,7)
7	Very strongly important	(5,7,9)
9	Extremely important	(7,9,9)

Then, using the triangular fuzzy numbers to make pair-wise comparison matrix for the basic attribute and sub-attribute. The form of fuzzy pairwise comparison matrix represented in formula (1) [14].

$$\tilde{A} = \begin{bmatrix} 1 & \cdots & \tilde{a}_{1n} \\ \vdots & \ddots & \vdots \\ \tilde{a}_{1n} & \cdots & 1 \end{bmatrix} \quad (1)$$

Step 2. Determining fuzzy geometric mean. The fuzzy geometric mean calculated by using formula (2) [14]:

$$\tilde{r}_i = \left(\tilde{a}_{i1} \otimes \tilde{a}_{i2} \otimes \cdots \otimes \tilde{a}_{in}\right)^{1/n} \quad (2)$$

where $\tilde{a}_{in}$ is estimation of fuzzy comparison matrix from attributes i to n. The outcome of the fuzzy geometric mean will be later to called local fuzzy number.

Step 3. Determining the fuzzy weight for each attribute. Calculate the global fuzzy number for each evaluated attribute with formula (3).

$$\tilde{w}_i = \tilde{r}_1 \otimes \left(\tilde{r}_1 \oplus \tilde{r}_2 \oplus \cdots \tilde{r}_n\right)^{-1} = \left(lw_i, mw_i, uw_i\right) \quad (3)$$

Step 4. Determining the best non fuzzy performance. The global fuzzy number transformed to crisp weight value using the sentry of area method to find the value of best non fuzzy performance (BNP) from the fuzzy weight in each attribute, determined using formula (4).

Step 5. Normalization decision matrix provided by using formula (4) [14].

$$r_{ij} = \frac{a_{ij}}{\sqrt{\sum_{k=1}^{m} a_{kj}^2}} \quad (4)$$

Step 6. Formulated weighted normalized decision matrix by using equation (5) [14].

$$V_{ij} = \begin{bmatrix} w_1 r_{11} & w_2 r_{12} & \cdots & w_n r_{1n} \\ w_1 r_{21} & w_2 r_{22} & \cdots & w_n r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_1 r_{m1} & w_2 r_{m2} & \cdots & w_n r_{mn} \end{bmatrix} \quad (5)$$

Step 7. Determine fuzzy positive ideal solution A^* and fuzzy negative ideal solution A^-. The positive ideal solution technique represents that each evaluated factor has a monotonically increasing or decreasing characteristic. For determination of the fuzzy positive ideal solution set used equation (6) [15]:

$$A^* = \left\{\max v_{ij} \mid j \in J\right\}, \left\{\min v_{ij} \mid j \in J'\right\}$$
$$A^* = \left\{v_1^*, v_2^*, \cdots, v_n^*\right\} \quad (6)$$

Fuzzy negative ideal solutions set selects the smallest of the column values in the matrix. For determination of the fuzzy negative ideal solution set used equation (7) [15]:

$$A^- = \left\{\min v_{ij} \mid j \in J\right\}, \left\{\max v_{ij} \mid j \in J'\right\}$$
$$A^- = \left\{v_1^-, v_2^-, \cdots, v_n^-\right\} \quad (7)$$

Step 8. Calculation positive ideal and negative ideal the seperation measures.

Positive ideal separation measure is calculated by using formula (8).

$$S_i^* = \sqrt{\sum_{j=1}^{n}\left(v_{ij} - v_j^*\right)^2} \quad (8)$$

Negative ideal separation measure is calculated by using formula (9) [16].

$$S_i^- = \sqrt{\sum_{j=1}^{n}\left(v_{ij} - v_j^-\right)^2} \quad (9)$$

Step 9. Calculation the relative closeness to the positive ideal solution [16].

$$C_i^* = \frac{S_i^-}{S_i^- + S_i^*} \quad (10)$$

Step 10. Ranking alternatives.

III. Result and Discussion

Selecting the appropriate facility for software systems involves choosing the infrastructure or platform where the software will be developed, deployed, and maintained. There are different functions for facility selection process in the context of software systems [17].

- Defining requirements. Clearly defining the requirements and objectives of the software system. Consider factors such as scalability, performance, security, reliability, compliance, and integration needs. Determine any specific infrastructure or platform requirements based on the nature of the software and its intended use.
- Evaluating hosting options. Assess different hosting options for the software system, such as on-premises, cloud, or hybrid solutions. Evaluate the advantages and disadvantages of each option in terms of cost, scalability, flexibility, reliability, security, and maintenance requirements.
- Considering cloud providers. If opting for a cloud-based solution, evaluate various cloud service providers based on factors such as pricing, services offered, performance, availability, security, compliance certifications, and geographical coverage.
- Assessing infrastructure requirements. Determine the infrastructure requirements for the software system, including computing resources, networking, databases, development tools, and middleware. Consider whether specialized hardware or software components are needed to support specific functionalities or performance requirements.
- Evaluating data storage and management. Assess requirements for data storage, management, and backup. Consider factors such as data volumes, access patterns, data consistency, latency, replication, disaster recovery, and compliance with data protection regulations.
- Considering development environment. Evaluate options for the development environment, including programming languages, frameworks, extension tools, control systems, continuous integration/continuous deployment pipelines, and collaboration platforms. Choose tools and technologies that align with the skills and preferences of the development team and support efficient software development practices.
- Assessing security and compliance. Ensure that the selected facility meets security and compliance requirements for the software system. Consider factors such as data encryption, access controls, identity management, audit logging, vulnerability management, and regulatory compliance.
- Evaluating support and maintenance. Consider support and maintenance requirements for the software system, including monitoring, troubleshooting, patch management, upgrades, backups, and technical support. Evaluate the availability of support services from vendors or service providers and assess their responsiveness and expertise.
- Cost analysis. Conduct a comprehensive cost analysis of different facility options, taking into account upfront costs, ongoing operational expenses, licensing fees, subscription costs, and potential cost savings or cost avoidance associated with each option.
- Risk assessment. Identify and assess potential risks associated with each facility option, such as vendor lock-in, service outages, security breaches, data loss, regulatory non-compliance, and changes in business requirements or market conditions. Develop mitigation strategies to address identified risks.
- Finalizing decision. Based on the evaluation criteria, prioritize the facility options and make a final decision on the most suitable facility for the software system. Document the rationale behind the decision and communicate it to stakeholders involved in the software development and deployment process. By following this systematic approach, organizations can make informed decisions when selecting facilities for software systems, ensuring that the chosen facility meets the requirements of the software system and supports its successful development, deployment, and maintenance. There may be a lack of comprehensive data or uncertainty surrounding key factors such as demand patterns, infrastructure availability, regulatory requirements, making it challenging to make well-informed decisions. Stakeholders may have different preferences, priorities, and risk tolerances when it comes to facility location selection. Vague or conflicting preferences can lead to uncertainty and disagreement in decision-making. The business and technological landscape is constantly evolving, introducing uncertainty and vagueness into facility location decisions. Factors such as market trends,

technological advancements, and regulatory changes can impact the suitability of a chosen location over time.

Assume that a multi criteria decision making problem of facility location selection for software systems involves three criteria - C_1, C_2, C_3, (C_1 - Labor climate, C_2 - Proximity to customers, C_3 -Proximity to suppliers and resources) and 3 alternatives - A, B, C. For determining best location of software systems integrated fuzzy AHP-TOPSIS approach is used. Graphical structure of this methodology given in fig 1.

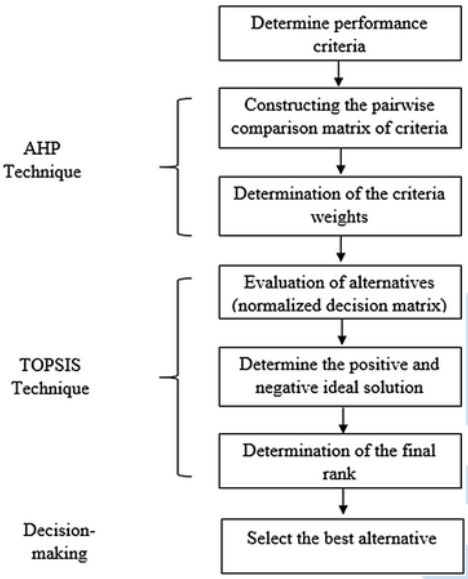


Fig 1. Graphical structure of fuzzy AHP-TOPSIS hybrid methodology

Step 1. Determining labor climate, proximity to customers and proximity to suppliers and resources as criteria for facility location selection [18]. Labor climate criteria in software systems refer to the conditions, environment, and factors that affect the morale, productivity, and satisfaction of software development teams. By considering and actively managing these labor climate criteria, organizations can create an environment that supports the well-being and productivity of their software development teams, ultimately leading to better outcomes for projects and the organization as a whole.

These criteria play a crucial role in determining the overall success of software projects. Proximity to customers in software systems refers to the degree to which software development teams interact, understand, and collaborate with end-users or customers throughout the development process. This criterion is essential for building successful software products that meet the needs and expectations of customers effectively. By prioritizing proximity to customers in software systems, organizations can build products that are not only technically robust but also resonate with users, drive adoption, and ultimately contribute to the success of the business. Proximity to suppliers and resources in software systems refers to the accessibility and availability of necessary resources, tools, and external partners that support the software development process. By effectively managing proximity to suppliers and resources in software systems, organizations can enhance their agility, scalability, and competitiveness in delivering high-quality software solutions that meet customer needs and expectations.

Step 2. Constructing the pairwise comparison matrix of criteria. For determining weights of criteria we use fuzzy triangle numbers. Using scale of interest that represented in Table I we construct pairwise comparison matrix that represented in Table II.

TABLE II. PAİRWİSE COMPARİSON MATRİX OF CRİTERİA

	C_1	C_2	C_3
C_1	(1,1,1)	(1,3,5)	(1/7,1/5,1/3)
C_2	(1/5,1/3,1)	(1,1,1)	(1/9,1/7,1/5)
C_3	(3,5,7)	(5,7,9)	(1,1,1)

Step 3. Determining weights of each criteria. The geometric mean of fuzzy comparison values of each attribute is determined by using formula (2). $\tilde{r}_i$ - geometric mean of fuzzy comparison values of "Labor climate" criterion calculated as down.

$$\tilde{r}_i = \left(\prod_{j=1}^{n} \tilde{d}_{ij}\right)^{1/n} = \left[\left(1*1*1/7\right)^{\frac{1}{3}}; \left(1*1/3*1/5\right)^{\frac{1}{3}}; \left(1*5*1/3\right)^{\frac{1}{3}}\right] =$$

$$= [0.53, 0.41, 1.18]$$

r_i -geometric mean of fuzzy comparison values of "Proximity to customers" criterion calculated as down

$$\tilde{r}_i = \left(\prod_{j=1}^{n} \tilde{d}_{ij}\right)^{1/n} = \left[\left(1/5*1*1/9\right)^{\frac{1}{3}}; \left(1/3*1*1/7\right)^{\frac{1}{3}}; \left(1*1*1/5\right)^{\frac{1}{3}}\right] =$$

$$= [0.29, 0.36, 0.59]$$

r_1 -geometric mean of fuzzy comparison values of "Proximity to suppliers and resources" criterion is calculated as down

$$\tilde{r}_i = \left(\prod_{j=1}^{n} \tilde{d}_{ij}\right)^{1/n} = \left[\left(3*5*1\right)^{\frac{1}{3}}; \left(5*7*1\right)^{\frac{1}{3}}; \left(7*9*1\right)^{\frac{1}{3}}\right] =$$

$$= [2.44, 3.23, 3.92]$$

The geometric means of fuzzy comparison values for different attributes shown in Table III.
In addition, the total values and the reverse values are also present

TABLE III. THE GEOMETRIC MEANS OF FUZZY COMPARISON VALUES

Criteria	$\tilde{r}_i$		
Labor climate	0.53	0.41	1.18
Proximity to customers	0.29	0.36	0.59
Proximity to suppliers and resources	2.44	3.23	3.92
Total	3.26	4	5.69
Reverse (power of -1)	0.31	0.25	0.17
Increasing Order	0.17	0.25	0.31

The fuzzy weight of "Labor climate" criteria $\tilde{w}_1$ found by using of equation (3)

$$\tilde{w}_1 = \left[(0.53*0.17);(0.41*0.25);(1.18*0.31)\right] = \\ = \left[0.09, 0.1, 0.36\right]$$

The fuzzy weight of "Proximity to customers" criterion $\tilde{w}_2$ calculated as down

$$\tilde{w}_2 = \left[(0.29*0.17);(0.36*0.25);(0.59*0.31)\right] = \\ = \left[0.05, 0.09, 0.18\right]$$

The fuzzy weight of "Proximity to suppliers and resources" criterion calculated as down.

$$\tilde{w}_3 = \left[(2.44*0.17);(3.23*0.25);(3.92*0.31)\right] = \\ = \left[0.41, 0.8, 0.9\right]$$

The weight of each criterion represented by fuzzy numbers, such as,

$$\tilde{w}_1 = (0.09,\ 0.1,\ 0.36)$$
$$\tilde{w}_2 = (0.04, 0.09, 0.2)$$
$$\tilde{w}_3 = (0.41, 0.8, 0.9)$$

A performance evaluation matrix, also known as performance assessment matrix, is an organized tool that is used to estimate and evaluate the performance of individuals, teams, projects, or processes against predefined criteria or objectives. It provides a systematic framework for measuring performance, identifying strengths, drawbacks, and facilitating decision-making related to performance improvement, recognition, rewards, and corrective actions. Performance rating decision matrix presented in Table IV.

TABLE IV. DECISION MATRIX OF PERFORMANCE RATING

	C_1	C_2	C_3
	w_1	w_2	w_3
	0.09, 0.1, 0.36	0.04,0.09,0.2	0.41,0.8,0.9
A	0.06, 0.2, 0.8	0.3, 0.8, 0.9	0.4, 0.65, 0.9
B	0.21, 0.6, 0.9	0.15, 0.5, 0.75	0.05, 0.07, 0.09
C	0.23,0.5,0.8	0.15,0.25,0.65	0.65, 0.74, 0.85

By implementing a performance evaluation matrix, organizations can effectively assess and manage performance, foster accountability, transparency, and drive continuous improvement and excellence across individuals, teams, and organizational units.
For solution of this problem and selection better location for facility used TOPSIS steps [19].
Step 4. Constructing the weighted decision matrix (Table V).

TABLE V. WEIGHTED DECISION MATRIX

	C_1	C_2	C_3
A	0.005, 0.02, 0.3	0.01, 0.07, 0.18	0.16, 0.52, 0.81
B	0.02, 0.06, 0.32	0.006, 0.04, 0.15	0.02, 0.06, 0.08
C	0.02, 0.05, 0.29	0.006, 0.02, 0.13	0.27, 0.6, 0.76

Step 5. Calculating the fuzzy positive ideal solution and fuzzy negative-ideal solution:

$$A^* = \{(0.02,\ 0.06,\ 0.32),(0.01,\ 0.07,\ 0.18),(0.27, 0.6, 0.76)\}$$
$$A^- = \{(0.005,\ 0.02,\ 0.29),(0.006, 0.02, 0.13),(0.02,\ 0.06,\ 0.08)\}$$

Determining separation measure for each alternative. For example, separation measure for first alternative can be determined as

$$S_A^* = \{\left[(0.005,\ 0.02,\ 0.29) - (0.005,\ 0.02,\ 0.29)\right]^2 + \\ + \left[(0.01,\ 0.07,\ 0.18) - (0.005,\ 0.02,\ 0.29)\right]^2 + \\ + \left[(0.16,\ 0.52,\ 0.81) - (0.005,\ 0.02,\ 0.29)\right]^2\}^{1/2} = \\ = (0.091, 0.34, 1.45)$$

$$S_A^* = (0.091, 0.34, 1.45), \qquad S_A^- = (0.08, 0.3, 1.05)$$

Similarly, for other alternatives separation measures are determined.

$$S_B^* = (0.19, 0.65, 1.98), \qquad S_B^- = (0.04, 0.19, 0.76)$$
$$S_C^* = (0.18, 0.60, 1.96), \qquad S_C^- = (0.035, 0.185, 0.73)$$

Step 6. Determination of the final rank by defining relative closeness to the ideal solutions for each alternative:

$$C_A^* = \frac{S_A^-}{S_A + S_A^-} = \frac{(0.091, 0.34, 1.45)}{(0.091, 0.34, 1.45) + (0.08, 0.31, 1.05)} = \\ = (0.05, 0.3, 5.6)$$

$$C_B^* = \frac{S_B^-}{S_B^* + S_B^-} = (0.02, 0.3, 3.4)$$

$$C_C^* = \frac{S_C^-}{S_C^* + S_C^-} = (0.07, 0.9, 9.1)$$

Step 7. Selecting best alternative. Relative closeness of each alternative to the ideal solution represents that alternative C is best alternative - $C > A > B$.

Discussion

Facility location for software systems in uncertain conditions involves determining optimal locations for data centers, servers, or other computing facilities considering potential uncertainties such as fluctuations in demand, network latency, power outages, and natural disasters. There are some basic strategies for addressing uncertainty in facility location for software systems. They are conducting a comprehensive risk analysis to identify potential sources of uncertainty, including environmental factors, market dynamics, and technical issues, assessing the probability and potential impact of each risk to inform decision-making, designing facilities with flexibility in mind to adapt to changing conditions [20]. This could involve modular designs that allow for easy expansion or relocation of computing resources based on demand fluctuations or unexpected events. Implementing redundancy and backup mechanisms to mitigate the impact of failures or disruptions is important in facility location [21]. This may include duplicating critical infrastructure components across multiple locations to ensure continuous operation in the event of failures [22]. Game-theoretic models can be used to address uncertainties in competitive environments where the actions of other agents (e.g., competitors, regulators) are uncertain [23].

This approach helps in making strategic decisions regarding facility locations. These methods, such as genetic algorithms and simulated annealing, are used to find near-optimal solutions for complex facility location problems under uncertainty. They are particularly useful when exact solutions are computationally infeasible [24]. Distributing computing facilities across geographically diverse locations for minimization the risk of localized disruptions such as natural calamity or regional infrastructure failures, considering factors such as proximity to target markets, regulatory requirements, and network connectivity when selecting locations also important issues in facility location of software systems. Implementation of real-time monitoring systems to track key performance indicators and environmental conditions and using this data to dynamically adjust resource allocation, routing decisions, and failover strategies to optimize performance and resilience in uncertain conditions are significant of facility location for software systems in uncertainty conditions. Cloud services, leverage cloud computing services are offered built-in redundancy, scalability, and geographic diversity. Cloud providers typically operate data centers in multiple regions, providing inherent resilience to failures and disruptions. Foster collaboration and communication between stakeholders, including IT teams, facility managers, and business units, ensure alignment of goals and priorities in addressing uncertainty. Regularly review and update facility location strategies based on evolving risks and opportunities. By incorporating these strategies into facility location decisions, software systems can better withstand uncertainty and maintain optimal performance and reliability in dynamic environments

IV. CONCLUSIONS

The selection of a facility location is a multifaceted decision that affects various operational, financial, and strategic aspects of a business. A thorough analysis considering these practical implications can lead to a more informed and effective location decision, ultimately contributing to the business's success. The application of the fuzzy AHP-TOPSIS hybrid method in facility location selection for software systems is not only theoretically sound but also practically applicable. It provides decision-makers with a systematic and structured approach to evaluate and select the most suitable location based on their specific requirements and constraints. To make decision on facility location selection for software systems, the methodology of fuzzy AHP-TOPSİS with fuzzy numbers is applied in this paper to take high vagueness that appropriates to the considered problem. Selection of the best location for software systems provided according to three attributes and three alternatives- A, B, C. The first attribute for selection is labor climate, the second attribute is proximity to customers, and third attribute is proximity to suppliers and resources. Results are defined from the relative closeness to the positive ideal solutions used for ranking the preference order and determined that alternative C is best alternative for facility location.

In conclusion, the application of the Fuzzy Analytic Hierarchy Process (AHP) and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) hybrid method in facility location selection for software systems, several key points can be highlighted. The hybrid method provides a robust framework for decision-making by incorporating both fuzzy AHP and TOPSIS. This integration allows for handling the inherent uncertainties and complexities involved in facility location selection for software systems. Facility location selection involves multiple

criteria such as labor climate, proximity to customers, proximity to suppliers and resources proximity. The fuzzy AHP-TOPSIS hybrid method enables the consideration of these diverse criteria and their relative importance in decision-making. Fuzzy logic is effective in capturing subjectivity and imprecision in expert judgments during pairwise comparisons of criteria and alternatives. This ensures a more accurate representation of decision-makers' preferences and enhances the reliability of the decision-making process. Sensitivity analysis allows for assessing the robustness of the selected facility location against changes in criteria weights and alternative rankings. This helps decision-makers understand the stability of their decisions and identify potential risks associated with the chosen location. Validation of the selected facility location through simulation or real-world experimentation validates the effectiveness of the hybrid method. Moreover, the method can be adapted and customized to suit different contexts and decision-making scenarios, making it a versatile tool for facility location selection in various industries.

In summary, the Fuzzy AHP-TOPSIS hybrid method offers a comprehensive and effective approach to facility location selection for software systems, integrating fuzzy logic, AHP, and TOPSIS to address the complexities and uncertainties inherent in decision-making processes. Its systematic framework, coupled with sensitivity analysis and validation, empowers decision-makers to make informed and reliable decisions regarding facility location selection.

REFERENCES

[1] Chen, Y. S., Chu, H. H., and Sangaiah, A. K., "Identifying Ambient Service LocationProblems and Its Application Using a Humanized Computing Model", *Journal of Am-bient Intelligence and Humanized Computing*, vol. 10, pp. 2345-2359, 2019. https://doi.org/10.1007/s12652-018-0.

[2] Yang J, Lee H., "An AHP decision model for facility location selection", *Facilities,* vol. 15, pp. 241–254, 1997. http://dx.doi.org/10.1108/02632779710178785.

[3] Krajewski L. J., Ritzman L. P., "*Operations Management: Strategy and Analysis*," 3rd Edition, Addison-Wesley, Boston, 1993.

[4] Kahraman C., Cebeci U., Ulukan Z. "Multi-criteria supplier selection using fuzzy AHP",*Logistics Information Management,* vol.16, no. 6, pp. 382–394, 2003. http://dx.doi.org/10.1108/09576050310503367.

[5] Hwang C., Yoon K. "Multiple attributes decision making methods and applications", *Springer,* Berlin, 1981. https://doi.org/10.1007/978-3-642-48318-9.

[6] Liang G.S., "Fuzzy MCDM based on ideal and anti-ideal concepts", *European Journal of Operational Research*, vol.112, pp. 682–691, 1999. https://doi.org/10.1016/S0377-2217(97)00410-4.

[7] Kahraman C., Ruan D., Doğan İ., "Fuzzy group decision making for facility location selection", *Information Sciences,* vol. 157, pp. 135–153, 2003. https://doi.org/10.1016/S0020-0255(03)00183-X

[8] Zadeh L.A., "Fuzzy sets", *Information and control,* vol. 8, pp. 338–353, 1965. http://dx.doi.org/10.1016/S0019-9958(65)90241-X

[9] Chang D.Y. "Applications of the extent analysis method on Fuzzy AHP", *European Journal of Operational Research* , vol. 95, no. 3, pp. 649–655, 1996. https://doi.org/10.1016/0377-2217(95)00300-2

[10] Chu T.C. "Selecting plant location via a fuzzy TOPSIS approach", *International Journal of Advanced Manufacturing Technology*, vol. 20, pp. 859–864, 2002. https://doi.org/10.1007/s001700200227.

[11] Ishii H., Lee Y.L., Yeh K.Y., "Fuzzy facility location problem with preference of candidate sites", *Fuzzy Sets Systems*, vol. 158, pp. 1922-1930, 2007. https://doi.org/10.1016/j.fss.2007.04.022.

[12] Sachin, P., Ravi K., "A fuzzy AHP-TOPSIS framwork for ranking the solutions of Knowledge Management adoption in Supply Chain to overcome its barriers", *Expert Systems with Applications*, vol. 41, pp. 679-693, 2014. https://doi.org/10.1016/j.eswa.2013.07.093.

[13] Kusumawardani R., Agintiara M., "Application of Fuzzy AHP-TOPSIS Method for Decision Making in Human Resource Manager Selection Process", *Procedia Computer Science*, vol. 72, pp. 638-646, 2015. https://doi.org/10.1016/j.procs.2015.12.173.

[14] Saaty, T.L., "The Analytic Hierarchy Process", *McGraw-Hill International*, New York, NY, USA, 1980.

[15] Hwang, C.L., Lai, Y.J., Liu, T.Y., "A new approach for multiple objective decision making", *Computers and Operational Research*, vol. 20, no. 8, pp. 889–899, 1993. doi:10.1016/0305-0548(93)90109-v.

[16] El Alaoui, M., "Fuzzy TOPSIS: Logic, Approaches, and Case Studies", *New York: CRC Press*, 2021. doi:10.1201/9781003168416.

[17] Liang G.S., Wang M.J., "A fuzzy multi-criteria decision making method for facility site selection", *The International Journal of Advanced Manufacturing Technology*, vol. 29, no. 11, pp. 2313–2330, 1991.

[18] Aliyeva K. R., "Applying Type-2 Fuzzy TOPSIS Method to Selection of Facility Location", *International Conference on Theory and Applications of Fuzzy Systems and Soft Computing*, pp. 180 – 185, 2022.

[19] Aliyeva K. R., "Decision on Facility Location on Base of Interval Data", *11th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions and Artificial Intelligence,* pp. 371-374, 2021.

[20] Snyder, L. V., "Facility location under uncertainty: A review", *IIE Transactions*, vol. 38, no. 7, pp. 547-564, 2006. https://doi.org/10.1080/07408170500216480.

[21] Ben-Tal, A., El Ghaoui, L., & Nemirovski, A., "Robust optimization", *Princeton University Press*, 2009.

[22] Santoso, T., Ahmed, S., Goetschalckx, M., Shapiro, A., "A stochastic programming approach for supply chain network design under uncertainty", *European Journal of Operational Research*, vol. 167, no. 1, pp. 96-115, 2005.

[23] Aghassi, M., Bertsimas, D., "Robust game theory", *Mathematical Programming*, vol. 107, no. 1, pp. 231-273, 2006. https://doi.org/10.1007/s10107-005-0686-0.

[24] Guo, D., Lim, G. J., Smith, J. C., "A multi-objective multi-matrix decomposition algorithm for large-scale facility location problems under uncertainty", *European Journal of Operational Research*, vol. 241, no.2, pp. 335-348, 2015.

ISSN 2085-4552

Development of Restful API Mental Health Application with Microservices Architecture Using Google Cloud Platform

Muhamad Zidan[1], Is Mardianto[2], Agus Salim[3]
[1,2,3] Universitas Trisakti, Jakarta, Indonesia
[1] aldy10ball@gmail.com, [2] mardianto@trisakti.ac.id, [3] agus@trisakti.ac.id

Accepted 08 January 2025
Approved 24 June 2025

***Abstract*— Circle application is an application on the Android operating system that is engaged in the field of mental health. The application requires a server that receives requests to fulfill the features contained in the application, therefore a server is needed one that can be scaled easily and quickly so that there are no long delays in requests, and it is hoped that the code base is easy to maintain and can be rapidly deployed.**

***Index Terms*— Cloud Computing, Google Cloud Platform, Microservices, Docker**.

I. INTRODUCTION

Mental health is a vital aspect of a person's life. Mental health is no less important than physical health for a person to have an everyday life and be able to adapt to the problems encountered throughout his life. According to WHO, mental health is a condition of well-being that individuals realize, in which they can manage reasonable life stress, work productively and productively, and participate in their community [1]. One in three teenagers (34.9%), equivalent to 15.5 million Indonesian teenagers, had one mental health problem in the last 12 months [2]. Of all primary caregivers who stated that their teenager needed help, more than two-fifths (43.8%) reported that they did not seek help because they preferred to handle the teenager's problems themselves or with support from family and friends. [2].

From the previously mentioned problems, there is the Circle application, Circle is a mental health application that has features, namely meditation, support groups, and online counseling with professional psychologists.

The application has a support group search feature, recommendation system, and payment system requires a server to process requests sent from the Circle application, so the application requires a scalable infrastructure, has an acceptable response time and is easy to maintain.

There are four main problems in developing and deploying a web application [3], namely: (1) Dependencies: a software depends on much other software, especially in the form of libraries for certain specific. (2) Incomplete documents or not solving initial installation and operational problems. (3) Code rot. Different versions of the library, operating system (kernel), or development language (interpreter) can also differ in the results provided by the application. For example, a bug update to a kernel or library can make software created to handle previous errors create new problems. (4) Barriers to adoption and reuse of pre-existing solutions.

The difficulties present in the above traditional approaches can be solved by virtualization technology. According to Zhang [4], virtualization is an integral part of modern cloud infrastructure such as Amazon's Elastic Compute Cloud (EC2) and Google's App Engine.

With the use of container-based virtualization technology in the form of Docker, web application development has several advantages [5], namely, making portable applications, more efficient use of computer resources, lightweight, fast, and suitable for developing microservices architecture.

This research improve reliability, fault tolerance, scalability and Flexibility with the use of virtualization technology and utilizing Google Cloud Platform robust infrastructure.

II. LITERATURE REVIEW

There are studies conducted by previous researchers who handled cases similar to this research. Here is a list of prior research used in this research as a reference in building solutions.

- Qalam AIlmiah dan Riko Virgiawan Z. PERANCANGAN ARSITEKTUR BACKEND MICROSERVICE PADA STARTUP CAMPAIGN.COM, The problem experienced by this research is that the system design built at campaign.com still uses a monolithic architecture. The user interface, logic processing, and data access are

combined into one program and placed in one database. So it isn't easy to maintain. The advantages in this study, namely the provision of methods that are easy to understand and implement in different studies, while the weaknesses in this study are that the endpoint is only accessed by the web, and the docker configuration and implementation are still done manually, there is no stress testing.

- Sinambela A dan Farady Coastera F, IMPLEMENTASI ARSITEKTUR MICROSERVICES PADA RANCANG BANGUN APLIKASI MARKETPLACE BERBASIS WEB. The problem experienced by researchers is the use of monolith architecture which makes the research server consume significant computing resources and also causes obstacles in developing new features on the system. The advantage of this research is the creation of an API gateway that can unite separate services. In contrast, the disadvantages contained in this research are deployments that are still done manually and there is no stress testing.
- Jhay Shah dan Dushyant Dubaria. Building modern clouds: Using Docker, Kubernetes Google Cloud Platform. The problem experienced by researchers is that researchers want to find a faster way of deployment, and need a facility that can scale well. The advantages of this research discuss the use of Kubernetes that can manage many docker containers efficiently and are fully documented. In contrast, the disadvantages of this research are the lack of performance tests on the infrastructure that has been created.

A. Microservice Architecture

Unlike a monolithic application, Microservice means dividing an application into smaller, interconnected services. Each microservice is a small application with a hexagonal architecture consisting of logic and various adapters, as illustrated in Figure 1.

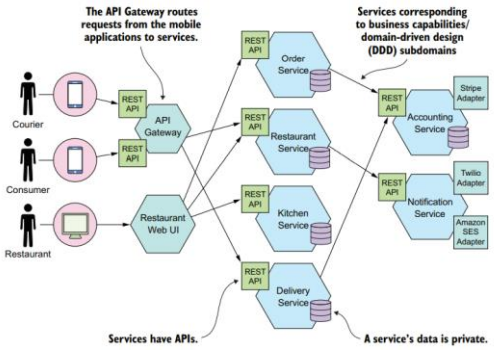


Figure 1. Microservices Architecture

Microservice architecture is a more scalable distributed alternative that provides more focused and specific services. Large problems will be broken down into several small solutions organized into a single service, where each service has its responsibilities. With this approach, an information system will consist of several services that can be managed and distributed independently, making it easier for the system to adapt to changing needs [6].

B. RestFul API

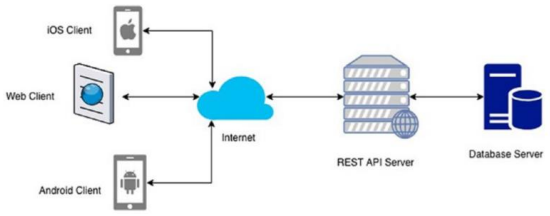


Figure 2. REST Architecture Diagram

Representational State Transfer (REST) is a style of software architecture for web services that provides standards for data communication between different types of systems [11]. In simple terms, REST is a standard for exchanging data over the Web for interoperability between computer systems. REST allows us to distinguish between client and server and implement client and server independently. The most important feature of REST is its statelessness, which means that neither client nor server needs to know each other's status to be able to communicate [7].

C. Docker

Docker is an open platform that can be used to build, distribute, and run applications and has a portable, lightweight packaging tool known as the Docker Engine. Docker also provides Docker Hub, a cloud service for sharing applications. The cost of using Docker containers is much more efficient than ordinary virtual machines. This reduces the cost of building applications on cloud computing platform providers [8]

Applications built on Docker are packed with all the dependencies they need into a standard form called a container. These containers continue to run in isolation on top of the operating system kernel, Docker containers can be easily deployed to cloud-based environments [3].

Containerization is a way of running multiple software applications on the same machine. Each is run in an isolated environment called a container. A container is a closed environment for software. It combines all the files and libraries the application needs to function correctly. Multiple containers can be deployed on the same machine and share resources. Docker uses images to create containers [9].

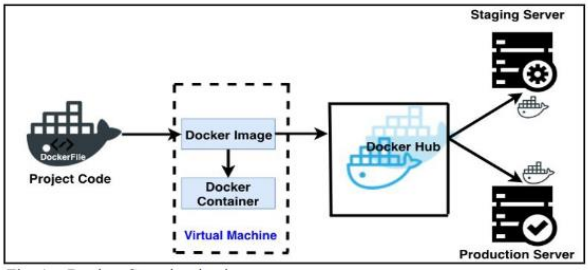


Figure 3. REST Docker Containerization

D. Google Cloud Platform

Google Cloud Platform (GCP) is a cloud computing service product owned by Google. GCP includes public cloud infrastructure and enterprise versions of G-Suite of Android, Chrome, and application programming interfaces for machine learning and enterprise mapping services. Google Cloud Platform provides over 100 services, such as computing, Networking, Storage and Databases, and others [10].

E. Cloud Build

Cloud Build is a service that runs app builds or dockers on Google Cloud. Cloud Build can import code from various repositories or cloud storage spaces, resulting in predefined specifications, and generate artifacts such as Docker containers or Java archives [11]. Cloud Build can easily be integrated into various public or private repositories such as existing repositories on GitHub this integration is done by Cloud Build utilizing pub-sub actions contained in the repository, Cloud Build can also automatically perform unit tests and integration tests to other Google Cloud Platform services and also Cloud Build has an auto deployment feature where the results of the build can be directly deployed to several Google Cloud services Platforms like Cloud Run.

F. Cloud Run

Cloud Run is a Serverless computing platform entirely managed by Google, Cloud Run can help run applications in highly scalable containers and can be called via web requests or from webhooks. Built on Knative, it enables high application portability. With Cloud Run, users can automatically scale up or down from zero to N. Cloud Run services are regional and automatically replicated across multiple zones. Cloud Run provides out-of-the-box integrations with Cloud Monitoring, Cloud Logging, Cloud Trace, and Error Reporting to monitor application performance [12].

Cloud Run offers some configurations that can be changed in each of the revisions of the container that we deploy, some of the configurations namely: (1) Port which request will be sent to the port of the container, (2) CPU allocation, which we can toggle if CPU is only allocated during request processing or CPU is always allocated, (3) Capacity of the Container including Memory, numbers of vCPUs, how long the request should be timed out, and maximum number of concurrent requests per container instance, (4) Execution environment which is First generation who prioritize speed from a cold start and Second generation that can utilize files system, full Linux compatibility, faster CPU, and Network performance, (5) Autoscaling that Bounds the number of container instances for the created container revision, and Lastly (6) Environment Variables

G. Cloud Storage

Cloud storage is one of Google Cloud Platform services to store unstructured data, Cloud storage uses an object-based storage system. An object is an immutable data consisting of files of any format. You store objects in containers called buckets. All buckets are associated with a project, and you can group your projects under an organization. Every project, bucket, and object in Google Cloud is a resource Google Cloud, as are things like Compute Engine instances. Cloud storage has several classes for its storage types:

Table 1. Cloud Storage Classes

Storage Class	API Name and CLI	Minimum Duration Storage	Typical Monthly Availability
Standard Storage	STANDARD	None	>99.99% di multi-regions dan dual-regions, 99.99% di regions
Nearline Storage	NEARLINE	30 Days	99.95% di multi-regions dan dual-regions, 99.9% di regions
Coldline Storage	COLDLINE	90 Days	99.95% di multi-regions dan dual-regions, 99.9% di regions
Archive Storage	ARCHIVE	365 Days	99.95% di multi-regions dan dual-regions, 99.9% di regions

H. Flask

Flask is a lightweight microframework for web applications built on Python, which provides an efficient framework for building web-based applications that use the flexibility of Python and

strong community support with scaling capabilities to serve millions of users [7]. Flask has two main components, Werkzeug and Jinja. While Werkzeug is responsible for providing routing, debugging, and Web Server Gateway Interface (WSGI). Flask leverages Jinja2 as a template engine. Natively, flask does not support database access, user authentication, or other high-level utilities. Still, it provides support for the integration of extensions to add all such functionality, making Flask a micro yet production-ready framework for developing applications and web services. A simple flask of an application can fit into a single Python file or it can be modulated to make the application production-ready. The idea behind Flask is to build a good foundation for all applications leaving everything else on the extension [7].

I. Firebase

Firebase is a BaaS (Backend as a Service) service provided by Google, Firebase is considered a web application platform. Firebase helps developers build quality apps quickly. then Firebase stores the data in JavaScript Object Notation Format (JSON) which does not use queries to insert, update, delete, or append data to it. The system's backend is used as a database to store data [13].

J. Firebase Auth

Firebase Auth supports social sign-in features such as Facebook, Google, GitHub, and Twitter. It is a service that can authenticate users by using client-side code and is a paid service. Firebase Auth also includes a user management system where developers can enable user authentication with email and password logins stored with Firebase [13].

K. Firestore

Firestore is a flexible and scalable mobile, web, and server development database from Firebase and Google Cloud Platform. Like Firebase's real-time databases, Firestore syncs your data across all client apps via real-time handlers and provides offline support for mobile and web. This way, you can build responsive and efficient applications without network latency or internet connection. Firestore also offers seamless integration with other Firebase and Google Cloud products, including Cloud Functions [14]. Firestore is a cloud-hosted NoSQL database that can be accessed directly by your Apple, Android, and web apps through native SDKs. Firestore is also available in the native SDKs of Node.js, Java, Python, Unity, C++, and Go, in addition to the REST API and RPC API [14].

III. METHODOLOGY

The steps in this research are carried out in the following stages:

1. Conduct interviews and surveys to the target market to get the wants and needs of the target market. This data retrieval is carried out intermittently to get an idea of what features need to be made.
2. Based on the results of interviews and surveys on the target market and get the features that are wanted, researchers design UML for each use case and dissect it into microservices respectively.
3. After determining the microservice division, researchers design the use of Google Cloud Platform services which are made into a diagram, the services used by researchers like Cloud Build, Cloud Storage, Cloud Run, Firebase, and Firestore.
4. Build API endpoints based on predefined use cases and microservices, at this stage, researchers build a structure and codebase in a local environment.
5. Testing the code base in the local environment to ensure the features in the microservice section have run well under the system design that has been made, with the black-box testing method.
6. Deploy code in the local environment to the Google Cloud Platform infrastructure with containers. At this stage, researchers directly use the Google Cloud Platform infrastructure to deploy code in the GitHub repository and install the container in the cloud run.

Load test API endpoints that have been deployed to the Google Cloud Platform. This stage is necessary to simulate real traffic and ensure that the system is not damaged under the specified load

IV. RESULT AND DISCUSSION

A. User Requirement

Based on the results of focus group discussions with several psychologists related to the use of support group features and mental health application features in general by users in building systems in the Circle application, here are some user needs needed in the system in the Circle application:

1. Users can search for the desired support group and enter the group chat
2. Users can order psychologist services in the form of consultations and pay according to the services that have been chosen
3. Users can do support group activities in the available group chat.
4. User can register and log in to the application.
5. User can do screening test.
6. Caregivers can access the results of the user screening test.
7. Caregivers can create support groups and moderate chat rooms.

B. System Requirement

Based on the results of focus group discussions with several psychologists related to the use of support group features and mental health application features in general by users in building systems in the Circle application, here are some of the system that needs to be needed in the Circle application:

1. Restful API will be deployed on the Google Cloud Platform.
2. Restful API has a simple security system in the form of an API Key.
3. Restful API can be accessed via HTTPS request.
4. The features required in creating a Restful API are as follows:
   - Create a Support Group in the form of a chat room.
   - Search Support Group for Users.
   - Create a login token for the GetStream service.
   - Booking Psychologist services.
   - User payment handling.
   - Create and evaluate screening test results for Users.

For the specification of each existing docker container unit to create a restful API, a minimum of the following specifications are required:

1. 1 Core CPU
2. 512 MB RAM
3. Access Port 8080

Then there is also a software environment that must be prepared, namely:

1. Python 3.9
2. Python libraries: Tensorflow, Sastrawi, GetStream-Client, Midtrans, Flask, and Firebase.

C. Use Case Diagram

Use case diagrams to illustrate a graphical display of the functioning of a system. For an explanation of the role of each actor performed on the Circle application can be seen in the following image:

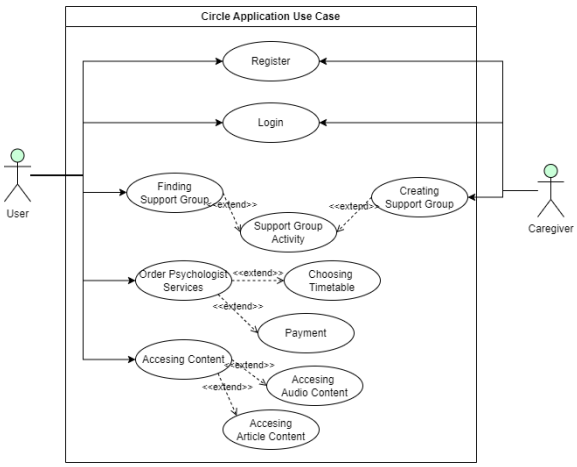


Figure 4. Use Case Diagram

D. Activity Diagram

The following stages of explanation of the role activities of each actor are contained in Figure 3.D.2 use case diagram:

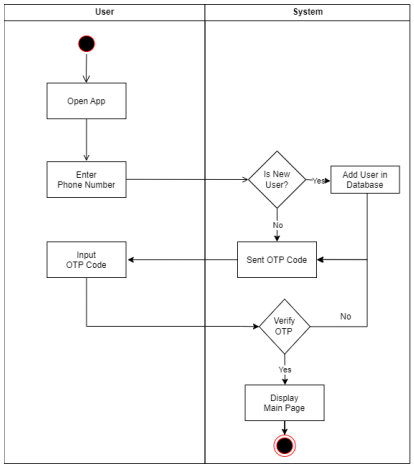


Figure 5. Activity Diagram: Log in and Register

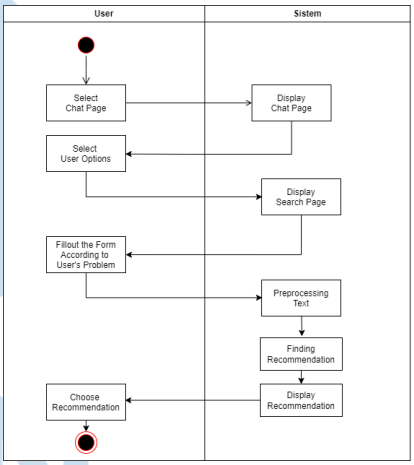


Figure 6. Activity Diagram: Search Support Group

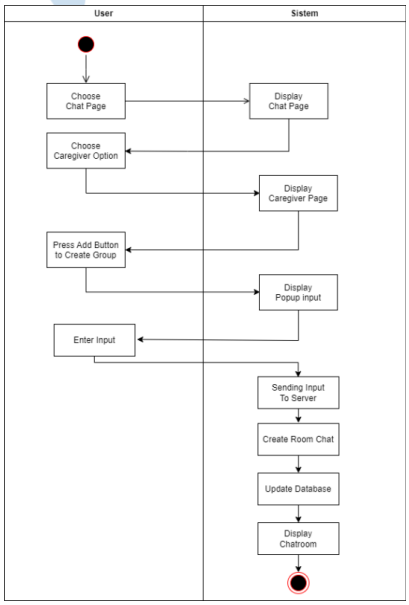


Figure 7. Activity Diagram: Support Group Creation

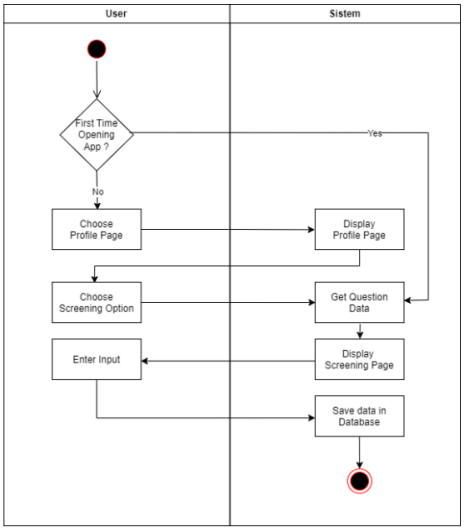


Figure 8. Screening: Psychological Tests

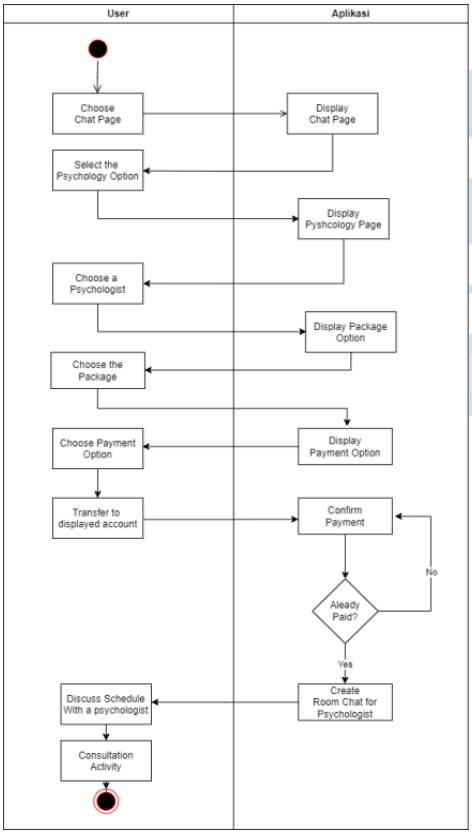


Figure 9. Psychologist Consultation

E. Microservices Architecture on Google Cloud Platform

Referring to user requirements and system requirements services in the Circle application can be divided into 4 systems, namely:

1. Chat Service
2. Payment Services
3. Search Service
4. Screening Test Service

With these 4 services can be made a picture of the system architecture in making it into Google Cloud Platform, here is an overview of the use of Google Cloud Platform products and microservices architecture that will be used:

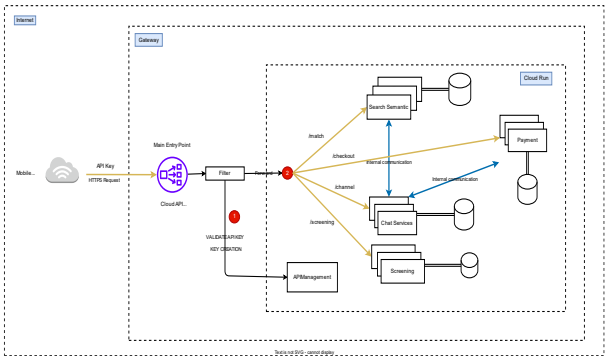


Figure 10. Microservice Architecture

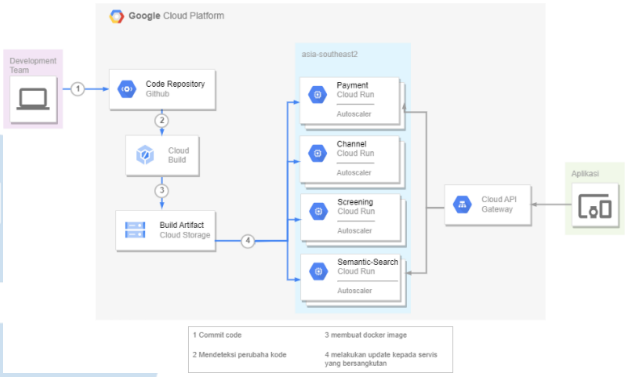


Figure 11. GCP Usage

In the GCP usage image above, there are several services that are used for Circle application needs, namely:

1. Cloud Build

   Cloud Build is used to create a Docker image that matches the contents of the repository on Github, each service has its repository so that they are mutually independent from other services
2. Cloud Storage

   Cloud Storage is used to store every Docker image version that Cloud Build has created, so this can help in storing stable backup versions
3. Cloud Run

   Cloud Run is used as a serverless platform which will be the main server for each of the microservices mentioned earlier, Cloud Run also has an autoscaler system so that it can multiply Containers so that it automatically scales requests received from the Circle Application, this happens on-demand so if there are no requests from the application, the resources from Cloud Run will not be used
4. Cloud API Gateway

   Cloud API Gateway allows us to manage URLs from different Cloud Runs into one URL domain making it easier for developers to access APIs, Cloud API Gateway also provides monitoring, alerts, logging, and tracing for each use

F. Cloud Run Configuration

Based on the Microservice Architecture in the previous figure we can utilize Cloud Run to divide the Services and use Firebase as the Database for each service, for the specification of the container for each service in Cloud Run is shown in the next tables

Table 2. Search Services Cloud Run Configuration

Search Service	
CPU allocation	CPU is only allocated during request processing
Startup CPU boost	Enabled
Maximum Concurrency per instance	80
Request timeout	300 seconds
Execution environment	Second generation
CPU	1
Memory	2 GiB
Minimum number of instances	0
Maximum number of instances	10
Port	8080
Build	Cloud Build
Environment Variables	Firebase_Key

Table 3. Chat Services Cloud Run Configuration

Chat Service	
CPU allocation	CPU is only allocated during request processing
Startup CPU boost	Enabled
Maximum Concurrency per instance	80
Request timeout	300 seconds
Execution environment	Second generation
CPU	2
Memory	512 MiB
Minimum number of instances	0
Maximum number of instances	5
Port	8080
Build	Cloud Build
Environment Variables	Firebase_Key, GetStream_Key

Table 4. Payment Service Cloud Run Configuration

Payment Service	
CPU allocation	CPU is only allocated during request processing
Startup CPU boost	Enabled
Maximum Concurrency per instance	80
Request timeout	300 seconds
Execution environment	First generation
CPU	1
Memory	512 MiB
Minimum number of instances	0
Maximum number of instances	5
Port	8080
Build	Cloud Build
Environment Variables	Firebase_Key, Midtrans_Key

Table 5. Payment Service Cloud Run Configuration

Screening Test Service	
CPU allocation	CPU is only allocated during request processing
Startup CPU boost	Enabled
Maximum Concurrency per instance	80
Request timeout	300 seconds
Execution environment	First generation
CPU	1
Memory	512 MiB
Minimum number of instances	0
Maximum number of instances	30
Port	8080
Build	Cloud Build
Environment Variables	Firebase_Key

From the tables above, we allocated the CPU and Memory of the container based on how expensive the computational power of each service is, from the tables we can conclude that the Search Service is the most demanding in terms of Memory which is caused by the Tensorflow Library.

With Cloud Run when there's a massive spike of request to the Search Service we can individually scale the necessary Service without affecting other services.

To save cost, all of the services have a minimum number of container instances set to zero which enables a cold start, the number of instances will go up if and only if there's a request coming to the port of the container by utilizing the on-demand feature of the container we can reduce the billing of the cloud significantly.

G. Containerization

Containerization is a technology that allows you to package an application and all its dependencies, libraries, and configurations into a single, isolated unit called a "container." This container can then be easily deployed and run consistently on any server that supports containerization without worrying about compatibility issues.

To create a container we need a configuration setup called Dockerfile, This file contains instructions on how to build your container step-by-step. It specifies the base image (a minimal operating system with pre-installed tools), adds your application code, and sets up

the required configurations. For Circle Services we use the following setup:

```
From python:3.9.11

COPY src/ src/
COPY requirements.txt.

RUN pip install --upgrade cython
RUN pip install --upgrade pip
RUN pip install --no-cache-dir -r requirements.txt


WORKDIR /src

ENV PORT 8080

CMD exec gunicorn --bind :$PORT --workers 1 --threads 8 app:app
```

From the Dockerfiles above we can configure the directory of the files by first copying the app files to the Docker directory, and then we can set up the environment suited to the requirements we need

H. Restful API Design

Based on activity diagrams, use case diagrams and microservices architecture design, several endpoints can be made that will be used in Circle Applications, namely:

Table 6. Restful API

No	Fungsi	URL	Parameter	Method	Expectation
1	Login and Register	/token	User id, password	GET	Return status and Token
2	Getting Screening Data	/screening	User_id,screening_name	GET	Return data about screening
3	Post the Result of Screening	/screening	User_id, screening_name, jawaban	POST	Return Feature Recommendation
4	Search Chatroom	/match	User_id, k_value, text	GET	Return Chatroom recommendation
5	Create Chat Room	/channel	User_id, Text, Title, Max_user	POST	Return Channel information
6	check out	/checkout	Appointment_id, user_id	GET	Return URL and transaction token
7	Notification Handling	/notification	Transaction_id,Transaction_status, Order_id	POST	Change the order status and return status code

I. Blackbox Test

To test the various feature of the API that have been designed we need some tests to confirm it. Blackbox test is designed to test the whole infrastructure to check if the integration between the services is correct:

Table 7. Blackbox Test

URLs	Test Case	Expected Result	Tested results
/token	Correct *User ID* and *Password*	Return Token Session	Valid
	Inccorect *User ID* or Password	Return incorrect message	Valid
	Missing parameter	Return incomplete parameter Message	Valid
/screening (GET)	Correct *User ID* and *Screening Name*	Return Screening Question Details	Valid
	Incorrect *User ID* and *Screening Name*	Return Incorrect message	Valid
	Missing Parameter	Return incomplete parameter Message	Valid
/screening (POST)	Correct *User ID, Answer Screening Name*	Return Feature Recommendation	Valid
	Incorrect *User ID, Answer and Screening Name*	Return Incorrect message	Valid
	Missing	Return	Valid

	Parameter	incomplete parameter Message	
/match	Correct *user id, text* and *k_value*	Return list of recommended room id	Valid
	K_value is not in parameter	Return top 3 in the list of recommended room id	Valid
	Inccorect *User ID or K_Value*	Return Incorrect message	Valid
/Channel	Correct *User ID,Text*, *Title,* and Max_user	Return Channel Information	Valid
	Inccorect *User ID*	Return Incorrect message	Valid
	Missing Parameter	Return incomplete parameter Message	Valid
/checkout	Correct *User ID* an *Appointment ID*	Return Redirect URL and transaction token	Valid
	Inccorect *User ID* and *Appointment ID*	Return Incorrect message	Valid
	Missing Parameter	Return incomplete parameter Message	Valid
/Notification	Correct *Transaction ID* and *Transaction Status*	Return Received Status Message	Valid
	Incorrect *Transaction ID* and *Transaction Status*	Return Incorrect Message	Valid
	Missing Parameter		

J. Load Test

For the Load Test, we use Locust as the Python framework of choice, the infrastructure of the microservices will be tested again with 5000 virtual users in a span of 10 minutes, and the following figures will describe how many users will be connected to the server and how many requests have been requested:

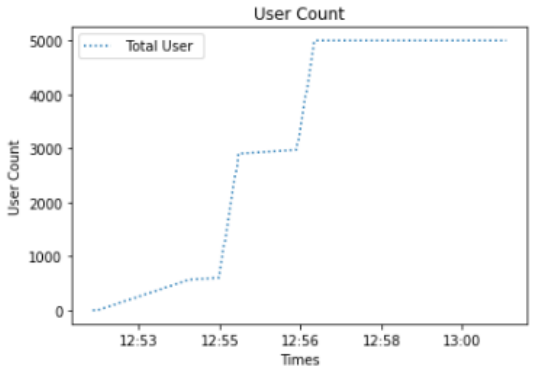


Figure 12. Total User Graph

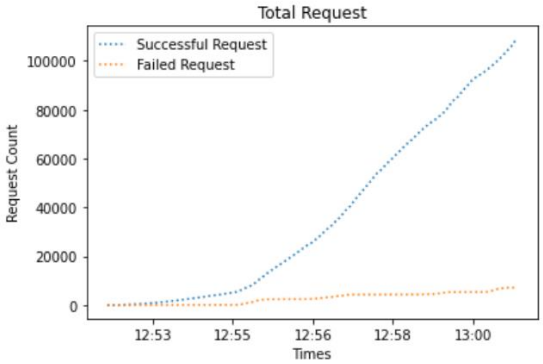


Figure 13. Total Request Graph

From the figures above, there are 5000 virtual users, which gradually increased from 0 to 5000 within 10 minutes. There's a total of 115910 requests that are coming from the virtual users, as you can see from figure 4.J.2 There is some error that occurs when the user spike to more than 1000 and the total of the error is 7284 error, the detail of the error will be explained in the following table:

Table 8. Total Error Table

URL	Error	Number of Occurences
/match	HTTPError('429 Client Error: Too Many Request)	2820
/token	HTTPError('500 Server Error: Internal Server Error)	3957
/channel	HTTPError('429 Client Error: Too Many Request)	507

From the error table above, we can infer that there are 2 types of errors that happen during the load test, which is HTTP Error 429 and HTTP Error 500, the reason that HTTP Error 429 appear in /match and /channel is because of too many requests have been created for their endpoint, and the initial container can't scale fast enough to keep up with the demands. As for

HTTP error 500 is caused by the limit rate from the GetStream library that is used for getting the user a token to chat with other users, the limit rate is caused by the free-tier plan from the GetStream services.

The load test will also measure the response time of the entire service and each endpoint, for the parameter of the response time, we will measure the average response time and median response time, for the unit of the response time, we will use milliseconds, for the detail of the response time of the server will be illustrated in the following figures:

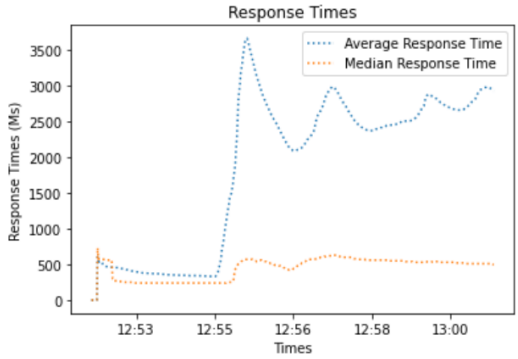


Figure 14. Response Time Graph

From the response time figures above, we can conclude that there's a huge spike when the use count reaches 5000, but it gradually lower, this is the result of the Cloud Run Autoscaler that automatically deploy new containers to serve the new Demands. The average response time when the server stabilize in 5000 users is 2817 milliseconds, we can dissect this further by exploring each endpoint response time as the following tables:

Table 9. Token URL Response Time

URL	/Token
Method	GET
Request Count	4999
Median Response Time	10000 ms
Average Response Time	10855 ms
Min Response Time	532 ms
Max Response Time	30011 ms
Average Content Size (Byte)	241 byte

Table 10. GET Screening URL Response Time

URL	/Screening
Method	GET
Request Count	26200
Median Response Time	260 ms
Average Response Time	384 ms
Min Response Time	219 ms
Max Response Time	5460 ms
Average Content Size (Byte)	4540 byte

Table 11. POST Screening URL Response Time

URL	/Screening
Method	POST
Request Count	25878
Median Response Time	280 ms
Average Response Time	400 ms
Min Response Time	223 ms
Max Response Time	5574 ms
Average Content Size (Byte)	16 byte

Table 12. Match URL Response Time

URL	/Match
Method	GET
Request Count	42012
Median Response Time	2700 ms
Average Response Time	4598 ms
Min Response Time	187 ms
Max Response Time	28278 ms
Average Content Size (Byte)	224 byte

Table 13. Channel URL Response Time

URL	/Channel
Method	POST
Request Count	8228
Median Response Time	5200 ms
Average Response Time	6731 ms
Min Response Time	185 ms
Max Response Time	30011 ms
Average Content Size (Byte)	547 byte

Table 14. Checkout URL Response Time

URL	/Checkout
Method	GET
Request Count	8526
Median Response Time	270 ms
Average Response Time	368 ms
Min Response Time	235 ms
Max Response Time	5597 ms
Average Content Size (Byte)	156 byte

Table 15. Notification URL Response Time

URL	/Notification
Method	POST
Request Count	67
Median Response Time	1500 ms
Average Response Time	2333 ms
Min Response Time	585 ms
Max Response Time	10044 ms
Average Content Size (Byte)	877 byte

From the seven tables, we can see a drastic change in the minimal response time to the max response time, this is because the warm up session with only 100 virtual user and gradually increase it to 5000 virtual user, we can use the minimal response time as the baseline and compare it to the median response time to get a sense how certain URL behave in this Load Test, Meanwhile the maximum response time corresponds to time when the user count spiked to 5000, and the server is rolling a new container, the time between finishing rolling the new container and the start of the request makes the max response time so much high compared to the average response time. In the previous

load test we can see /match, /token, and /channel URLs have massive differences in minimal response time to average response time, for /match and /channel the high response time is caused by the resource-intensive system created by Tensorflow library and to prevent using too many resources of Cloud Run, limiting the number of containers is vital, as for /token case many of the error is caused by internal server error is caused by the limit rate from the GetStream library that is used for getting the user a token to chat with other users, the limit rate is caused by the free-tier plan from the GetStream services.

We can see how the container of each services automatically scale when facing a large request in the following figures:

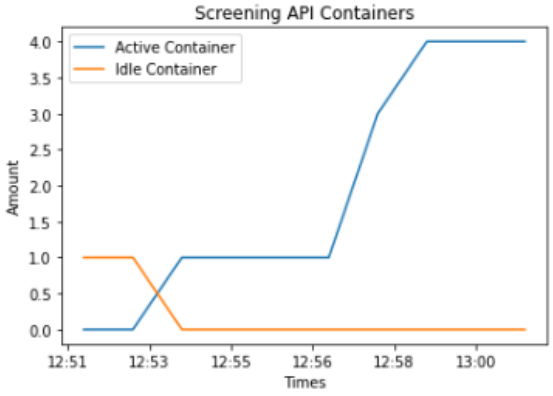


Figure 15. Screening Services Container Graph

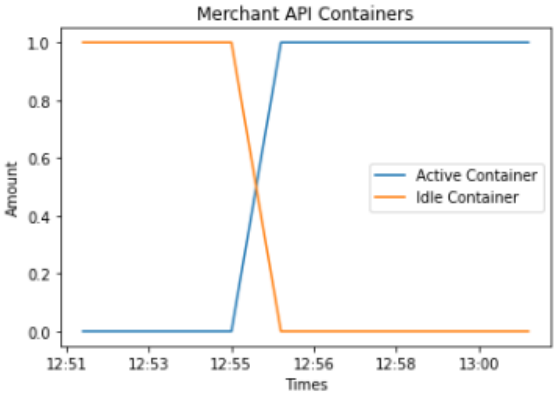


Figure 16. Merchant Services Container Graph

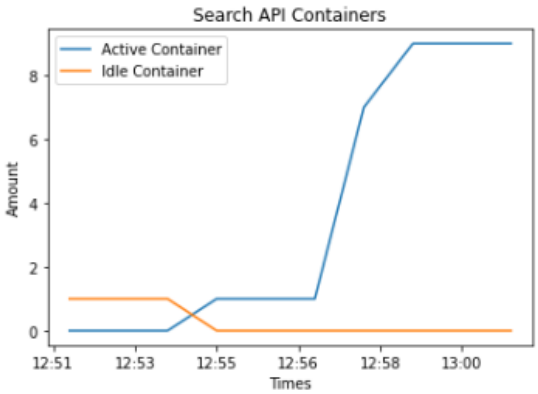


Figure 17. Search Services Container Graph

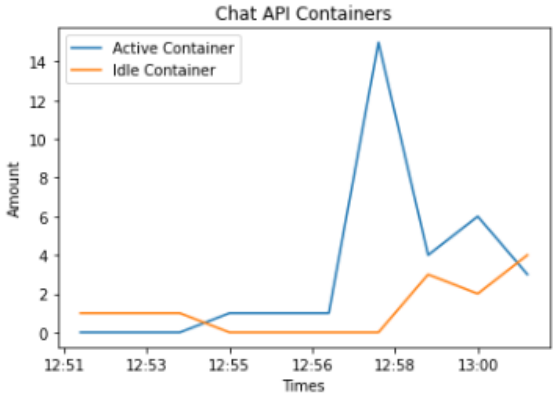


Figure 18. Chat Services Container Graph

From the figures above we can see how the Container automatically increases when the demand is high.

V. CONCLUSIONS

Based on the result of the research "Development of Restful API Mental Health Application with Microservices Architecture Using Google Cloud Platform" that has been carried out, there are several conclusions:

1. The making of microservices pattern in API for Circle Application that utilizes Google Cloud Platform as its cloud provider can be used and scaled based on the demand of the request from the application
2. Using a microservices pattern can remove dependencies from other services and as a result other services' load won't cause any disturbance to unrelated services, this also makes the scaling adjustable to only the most used service which raises the efficiency of the server and lowers the cost of the cloud computing cost
3. Using Cloud Build makes it easy to deploy a docker container and make changes to an existing container
4. Dockerization makes it easy to isolate bugs and faults within the system, and because we can set up the environment and virtualize it, it's easy to set up the same environment across multiple instances

There are vital factors that need to be considered when using Cloud Run in Google Cloud Platform that is: the computing cost of the services, the service Resource (CPU, Ram) requirement, how frequently it will be accessed, and how it handle shared data.

REFERENCES

[1] WHO, Basic documents, 43rd Edition. Geneva: World Health Organization, 2014.

[2] Indonesia National Adolescent Mental Health Survey (I-NAMHS), 2023, [Online]. Available: https://qcmhr.org/~teoqngfi/outputs/reports/12-i-namhs-report-bahasa-indonesia/file.

[3] C. Boettiger, "An introduction to Docker for reproducible research," in *Operating Systems Review (ACM)*, 2015. doi: 10.1145/2723872.2723882.

[4] Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: State-of-the-art and research challenges," *Journal of Internet Services and Applications*, vol. 1, no. 1, pp. 7–18, May 2010, doi: 10.1007/s13174-010-0007-6

[5] S. Singh and N. Singh, "Containers & Docker: Emerging roles & future of Cloud technology," in *Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology, iCATccT 2016*, 2017. doi: 10.1109/ICATCCT.2016.7912109.

[6] M. Fadlulloh and R. Bik, "IMPLEMENTASI DOCKER UNTUK PENGELOLAAN BANYAK APLIKASI WEB (Studi Kasus : Jurusan Teknik Informatika UNESA)," 2017.

[7] K. Relan, *Building REST APIs with Flask*. 2019. doi: 10.1007/978-1-4842-5022-8.

[8] B. B. Rad, H. J. Bhatti, and M. Ahmadi, "An Introduction to Docker and Analysis of its Performance," *IJCSNS International Journal of Computer Science and Network Security*, vol. 17, no. 3, 2017, [Online]. Available: https://www.researchgate.net/publication/318816158

[9] J. Shah and D. Dubaria, "Building modern clouds: Using docker, kubernetes google cloud platform," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference, CCWC 2019*, 2019. doi: 10.1109/CCWC.2019.8666479.

[10] Google, "Google Cloud Products." https://cloud.google.com/products (accessed Oct. 21, 2022).

[11] Google, "Overview of Cloud Build." https://cloud.google.com/build/docs/overview (accessed Oct. 21, 2022).

[12] I. Barokah and A. Asriyanik, "Analisis Perbandingan Serverless Computing Pada Google Cloud Platform," *Jurnal Teknologi Informatika dan Komputer*, vol. 7, no. 2, 2021, doi: 10.37012/jtik.v7i2.662.

[13] C. Khawas and P. Shah, "Application of Firebase in Android App Development-A Study," *Int J Comput Appl*, vol. 179, no. 46, pp. 49–53, Jun. 2018, doi: 10.5120/ijca2018917200.

[14] Google, "Cloud Firestore Documentation." https://firebase.google.com/docs/firestores (accessed Oct. 22, 2022).

ISSN 2085-4552

Association Rule Mining of Consumer Behavior at MOY Supermarket Using Apriori Algorithm

Eka Putri Oktavia[1], Amalia Nur Alifah[2], Mediliadita Dwi Dimastia[3], Yoseph Friden Bani Wodju[4], Febrian Farda Hasdikiyah[5]
[1,2,3,4,5] Dept. of Data Science, Telkom University, Surabaya, Indonesia
[1] ekaptroktavia@student.telkomuniversity.ac.id, [2] amalialifah@telkomuniversity.ac.id, [3]mediliadita@student.telkomuniversity.ac.id, [4] fridenwod@student.telkomuniversity.ac.id , [5]febrianfh@student.telkomuniversity.ac.id

Accepted 28 May 2025
Approved 04 June 2025

***Abstract*— MOY Frozen Food is a retail business located in Kediri Regency, specializing in selling frozen food, beverages, and necessities. In recent years, the retail industry has faced numerous challenges, including shifts in consumer behavior, technological advancements, and increasing competition. This study addresses the issue of identifying which products are frequently purchased together and determining appropriate recommendations for consumers. To achieve this goal, association rules are employed to discover correlations and co-occurrences among data sets, which facilitate identifying product relationships within a single transaction. Using the Apriori algorithm with a minimum support threshold of 0.01 and a confidence level of 0.5, implemented in Python, this research successfully generates association rules. The insights derived from these association rules can be leveraged to develop various sales strategies, ultimately enhancing product sales at MOY Frozen Food.**

***Index Terms*— *association rule mining; apriori algorithm; consumer behavior; product recommendation; sales strategy*.**

I. INTRODUCTION

Making quick and accurate decisions to determine business strategies is one of the many influential factors that business owners, including those at MOY Supermarket, must consider. The growth and competition in global trade, driven by free-market economies and advances in information technology, have intensified the level of competition, making it more transparent and challenging for companies to meet increasing customer demands [1]. These factors highlight the importance of making business decisions based on data [2], rather than relying solely on intuition. The data used should be derived from the collection of every transaction that occurs.

There are many ways to understand market conditions in the retail business, one of which is by analyzing sales transaction data. Daily, weekly, monthly, or yearly, vast transaction data are stored in the company's database [3]. In the digital era, competition in the retail business is becoming increasingly intense, necessitating new marketing strategies to attract consumers [4]. Digital marketing significantly influences the growth of retail businesses, requiring companies or retail stores to keep up with technological advancements in business, which are key to achieving success [5]. Consequently, the amount of data stored in digital product sales databases will continue to grow, becoming big data, which forms the foundation for the emergence of data mining.

Data mining in economics, particularly in retail business, can help solve business problems, such as identifying patterns in customer transaction behavior. By understanding these patterns, a retail business can more easily make decisions regarding marketing strategies aimed at consumers. One customer behavior that can be identified and analyzed is the combination of products frequently purchased together. Data mining techniques can be used to discover association rules, specifically the association rules between combinations of items, to determine which products are bought simultaneously. This association process utilizes the Apriori algorithm. The Apriori algorithm, which was introduced by Agrawal and Srikant in 1994 [6], is used to determine frequent item sets for Boolean association rules [7].

In a previous study [8] titled " Penerapan Algoritma Apriori Terhadap Data Mining Penjualan di Alfamart Berastagi" 75 association rules were generated with a minimum support of 50% and a confidence of 0.9. Meanwhile, another study [3] titled "Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Penjualan Bisnis Ritel" produced 3 association rules with a minimum support of 0.6 and a confidence of 0.8. Both studies focused on retail companies; study [8] utilized RapidMiner software, while study [3] was tested using the Orange tool. In contrast, the current research focuses on a medium-

scale retail business and uses Python software for testing.

Business activities in Indonesia are largely conducted by the community through Micro, Small, and Medium Enterprises (MSMEs) [9]. MSMEs also play a crucial role in Indonesia's economy [10], necessitating efforts to enhance their strategies to stay competitive with larger businesses. MOY Frozen Food Supermarket is one such MSME operating in the retail sector, primarily focusing on food products, particularly frozen foods, beverages, and daily necessities. This supermarket must meet the needs of its customers daily and be driven to develop the right sales strategies. To achieve this, the company needs sufficient information for further analysis. For example, data from recorded sales transactions can be used to understand consumer habits or behavior and develop strategies to achieve the desired business goals.

II. METHODOLOGY

This study employs the association rule-mining method to analyze consumer purchasing patterns at MOY Frozen Food Supermarket. By uncovering hidden patterns and correlations within sales transaction data, this method allows for a deeper understanding of consumer behavior, which can be leveraged to enhance marketing strategies and inventory management. The Apriori algorithm, a well-known technique in data mining, is utilized in this research due to its efficiency in discovering frequent item sets and generating association rules from large datasets. The process begins with data collection, followed by data pre-processing and mining, where meaningful patterns are extracted. The flowchart illustrating the entire research methodology, including each step of the data analysis and implementation, is presented in **Figure 1.** This structured approach aims to provide actionable insights that can help MOY Frozen Food Supermarket optimize its operations and improve customer satisfaction.

The flowchart in **Figure 1** illustrates the process of applying the Apriori algorithm to analyze consumer purchasing patterns at MOY Frozen Food Supermarket. The process begins with the Data Collection stage, where all relevant sales transaction data is gathered. Following this, a Literature Review is conducted to understand previous studies and establish a theoretical foundation for the research. The next phase involves the Implementation of the Apriori Algorithm, which is divided into two key steps: Data Pre-processing and Data Mining. During data pre-processing, the collected data is cleaned and prepared for analysis to ensure accuracy and consistency. In the data mining step, the Apriori algorithm is applied to identify frequent item sets and extract meaningful association rules. The results of this analysis are then summarized in the Conclusion and Recommendations stage, providing insights and strategic suggestions for the supermarket. Finally, the process ends at the Completion stage, where the research findings are documented and reviewed.

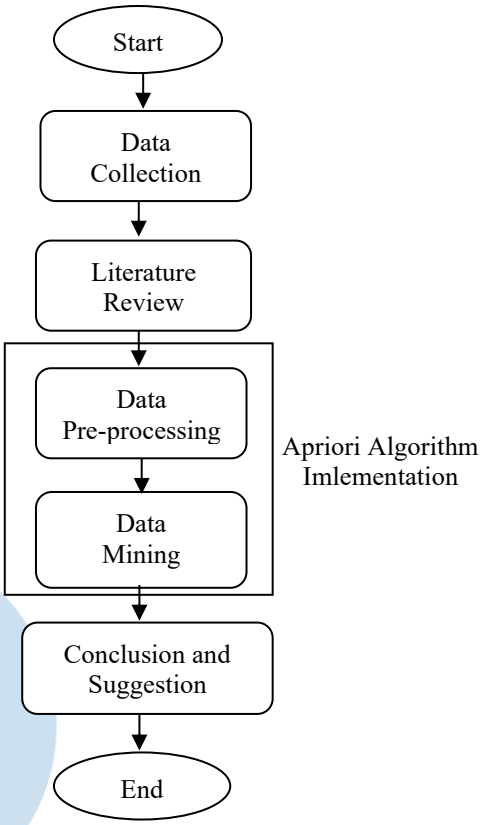


Figure 1. Research Flow

A. *Data Collection*

The data for this study was obtained from the transaction records of MOY Frozen Food Supermarket, located in Kediri Regency, East Java, covering the period from March 1, 2024, to March 13, 2024. A total of 2,544 rows of data were processed, encompassing several key variables: transaction number, transaction date, transaction day, transaction time, customer type, item code, item name, item quantity, and item category. This comprehensive dataset provides a detailed overview of consumer purchasing activities within the specified timeframe, offering a rich source of information to identify purchasing patterns and trends that can inform business strategy and decision-making at the supermarket.

B. *Literature Review*

At this stage, relevant literature sources are sought and gathered, including articles, journals, books, and other references related to the research, particularly those focusing on association methods using the Apriori algorithm. This comprehensive literature review helps to establish a solid theoretical foundation for the study, ensuring that the research is grounded in existing knowledge and methodologies. By examining previous studies and findings, the research can identify gaps in

the current literature and better position its contribution to the field of consumer behavior analysis using data mining techniques.

C. Implementasi Algoritma Apriori

The implementation of the Apriori algorithm is carried out using Python tools. The data processing stage is divided into two main phases: data pre-processing and data mining. Data pre-processing is a critical step in data mining analysis, aimed at cleaning [11], reformatting, and preparing the data for analysis, making it easier to handle and more accurate for subsequent steps. This stage involves removing any inconsistencies, errors, or irrelevant information to ensure that the dataset is of high quality and suitable for mining. On the other hand, data mining focuses on extracting previously unknown patterns or useful information from large datasets, transforming them into understandable and actionable insights that can be used to make crucial business decisions [11]. In the context of implementing the Apriori algorithm, several key terms are frequently used, including:

1. *Support*

Support represents the percentage of a particular item combination within a database [12]. The support value of an item is calculated using the following formula [13]:

$$Support(A) = \frac{Number\ of\ transactions\ containing\ item\ A}{Total\ number\ of\ transactions} \quad (1)$$

Similarly, the support value of 'n' items is determined using [13]:

$$Support(A \cup B) = \frac{Number\ of\ transactions\ containing\ both\ A\ and\ B}{Total\ number\ of\ transactions} \quad (2)$$

2. *Minimum support*

Minimum support is a process used to find all association rules that meet a specified minimum threshold for support [14]. It is expressed as a percentage or proportion of the total number of transactions. If an item set has a support value (frequency of occurrence) equal to or greater than the minimum support threshold, it is considered a frequent item sets. Minimum support helps filter out infrequent item sets, ensuring the algorithm only processes significant and relevant item sets.

3. *Support count*

Support count refers to the number of times a group of products appears together in the transaction data of a shopping cart [15]. It is a fundamental metric in the Apriori algorithm for determining how often a particular combination of items occurs together across all transactions. For example, if the item sets {Bread, Milk} appears in three out of five transactions, its support count is 3. The support count calculates the support value, which then helps identify frequent item sets that meet a certain threshold.

4. *Confidence*

Confidence measures the certainty or strength of the relationship between items in the Apriori algorithm [16]. A high confidence value indicates that item B is also likely to appear when item A appears, thereby validating the reliability of the association rule. Mathematically, confidence is defined as the ratio of the number of transactions that contain both the antecedent and the consequent to the number of transactions that contain only the antecedent. The formula for confidence is given by equation 3 [17].

$$Confidence(A \rightarrow B) = \frac{Support(A \cup B)}{Support(A)} \quad (3)$$

Where:

A is the antecedent (the item or set of items on the left-hand side of the rule).

B is the consequent (the item or set of items on the right-hand side of the rule).

5. *Minimum confidence*

The minimum confidence is the value that meets the minimum requirement for confidence [18]. This value measures how often item Y appears in transactions that contain item X. If the confidence of the rule X → Y is greater than or equal to the specified minimum confidence, then the rule is considered strong and significant. For example, if the minimum confidence is set at 70%, the rule X → Y will only be considered if at least 70% of the transactions containing X also contain Y.

6. *Itemsets*

An item sets is a set of items contained within a set processed by the system [17]. In the context of data mining and association analysis, item sets are used to discover patterns or relationships between items that frequently appear together in transaction datasets. An item sets can consist of a single item or a combination of multiple items.

7. *Frequent item set*

A frequent item set is a collection of items that meet a minimum support threshold. In the context of the Apriori algorithm, frequent item sets are used to identify patterns or relationships between items that frequently appear together in transactions. For example, if {Bread, Milk} often appears together in many transactions, {Bread, Milk} is considered a frequent item set. The discovery of frequent item sets is useful in various applications such as product recommendation, shopping basket analysis, and product placement optimization in retail stores.

8. *Lift ratio*

The Lift Ratio is a measure used to evaluate the strength of an association rule derived from support and confidence values [12]. The lift ratio indicates the likelihood that when the first item is purchased, the second item will also be purchased, considering the level of support. If the lift of {X → Y} = 1, there is no association. A lift value greater than 1 indicates that item Y is more likely to be purchased if item X is purchased, whereas a value less than 1 means that item Y is less likely to be purchased if item X is purchased. The lift ratio value is calculated using the formula below [17]:

$$Lift\ ratio = \frac{Confidence\ (X,Y)}{Benchmark\ Confidence\ (X,Y)} \quad (4)$$

$$Benchmark\ Confidence\ = \frac{Nc}{N} \quad (5)$$

Where:

N = the total number of transactions in the database.

Nc = the number of transactions that contain the consequent item.

9. *Threshold*

In the Apriori algorithm, the threshold is used to set the minimum support and minimum confidence that must be met by item sets and association rules [14]. This threshold value allows researchers to determine the level of importance and the strength of the associations generated by the algorithm. By doing so, researchers can filter out irrelevant results and focus on the most significant and valuable associations.

III. Result and Discussion

The Apriori algorithm can be applied to the sales transaction data of MOY supermarket, with the accumulated sales transactions obtained from daily sales over a period of 12 days through the following stages:

A. Data Pre-processing

The data used in this research consists of transaction data that has already been organized in Microsoft Excel. Before proceeding to the data mining stage, the data undergoes pre-processing. This involves correcting any spelling errors, standardizing the format in the item name column, and categorizing the item names into more general or common item groups.

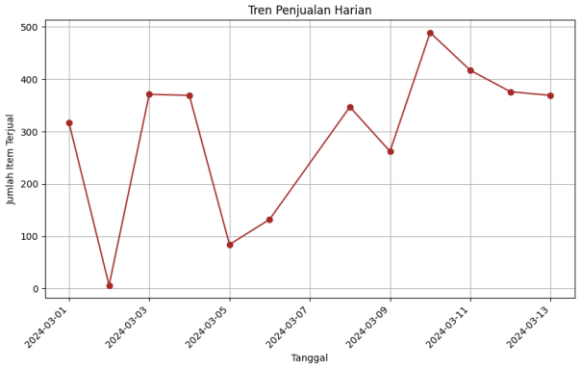


Figure 2. The Daily Sales Trends at Moy Supermarket

The graph in Figure 2. represents the daily sales trend at Moy Supermarket over a specified period. The x-axis indicates the dates, while the y-axis represents the number of items sold. The data points are connected by a line, showing the fluctuation on a sales basis. Based on Figure 2, there are noticeable ups and downs in daily sales, indicating varying consumer demand across the observed dates. On 2024-03-05, sales plummeted to a low point, reaching almost zero, suggesting an unusual drop in sales that day, which could have been due to operational challenges, lower consumer turnout, or external factors. The highest sales occurred on 2024-03-10, with a total of nearly 500 items sold. This spike could be attributed to specific promotions, events, or an increase in customer demand. Despite some sharp declines, the overall sales trend seems to rise over the observed period, suggesting a potential improvement in sales or better market conditions as the dates progress. While there are significant fluctuations in daily sales, the general upward trend provides a positive outlook for Moy Supermarket, highlighting potential opportunities for optimizing marketing strategies or better inventory management.

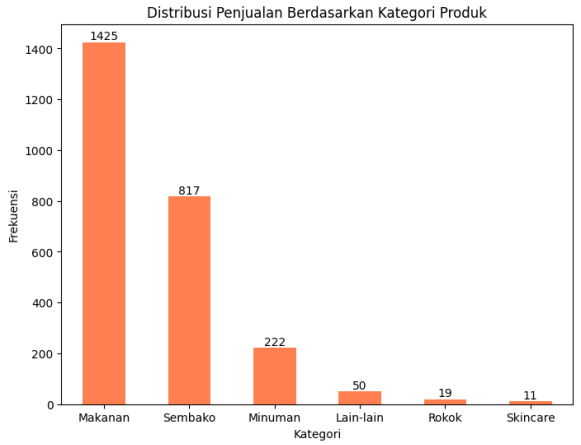


Figure 3. Sales Distribution by Product Category

The bar chart in Figure 3 illustrates the distribution of product sales at Moy Supermarket, categorized by product types. The x-axis represents the product

categories, while the y-axis shows the frequency or number of items sold within each category. The height of each bar reflects the total sales for the corresponding category. The highest sales were recorded in the "Makanan" (Food) category, with 1,425 items sold. This suggests that food products are the most popular and most frequently purchased by customers. The "Sembako" (Basic Necessities) category follows with 817 items sold, indicating significant demand for essential products like rice, sugar, and oil. The "Minuman" (Beverages) category comes next, with 222 items sold, showing a moderate level of sales compared to the top two categories. Categories like "Lain-lain" (Others), "Rokok" (Cigarettes), and "Skincare" show relatively low sales, with 50, 19, and 11 items sold, respectively. These Figures indicate that these product categories are less in demand or purchased less frequently by Moy's customers. The data reflects that food and basic necessities dominate Moy Supermarket's sales, while categories like beverages, skincare, and cigarettes contribute much less to overall sales. This insight can help the supermarket adjust its inventory and marketing focus on higher-demand products to maximize sales potential.

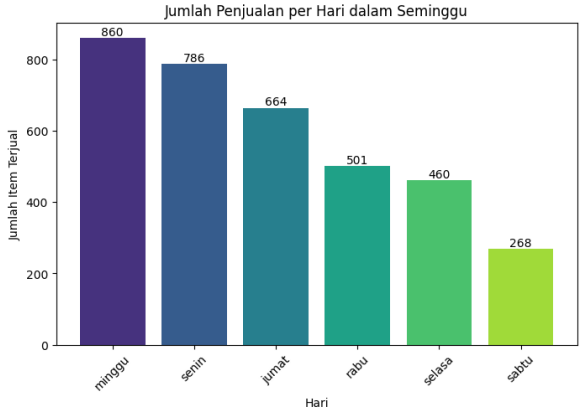


Figure 4. Total Sales per Day in a Week

The bar chart in Figure 4 illustrates the total number of items sold each day over a week at Moy Supermarket. The x-axis represents the days of the week, while the y-axis shows the number of items sold. The highest sales occurred on Sunday (Minggu), with 860 items sold, indicating that customer traffic is significantly higher on this day, possibly due to weekend shopping habits. Monday and Friday Follow Closely: Monday (Senin) and Friday (Jumat) also show substantial sales, with 786 and 664 items sold, respectively. This could indicate that shoppers are likely replenishing their supplies at the start of the week or preparing for the weekend. Sales gradually decrease as the week progresses. Wednesday (Rabu) recorded 501 items sold, followed by Tuesday (Selasa) with 460 items. These days show moderate but still steady sales activity. Saturday (Sabtu) has the lowest sales Figures, with only 248 items sold, which is in stark contrast to Sunday. This suggests that Saturday may not be a preferred shopping day for Moy's customers. The data suggests a clear pattern of higher sales on weekends, especially on Sundays, and a noticeable dip toward the end of the week, with Saturday being the least busy day. These trends could help the supermarket optimize staffing, promotions, and stock levels according to consumer shopping behaviors throughout the week.

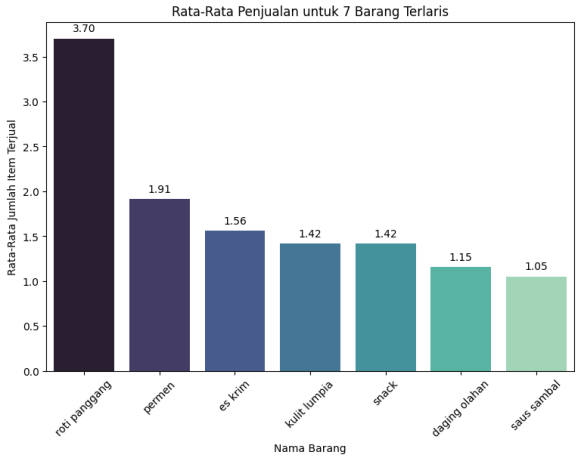


Figure 5. Average Sales for the 7 Best-Selling Products

The bar chart in Figure 5 displays the average number of items sold for the seven best-selling products at Moy Supermarket. The x-axis represents the product names, while the y-axis shows the average number of items sold per transaction. The product Roti Panggang (toast) stands out as the best-selling item with an average of 3.70 items sold per transaction. This suggests a high demand for this product, potentially making it a staple item for many customers. Following Roti Panggang are Permen (candy) with an average of 1.91 items sold and Es Krim (ice cream) with an average of 1.56 items sold. These items likely have steady demand but not as high as Roti Panggang. Kulit Lumpia (spring roll wrappers) and Susu (milk) both have an average of 1.42 items sold, showing moderate popularity. Meanwhile, Daging Cincang (minced meat) and Saus Sambal (chili sauce) have the lowest average sales with 1.15 and 1.05 items sold, respectively. The chart highlights that Roti Panggang is the top-performing product, far surpassing the other items in terms of average sales. Other products like Permen and Es Krim are also popular but are sold in fewer quantities per transaction. The insights from this data can help the supermarket focus on promoting high-performing products while adjusting inventory for lower-selling items.

B. Data Mining

1) Building Frequent Item's Results

At this stage, the researcher creates a variable consisting of several frequently purchased items. This variable is analyzed based on the transaction table in Python software using the Apriori command with a minimum support of 0.01 or 1%, resulting in item sets and their respective support values as shown in Table 1.

Table 1. Frequent Transaction Patterns of Customers

No.	Item sets	Support
1.	Frozenset({'baso aci', 'daging olahan'})	0,010256
2.	Frozenset({'mayones', 'saus sambal', 'daging olahan'})	0,010256
3.	Frozenset({'minyak goreng', 'daging olahan'})	0,010256
4.	Frozenset({'selai', 'roti panggang'})	0,010256
5.	Frozenset({'permen', 'jelly'})	0,010256
6.	Frozenset({'susu', 'permen'})	0,010256
7.	Frozenset({'susu', 'daging olahan', 'snack'})	0,011282
8.	Frozenset({'daging olahan', 'saus tomat'})	0,011282
9.	Frozenset({'bon cabe'})	0,011282
10.	Frozenset({'siomay'})	0,011282
11.	Frozenset({'bumbu', 'daging olahan'})	0,011282
12.	Frozenset({'mayones', 'saus sambal'})	0,012308
13.	Frozenset({'tusuk sate', 'daging olahan'})	0,012308
14.	Frozenset({'keripik'})	0,012308
15.	Frozenset({'permen', 'es krim'})	0,012308
16.	Frozenset({'daging olahan', 'yoghurt'})	0,013333
17.	Frozenset({'bumbu'})	0,013333
18.	Frozenset({'saus sambal', 'daging olahan', 'snack'})	0,013333
19.	Frozenset({'kulit lumpia', 'saus sambal'})	0,013333
20.	Frozenset({'tahu'})	0,014359
21.	Frozenset({'daging olahan', 'minuman rasa'})	0,014359
22.	Frozenset({'gula'})	0,014359
23.	Frozenset({'es lilin'})	0,014359
24.	Frozenset({'meses coklat 250gr'})	0,014359
25.	Frozenset({'es tube'})	0,015385
26.	Frozenset({'roti panggang', 'snack'})	0,015385
27.	Frozenset({'baso aci'})	0,015385
28.	Frozenset({'tusuk sate'})	0,015385
29.	Frozenset({'snack', 'es krim'})	0,01641
30.	Frozenset({'saus sambal', 'snack'})	0,01641
31.	Frozenset({'teh'})	0,01641
32.	Frozenset({'susu', 'snack'})	0,017436
33.	Frozenset({'donat', 'daging olahan'})	0,017436
34.	Frozenset({'saus tomat'})	0,018462
35.	Frozenset({'selai'})	0,018462
36.	Frozenset({'permen', 'daging olahan', 'snack'})	0,018462
37.	Frozenset({'masker'})	0,018462
38.	Frozenset({'jelly', 'daging olahan'})	0,019487
39.	Frozenset({'brownies', 'daging olahan'})	0,020513
40.	Frozenset({'mayones', 'daging olahan'})	0,021538
41.	Frozenset({'mie instan'})	0,021538
42.	Frozenset({'kopi'})	0,022564
43.	Frozenset({'roti panggang', 'daging olahan'})	0,022564
44.	Frozenset({'rokok'})	0,02359
45.	Frozenset({'yoghurt'})	0,024615
46.	Frozenset({'cireng', 'daging olahan'})	0,024615
47.	Frozenset({'mayones'})	0,025641
48.	Frozenset({'roti frozen', 'daging olahan'})	0,026667
49.	Frozenset({'minuman rasa'})	0,029744
50.	Frozenset({'cireng'})	0,029744
51.	Frozenset({'permen', 'snack'})	0,029744
52.	Frozenset({'brownies'})	0,030769
53.	Frozenset({'minyak goreng'})	0,030769
54.	Frozenset({'jelly'})	0,032821
55.	Frozenset({'kulit lumpia', 'daging olahan'})	0,032821
56.	Frozenset({'susu', 'daging olahan'})	0,034872
57.	Frozenset({'roti frozen'})	0,035897
58.	Frozenset({'donat'})	0,036923
59.	Frozenset({'permen', 'daging olahan'})	0,037949
60.	Frozenset({'daging olahan', 'es krim'})	0,050256
61.	Frozenset({'susu'})	0,058462
62.	Frozenset({'air mineral'})	0,066667
63.	Frozenset({'roti panggang'})	0,068718
64.	Frozenset({'daging olahan', 'snack'})	0,071795
65.	Frozenset({'permen'})	0,071795
66.	Frozenset({'saus sambal', 'daging olahan'})	0,08
67.	Frozenset({'es krim'})	0,094359
68.	Frozenset({'saus sambal'})	0,104615
69.	Frozenset({'kulit lumpia'})	0,121026
70.	Frozenset({'snack'})	0,140513
71.	Frozenset({'daging olahan'})	0,526154

Support value refers to the percentage of the average popularity of a product or item in the dataset. With a minimum support value of 0.01 or 1%, a buyer (consumer) who purchases one item or multiple items together will have a certain support value. For example, in the first row of Table 1, a consumer who buys 250g of chocolate sprinkles will only purchase that item, where 1.02% of the transactions in the database contain only the 250g chocolate sprinkles.

2) Association Rule

The determination of support and confidence in this study is based on a minimum threshold where the lift ratio metric is set to 1. A filter is then applied, requiring a minimum lift ratio of 1.5 and a minimum confidence level of 0.5 or 50%, which results in six association rules, as detailed in Table 2.

Table 2. Association Rules Result

No.	Antecedents	Consequents	Support
1.	Frozenset({‘Bumbu})	Frozenset({‘Daging Olahan’})	0.011282
2.	Frozenset({'Mayones’})	Frozenset({‘Daging Olahan’})	0.021538
3.	Frozenset({‘Saus Sambal, Mayones’})	Frozenset({‘Daging Olahan’})	0.010256
4.	Frozenset({‘Cireng’})	Frozenset({‘Daging Olahan’})	0.024615
5.	Frozenset({‘Snack, Saus Sambal’})	Frozenset({‘Daging Olahan’})	0.013333
6.	Frozenset({‘Tusuk Sate })	Frozenset({‘Daging Olahan’})	0.012308
7.	Frozenset({'Selai’})	Frozenset({‘Roti Panggang })	0.010256

Table 2 above presents the association rules derived from the dataset at Moy Supermarket. The goal is to identify products that are frequently purchased together. Each rule represents an antecedent (the product or set of products) and a consequent (the product likely to be purchased together). The support value quantifies the proportion of transactions that contain both the antecedent and the consequent.

Frequent Co-Purchases with Processed Meat: Several antecedents, such as Bumbu (seasoning), Mayones (mayonnaise), Saus Sambal (chili sauce), Cireng, Snack, and Tusuk Sate (satay skewers), lead to the purchase of Daging Olahan (processed meat). The support values for these associations range between 0.010256 and 0.024615, indicating that in 1.03% to 2.46% of transactions, these products are bought together with processed meat. Cireng has the highest support (0.024615), suggesting that customers who purchase Cireng are more likely to also buy processed meat. The combination of Snack and Saus Sambal has a support of 0.013333, highlighting a notable association with processed meat.

The combination of Saus Sambal and Mayones has a support value of 0.010256, indicating that around 1% of transactions involve the purchase of both sauces, which also leads to the purchase of processed meat. This may suggest that these sauces are often used together with processed meat for meal preparation.

There is a notable rule showing that customers who purchase Selai (jam) are likely to also buy Roti Panggang (toast). This rule has a support of 0.010256, meaning that in around 1.03% of transactions, both jam and toast are bought together. This reflects a common combination of items typically consumed as a breakfast or snack.

The association rules provide valuable insights into consumer behavior at Moy Supermarket. The rules suggest that processed meat is often bought alongside condiments and related items like Cireng and Tusuk Sate. Additionally, there is a strong relationship between Selai and Roti Panggang, reflecting common consumption patterns. These findings can help the supermarket optimize product placements, bundle promotions, or cross-selling strategies to enhance sales.

Table 3. Confidences Association Rules

Antecedents	Consequents	Confidence
frozenset({'cireng'})	frozenset({'daging olahan'})	0,85185
frozenset({'saus sambal', 'snack'})	frozenset({'daging olahan'})	0,83333
frozenset({'mayonais'})	frozenset({'daging olahan'})	0,76471
frozenset({'saus sambal'})	frozenset({'daging olahan'})	0,75949
frozenset({'ilm tempura 500gr '})	frozenset({'daging olahan'})	0,73077
frozenset({'selai'})	frozenset({'roti frozen'})	0,55556
frozenset({'permen', 'daging olahan'})	frozenset({'snack'})	0,5

Table 3 is the result of association rule mining which displays the relationship between products in purchase transactions. Each line shows the association rule with antecedents (initial purchase items), consequents (items that tend to be purchased together), and confidence ("if-then" condition probability). The first frozenset rule showed that if a customer bought Cireng, there was an 85.19% chance that they would also buy Daging Olahan (processed meat), indicating a very strong relationship. A similar pattern can be seen in the combination of Saus Sambal and Snack that have a high impetus for the purchase of Daging Olahan. Meanwhile, the rules of frozenset ({'mayonnaise'}) → frozenset ({'daging olahan'}) with a confidence of 76.47% and frozenset ({'saus sambal'}) → frozenset ({'daging olahan'}) with a confidence of 75.95% still showed a significant correlation although slightly lower. On the other hand, Selai and Roti Frozen with a confidence of 55.56% revealed a moderate relationship, while frozenset({'permen', 'daging olahan'}) → frozenset ({'snack'}) with a confidence of 50% indicated a weaker relationship, where only half of the transactions followed this pattern. This analysis is useful for marketing strategies such as placing products adjacent to each other or bundling discounts based on dominant purchasing patterns.

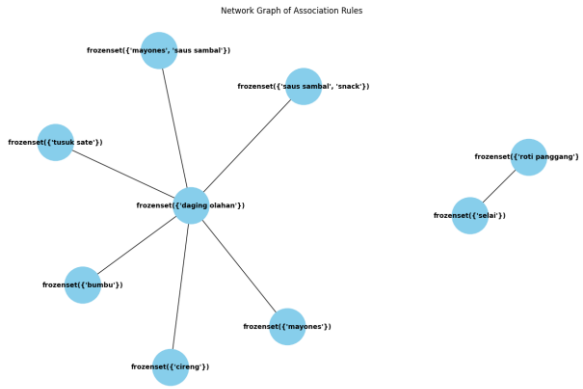


Figure 6. *Network Graph of Association Rules*

The image in Figure 6 above represents a network graph of the association rules derived from consumer transaction data at Moy Supermarket. The graph visualizes the relationships between products (as nodes) based on the association rules, with lines (edges) indicating co-purchase patterns. Daging Olahan (processed meat) appears as the central node in the graph with multiple connections to other product sets. This indicates that processed meat is frequently purchased together with various other products. Products like Bumbu (seasoning), Mayones (mayonnaise), Saus Sambal (chili sauce), Tusuk Sate (satay skewers), Cireng, and the combination of Snack with Saus Sambal all connect to processed meat, reinforcing the high likelihood of co-purchases involving these products.

The combination of Saus Sambal and Mayones forms a node that links to Daging Olahan, suggesting that these two products are often bought together, potentially as accompaniments to processed meat. Another notable combination is Snack and Saus Sambal, which also connects to processed meat. In a separate cluster, Roti Panggang (toast) is connected with Selai (jam), indicating a strong association between these two products. This cluster is isolated from the processed meat cluster, showing that these products form their own distinct co-purchase pattern.

The network graph effectively visualizes the strong relationships between processed meat and various condiments or side dishes, reflecting consumer preferences for meal preparation. The connection between Selai and Roti Panggang in a separate cluster suggests a different type of consumer behavior, likely centered around breakfast or snack items. These insights can help Moy Supermarket strategically organize product placements and cross-promotions, enhancing customer experience and boosting sales through co-purchase tendencies.

IV. CONCLUSIONS

A. *Conclusions*

Based on the analysis conducted in this study, the following conclusions can be drawn:

1. The association rule-mining method using the Apriori algorithm can help supermarkets identify consumer shopping patterns.
2. The results of the Apriori algorithm analysis using a dataset of 2,544 transactions produced the following association rules:
   - When customers buy spices, they also purchase processed meat, and they are 85% confident in the products they buy simultaneously.
   - When customers buy mayonnaise, they purchase preserved processed meat with reliable product quality, and the likelihood of buying these products together increases by 84%.
   - When customers buy chilli sauce and mayonnaise, they purchase processed meat with an 83% confidence level in what product they will buy simultaneously.
   - When customers buy cireng (fried dough), they purchase snacks and processed meat with an 83% confidence level for buying these products together.
   - When customers buy snacks and chili sauce, they purchase processed meat with an 81% confidence level regarding the products they buy simultaneously.
   - When customers buy skewers, they purchase processed meat with an 80% confidence that these products will be purchased together.
   - When customers buy jam, they purchase frozen or toasted bread with a 55% confidence level regarding the products they will buy simultaneously.

B. *Recommendations*

After identifying consumer shopping patterns, MOY Supermarket can stock more items such as spices, chili sauce, mayonnaise, snacks, jam, frozen bread, cireng (fried dough), skewers, and processed meat to avoid stock shortages. On the other hand, MOY Supermarket should reduce stock for items that are rarely purchased to prevent overstocking. MOY Supermarket can implement sales strategies such as offering discounts or promotions on product bundles that are likely to be purchased together and offering complementary products.

Recommendations for future research include using a larger dataset and applying the Apriori algorithm with a higher minimum support threshold. Future studies are also encouraged to use other algorithms for association rule-mining and compare them to determine which algorithm is more accurate.

REFERENCES

[1] D. Listriani, A. H. Setyaningrum, and F. Eka, "Penerapan Metode Asosiasi Menggunakan Algoritma Apriori Pada Aplikasi Analisa Pola Belanja Konsumen (Studi Kasus Toko Buku Gramedia Bintaro)," *Jurnal Teknik Informatika*, vol. 9, no. 2, 2016.

[2] Irma T, "PENERAPAN DATA MINING DALAM DUNIA BISNIS MENGGUNAKAN METODE CLUSTERING," 2019.

[3] R. Takdirillah, "Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Penjualan Bisnis Ritel," *Edumatic: Jurnal Pendidikan Informatika*, vol. 4, no. 1, pp. 37–46, 2020.

[4] V. G. Shabrina, "Pengaruh Revolusi Digital terhadap Pemasaran dan Perilaku Konsumen," *Jurnal Pewarta Indonesia*, vol. 1, no. 2, pp. 131–141, 2019.

[5] W. Ardani, "Pengaruh Digital Marketing Terhadap Perilaku Konsumen," *Jurnal Tadbir Peradaban*, vol. 2, no. 1, pp. 40–47, 2022.

[6] R. Srikant, Q. Vu, and R. Agrawal, "Mining association rules with item constraints.," in *Kdd*, 1997, pp. 67–73.

[7] M. Badrul, "Algoritma asosiasi dengan algoritma apriori untuk analisa data penjualan," *Jurnal Pilar Nusa Mandiri*, vol. 12, no. 2, pp. 121–129, 2016.

[8] L. C. Dewi, E. Irawan, and R. K. Sormin, "Penerapan Algoritma Apriori Terhadap Data Mining Penjualan Di Alfamart Berastagi," *FATIMAH: Penerapan Teknologi dan Sistem Komputer*, vol. 1, no. 1, pp. 38–47, 2022.

[9] S. Dewi, "Sosialisasi Digitalisasi Dan Inovasi Usaha Mikro Kecil Dan Menengah (Umkm) Kerupuk Rajungan 'Sari Mina' Di Desa Sukajaya Kecamatan Cilamaya Kulon," *Prosiding Konferensi Nasional Penelitian Dan Pengabdian Universitas Buana Perjuangan Karawang*, vol. 3, no. 1, pp. 1261–1272, 2023.

[10] S. Al Farisi and M. I. Fasa, "Peran UMKM (Usaha Mikro Kecil Menengah) dalam Meningkatkan Kesejahteraan Masyarakat," *Jurnal Dinamika Ekonomi Syariah*, vol. 9, no. 1, pp. 73–84, 2022.

[11] C. Zai, "Implementasi Data Mining Sebagai Pengolahan Data," *Jurnal Portal Data*, vol. 2, no. 3, 2022.

[12] A. N. Rahmi and Y. A. Mikola, "Implementasi Algoritma Apriori Untuk Menentukan Pola Pembelian Pada Customer (Studi Kasus: Toko Bakoel Sembako)," *Information System Journal*, vol. 4, no. 1, pp. 14–19, 2021.

[13] N. N. Merliani, N. I. Khoerida, N. T. Widiawati, L. A. Triana, and P. Subarkah, "Penerapan Algoritma Apriori Pada Transaksi Penjualan Untuk Rekomendasi Menu Makanan Dan Minuman," *J. Nas. Teknol. dan Sist. Inf*, vol. 8, no. 1, pp. 9–16, 2022.

[14] A. W. O. Gama, I. Putra, and I. P. A. Bayupati, "Implementasi Algoritma Apriori Untuk Menemukan Frequent Item sets Dalam Keranjang Belanja," *Maj. Ilm. Teknol. Elektro*, vol. 15, no. 2, pp. 21–26, 2016.

[15] R. Wahyusari, "Penerapan Algoritma Apriori Untuk Menemukan Pola Peminjaman Buku Di Perpustakaan (Studi Kasus: Sekolah Menengah Atas Negeri 2 Cepu)," *SIMETRIS*, vol. 16, no. 1, pp. 17–23, 2022.

[16] E. Alma'arif, E. Utami, and F. W. Wibowo, "Implementasi Algoritma Apriori Untuk Rekomendasi Produk Pada Toko Online," *Creative Information Technology Journal*, vol. 7, no. 1, p. 63, Mar. 2021, doi: 10.24076/citec.2020v7i1.241.

[17] M. Fauzy and I. Asror, "Penerapan metode association rule menggunakan algoritma apriori pada simulasi prediksi hujan wilayah kota bandung," *Jurnal Ilmiah Teknologi Infomasi Terapan*, vol. 2, no. 3, 2016.

[18] S.-H. Bae and K.-J. Yoon, "Confidence-based data association and discriminative deep appearance learning for robust online multi-object tracking," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 3, pp. 595–610, 2017

ISSN 2085-4552

Comparison of Cosine Similarity, Rabin-Karp, and Levenshtein Distance Algorithms for Plagiarism Detection in Document

Jasman Pardede[1], Agil Yudistira[2]
[1,2] Deparment of Informatics, Intitut Teknologi Nasional (Itenas), Bandung, Indonesia
[1]jasman@itenas.ac.id, [2]agil.yudistira@outlook.com

Accepted 13 June 2025
Approved 24 June 2025

***Abstract*— Prevention and detection of plagiarism are crucial. There are several algorithms that can be used to detect plagiarism in documents, including the Cosine Similarity, Rabin-Karp, Levenshtein Distance, Hamming Distance, Euclidean Distance, Edit Distance, Jaccard Similarity, Ratcliff/Obershelp, Winnowing, Brute Force, Boyer-Moore, and Knuth Morris Pratt algorithms. Based on the literature review from previous research, three best algorithms were identified: Cosine Similarity, Rabin-Karp, and Levenshtein Distance. However, there has been no study analyzing the comparison of these three algorithms. Therefore, this study will compare the performance of each algorithm and determine the best algorithm for plagiarism detection in documents based on similarity scores and execution time. The research objects use a sample of documents consisting of titles and abstracts from Indonesian-language informatics journals. Cosine Similarity is superior to others for plagiarism detection in documents, as it produces the highest average similarity score with a relatively fast execution time. The similarity values obtained using Cosine Similarity, Rabin-Karp 4-grams, and Levenshtein Distance were found to be 48.80%, 47.13%, and 20.61%, respectively. The average execution time of Cosine Similarity, Rabin-Karp with 4-grams, and Levenshtein Distance are 0.22 s, 0.45 s, and 39.15 s, respectively.**

***Index Terms*— *comparison; cosine similarity; rabin-karp; levenshtein distance; plagiarism; similarity*.**

I. INTRODUCTION

Plagiarism constitutes a substantial issue within academic and professional contexts, as it compromises the authenticity and credibility of scholarly endeavors. Plagiarism refers to the practice of taking, utilizing, or copying the work, concepts, or discoveries of another individual, either in full or in part, without sufficient attribution or appropriate referencing, thereby misrepresenting it as one's own intellectual product or achievement [1], [2]. Prevention and detection of plagiarism are crucial, especially in the academic world [2], [3]. With the ease of information exchange through the internet, academic community comprising students and lecturers can engage in copy-paste practices that may lead to plagiarism.

According to Regulation of the Minister of Education, Culture, Research, and Technology Number 39 of 2021 regarding Academic Integrity in Producing Scientific Works, article 10 paragraph (3), Plagiarism is the act of taking and rewriting part or all of someone else's work without using one's own language, even if the source is cited correctly [2],[5]. Furthermore, in the Tridharma of Higher Education, academic community members are also obligated to conduct research and produce scientific works [5]. Therefore, it is crucial to prevent and identify acts of plagiarism [2],[3].

TABLE I. LITERATURE REVIEW

Researcher	Year	Algorithms
Alvi, F. et al.	2017	Hamming Distance
Hartanto, A. D., et al.	2019	Rabin-Karp
Süzen, N., et al.	2020	Euclidean distance, Needleman-Wunsch Distance
Wahyuningsih, T., et. al.	2021	Cosine Similarity, Jaccard Similarity
Alobed, M., et al.	2021	Cosine Similarity, Jaccard Similarity
Nalawati, R.E., & Yuntari, A.D.	2021	Ratcliff/Obershelp
Astuti, Y., & Wulandari, I.,	2022	Rabin-Karp
Hidayat, W., et al.	2022	Rabin-Karp
Al-Hagree, S., & Al-Gaphari, G.	2022	Levenshtein Distance
Nandurkar, D. A.,et al.	2023	Levenshtein Distance
Amalia, E.L., et al.	2023	Winnowing
Alfat, L., et al.	2023	Knuth Morris Pratt
Barut, Z. & Altuntas, V.	2023	Boyer-Moore
Setu, D.M., et al.	2025	Cosine Similarity
Madhan, N., et al.	2025	Manhattan Distance

One way to detect plagiarism is by using the concept of string matching to calculate the similarity between documents [1],[3],[6],[7]. To achieve the best results, the use of the best algorithm is essential for detecting plagiarism in documents. There are several algorithms that can be used for plagiarism detection in document, including Cosine Similarity [8]-[10], Jaccard Similarity [8],[9], Rabin-Karp [11]-[13], Levenshtein Distance [14],[15], Hamming Distance [16], Euclidean distance [17], Winnowing [18], Knuth Morris Pratt [19], Ratcliff/Obershelp [20], Needleman-Wunsch Distance [17], Boyer-Moore [21], Manhattan Distance [22], and others. Several previous studies have been conducted using those algorithms. A literature review of research relevant to those algorithms is presented in Table I. Among those methods, three were selected based on their consistent superiority in various aspects of plagiarism detection, as explained in the literature review.

Cosine Similarity with pre-processing is superior to Jaccard Similarity in measuring similarity because it produces the highest correlation values [8]. Cosine Similarity has a better accuracy rate compared to the Euclidean Distance algorithm for calculating similarity [9]. Although the Rabin-Karp algorithm exhibits limitations in accurately distinguishing between similar words, it demonstrates superior performance in plagiarism detection when compared to the Brute Force and Boyer-Moore algorithms [23]. Specifically, while the Brute Force algorithm excels in single-pattern searches, it remains less effective for multiple-pattern searches. Additionally, the Boyer-Moore algorithm performs efficiently by shifting the last two characters but is less effective with earlier shifts. In contrast, the Rabin-Karp algorithm effectively addresses these limitations [24].

The Rabin-Karp algorithm can be used for searching long pattern strings but has a longer execution time compared to Ratcliff/Obershelp. Similarity values are highly influenced by the sentence structure in Rabin-Karp, whereas it has no impact at all on Ratcliff/Obershelp [11]. From the testing of Rabin-Karp with ***k***-grams ranging from 2 to 10, the one with the highest accuracy is ***3***-gram [13]. The Rabin-Karp algorithm is better than the Winnowing for detecting plagiarism because it produces higher similarity scores. Therefore, for this research, the Rabin-Karp algorithm is proposed for detecting plagiarism.

The Levenshtein Distance algorithm, which checks each character one by one, produces perfect results on simple short sentences but does not perform well on long documents with irregular positions [25]. The Levenshtein Distance algorithm is better than the Knuth Morris Pratt algorithm because it has higher speed and accuracy and can minimize errors when searching for data [26]. The Levenshtein Distance algorithm is superior and more efficient compared to Ratcliff/Obershelp for detecting plagiarism [27].

Based on the literature review, there are three best algorithms: Cosine Similarity, Rabin-Karp, and Levenshtein Distance. These algorithms have their respective strengths and weaknesses, but there has been no research analyzing the comparison of these three algorithms for detecting plagiarism in documents. Therefore, a comparative analysis among these algorithms is needed to determine the best algorithm by measuring the level of similarity for detecting plagiarism in documents and execution time as a parameter to assess the speed of each algorithm.

Therefore, the purpose of this research is to compare the similarity scores and execution times to know the performance of each algorithm and determine the best algorithm for plagiarism detection in documents based on similarity scores and execution times. This study is a novelty compared to previous research, as there has been no study comparing the Cosine Similarity, Rabin-Karp, and Levenshtein Distance algorithms for plagiarism detection in documents.

II. METHODOLOGY

The research method used is a comparative analysis, where this study will compare performance of Cosine Similarity, Rabin-Karp, and Levenshtein Distance algorithms for detecting plagiarism in documents. The research begins with data collections, followed by comparison document collections processing. After that, testing is conducted using the Cosine Similarity, Rabin-Karp, and Levenshtein Distance algorithms. Once the testing is completed, an analysis will be performed to determine the best algorithm based on similarity scores and execution time. The block diagram of this research can be seen in Fig. 1.

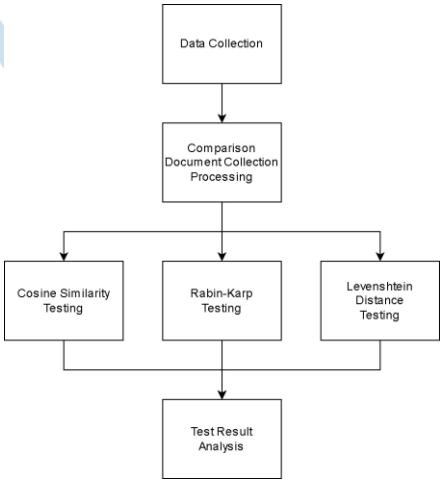


Fig. 1. Block diagram system

A. Data Collections Technique

The data collections technique used for testing in this research is by sampling documents consisting of titles and abstracts from Indonesian-language journals. The chosen journals are in the field of informatics, particularly those with high citation numbers and

publication years ranging from 2018 to 2023. This is to obtain data that is relevant to the research objectives and of good quality. The process of collecting sample documents is conducted through Google Scholar using the keywords {"Information Technology," "Artificial Intelligence," "Machine Learning," "K-Means Clustering," and "Naive Bayes"}. Then, the titles and abstracts from each journal will be extracted and saved in *.txt* file format to facilitate data pre-processing.

The total sample documents in this research amount to 440 documents. The 440 documents were then divided into two groups: 400 documents were used as the comparison document collections (reference corpus) without specific labeling, which serve as the base dataset for similarity comparison, while 40 documents were designated as test documents. The test documents were selected randomly from the collected sample to evaluate the performance of the plagiarism detection algorithms. Since all documents are from the same domain and collected via the same keywords, this setup simulates real-world scenarios of document similarity detection without explicit labeling of plagiarism cases.

B. Data Pre-processing

Pre-processing is performed as part of data cleansing to simplify and standardize the text, allowing it to be processed more effectively in the main process [28]. The pre-processing conducted in this research consists of only two stages, namely Case Folding and Filtering.

The case folding stage is performed to eliminate differences in letter case, thus converting the entire text to lowercase. The filtering stage is carried out to enhance accuracy and speed by focusing on more significant words. In this study, the filtering process does not remove stop words because, based on the adopted definition of plagiarism, paraphrasing is considered important. Therefore, only punctuation marks and special characters will be removed. The flowchart of the data pre-processing can be seen in Fig. 2.

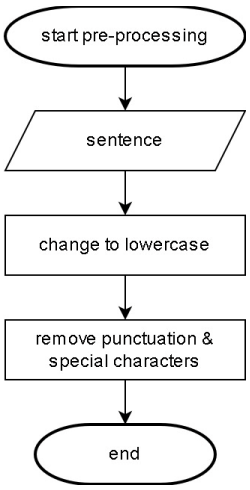


Fig. 2. Flowchart of data pre-processing

C. Comparison Document Collections Processing

In the testing process, the test documents will be compared with the collections of comparison documents, while the text content in each document must undergo parsing or text segmentation based on sentences and pre-processing. The purpose of processing the collections of comparison documents is to avoid repetitive parsing and pre-processing of the document collections every time a test is conducted.

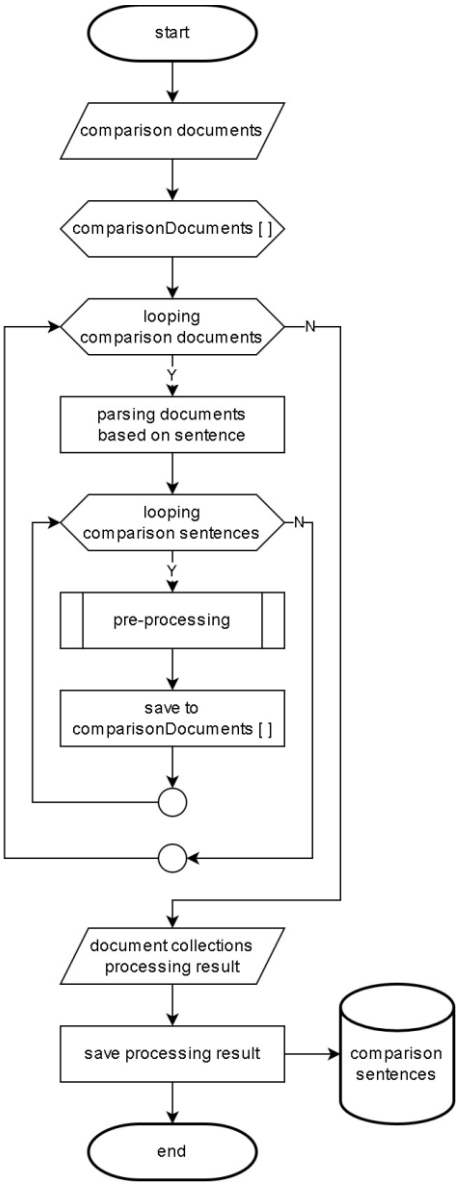


Fig. 3. Flowchart of comparison document collections processing

The system will read the folder containing the collections of comparison documents. Each document will undergo parsing to separate the text based on sentences, and each sentence will undergo pre-processing. Then, the processed sentences will be stored in the variable "comparisonDocuments" with the document index and its corresponding sentence. Once all the sentences in the first document are processed, the system will proceed to the next document until all documents in the folder are processed. After that, the

results of the comparison document collections processing will be stored in the comparison sentence database, allowing access for each testing session. The flowchart of the comparison document collections processing can be seen in Fig. 3.

D. Algorithm Testing Technique

The purpose of this testing is to obtain the similarity scores for each test sentence with all the sentences in the comparison document collections. Then, the highest similarity score for each test sentence will be determined to identify which comparison document sentence it is most similar to and at which sentence position. The average similarity score for each test sentence will be calculated to determine the similarity value between the test document and the entire collections of comparison documents.

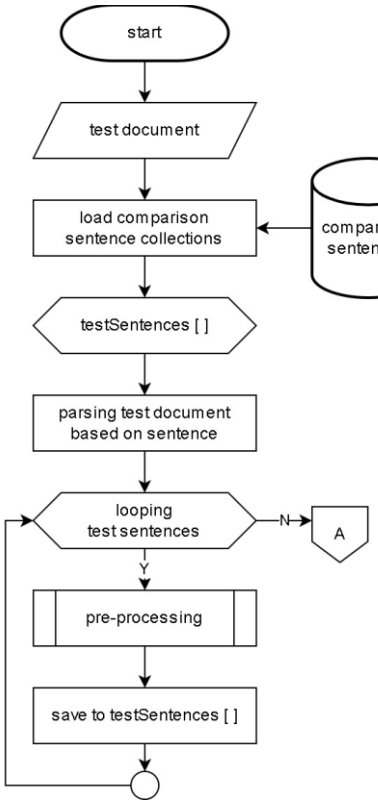


Fig. 4. Flowchart of algorithm testing

The similarity scores between the test documents generated by each algorithm will be compared to determine which algorithm produces the highest similarity score. Additionally, the execution time of each algorithm will also be measured to determine which algorithm is the fastest in its usage. The testing phase will be conducted using 40 test documents. Each algorithm will be applied to the same test documents. Analyze each of the results. Comparing the performance of each algorithm in detecting plagiarism.

In this study, testing will be conducted one by one in sequential order for each algorithm, namely: (1) Cosine Similarity, (2) Levenshtein Distance, (3) Rabin-Karp 2-gram, (4) Rabin-Karp 4-gram, and (5) Rabin-Karp 6-gram. The selection of these *k*-gram values is based on previous studies that have also used similar *k*-gram values [13], to observe the impact of *k*-gram size on Rabin-Karp algorithm performance. Therefore, one document will not be tested using all algorithms simultaneously. This means that one test document will be repeatedly used in five different tests. As a result, out of 40 test documents, a total of 200 tests will be conducted. The testing flowcharts can be seen in Fig. 4 to Fig. 6.

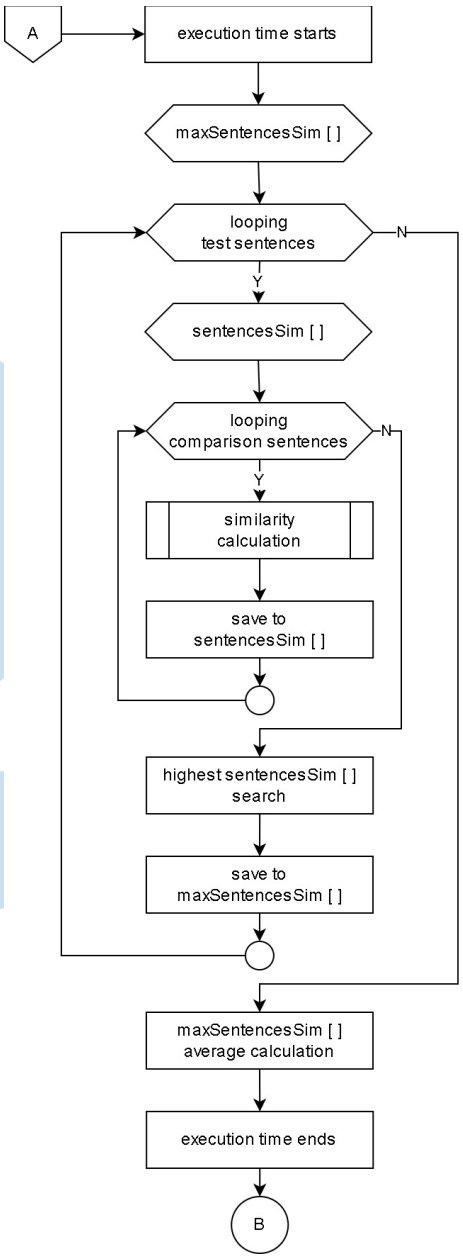


Fig. 5. Flowchart of algorithm testing (continued)

The testing begins by uploading the test documents. First, the system will load the comparison sentence collections from the database. Then, the test document will be parsed based on sentences, and each test sentence will undergo pre-processing and will be saved in the variable "testSentences". This is done to facilitate the execution time calculation, as the time will be

measured only during the algorithm testing process. Once all test sentences have completed pre-processing, the next stage will proceed to the main testing process. The flowchart of the main testing process can be seen in Fig. 5.

At this stage, the execution time calculation begins. Each test sentence will be compared with all the comparison sentences. The first test sentence will be compared with the first comparison sentence to calculate the similarity. As the testing is conducted one by one for each algorithm, at this stage of similarity calculation, one algorithm function will be called based on the selected algorithm testing.

The similarity calculation results will be stored in the "sentencesSimilarity" variable along with the index of the test sentence, the comparison document, and the comparison sentence. This stage will be repeated continuously until the first test sentence is compared with all the comparison sentences. After that, the highest similarity of the first test sentence will be determined and saved in the variable "maxSentencesSimilarity". This is intended to determine which test sentence has the highest similarity with which comparison sentence in which comparison document. If the first test sentence has been compared with all the comparison sentences and the index with the highest similarity value is obtained, the process will be repeated for the second test sentence until all the test sentences are compared with all the comparison sentences.

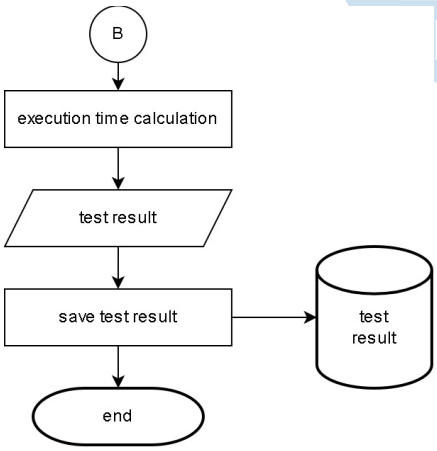


Fig. 6. Flowchart of algorithm testing (continued)

Afterward, the similarity score for the test document will be determined by calculating the average of the highest similarity scores from all test sentences. This stage is the end of the main testing, thus the execution time calculation is completed at this stage. The process proceeds to the final stage, which can be seen in Fig. 6.

The final stage is the execution time calculation, which is obtained by subtracting the end execution time from the start execution time. After that, all test results will be saved to the test result database. These tests are conducted on each algorithm.

E. Cosine Similarity Algorithm

The working principle of the Cosine Similarity algorithm is to measure the proximity between two vectors by calculating their dot product and then dividing it by the Euclidean distance between the two vectors for normalization. Equation (1) represents the formula for Cosine Similarity used to calculate the similarity score [9],[29].

$$S_{(A,B)} = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} {A_i}^2}\sqrt{\sum_{i=1}^{n} {B_i}^2}} \times 100 \quad (1)$$

S represents the similarity score, ***A*** and ***B*** are vector representations of two documents being compared, where each vector represents the term frequency (TF) values of words in the documents, ***n*** is the number of vectors, and the ***i*** is the vector order.

F. Rabin-Karp Algorithm

The working principle of the Rabin-Karp algorithm is string searching based on patterns (substrings) and comparing the hash values of each pattern [23]. First, the text will be divided into grams based on the value of k to obtain substrings. Then, the substrings will be converted into hash values, and the hash values between two texts will be compared. If both hash values are the same, the comparison will be done once again on the characters of the substring to ensure that the hash value represents the same substring. If they are not the same, the substring will be shifted to the right [13]. Afterward, the similarity score is calculated using the Dice Similarity Coefficient equation. Equation (2) represents the formula used in the Rabin-Karp algorithm to calculate the hash value.

$$H_{(p)} = \sum_{i=1}^{k} ascii(p_i)\ b^{(k-1)} \quad (2)$$

H represents the hash value, ***p*** is the pattern, ***k*** is the value of ***k***-gram, ***b*** is the base or radix value (a prime number), and ***i*** is the character order in the pattern. Equation (3) is the Dice Similarity Coefficient formula used for calculating the similarity in the Rabin-Karp algorithm [12].

$$S_{(A,B)} = \frac{2 \times |H_A \cap H_B|}{|H_A| + |H_B|} \quad (3)$$

S represents the similarity score, ***A*** and ***B*** are the texts being compared, and ***H*** is the hash. Example calculations using the Rabin-Karp algorithm can be found in [11],[23].

G. Levenshtein Distance Algorithm

The working principle of the Levenshtein Distance algorithm is by creating a matrix to calculate the number of operations required to change one character using addition, deletion, or substitution operations. This calculation is done by comparing each character [14].

However, based on the research conducted [18], distance calculations can be done by directly calculating the word differences without creating a matrix, making the steps in the distance calculation simpler and can improve execution time. Equation (4)

represents the formula used for calculating the Levenshtein distance value [15].

$$D_{(A,B)} = \sum_{i=1}^{n} d(A_i, B_i) \quad (4)$$

D is the Levenshtein distance, ***A*** and ***B*** are the texts being compared, ***d*** is the comparison value (if equal = 0, if different = 1), ***n*** is the maximum number of texts between ***A*** and ***B***, and the ***i*** is the word order in the text. Equation (5) represents the Levenshtein Distance formula for calculating the similarity.

$$S_{(A,B)} = \left(1 - \frac{D_{(A,B)}}{n}\right) \times 100 \quad (5)$$

S is the similarity value, ***D*** is the Levenshtein distance, ***A*** and ***B*** are the texts being compared, and ***n*** is the maximum number of texts between ***A*** and ***B***. Examples of calculations using the Levenshtein Distance algorithm can be found in [18].

H. Test Result Analysis Technique

The data analysis technique used in this study is a comparative analysis, where the test results of each algorithm will be presented in the same table, allowing the performance of each algorithm to be directly compared. Additionally, the average similarity values and execution times of each algorithm will be presented in the form of bar charts.

III. Result And Discussion

A. Experimental Result

After comparison document collections processing is completed, out of the 400 documents used, a total of 3,471 comparison sentences were obtained. Therefore, in this study, each test sentence will be compared and its similarity calculated with the 3,471 comparison sentences. Table II shows the test results on the same 10 test documents using Cosine Similarity, Levenshtein Distance, and Rabin-Karp algorithms

TABLE II. Experimental Result

No	Document	Cosine Similarity		Levenshtein Distance		Rabin-Karp		
		Similarity (%)	Time (s)	Similarity (%)	Time (s)	k	Similarity (%)	Time (s)
1	2020-403 TI	49.63	0.50	15.29	0.23	2 4 6	89.71 45.69 29.28	24.74 51.95 62.37
2	2021-10 ML	53.81	0.36	22.88	0.17	2 4 6	89.77 52.12 33.16	19.00 40.07 48.06
3	2021-14 KM	47.09	0.23	14.94	0.12	2 4 6	89.00 47.33 29.52	13.07 26.15 32.10
4	2021-84 NB	53.01	0.52	18.76	0.22	2 4 6	86.14 44.63 30.72	22.66 41.39 49.02
5	2021-86 TI	41.86	0.40	17.82	0.18	2 4 6	84.90 42.28 27.04	17.47 31.98 38.40
6	2021-105 NB	45.77	0.54	18.34	0.26	2 4 6	84.46 41.80 27.40	24.42 44.66 52.56
7	2021-145 NB	53.18	0.46	25.23	0.22	2 4 6	84.06 46.41 31.79	19.40 35.35 42.10
8	2021-155 NB	48.89	0.50	21.47	0.24	2 4 6	83.82 45.07 28.40	20.72 38.51 45.63
9	2021-170 NB	48.30	0.53	26.33	0.24	2 4 6	83.36 48.50 33.97	22.41 40.35 48.68
10	2021-198 TI	38.53	0.61	10.61	0.29	2 4 6	82.96 39.44 20.03	29.48 56.42 68.06

The test result using the Cosine Similarity algorithm resulted in an average similarity of 48.80% and an execution time of 0.45 seconds. The test result using the Levenshtein Distance algorithm resulted in an average similarity of 20.61% and an execution time of 0.22 seconds. The test result using the Rabin-Karp algorithm with 2-grams resulted in an average similarity of 83.76% and an execution time of 19.83 seconds. With 4-grams, the average similarity was 45.96% and the execution time was 37.71 seconds. Meanwhile, with 6-grams, the average similarity was 30.76% and the execution time was 45.70 seconds. The comparison of

average similarity values and execution times of each algorithm is presented in Table III.

TABLE III. COMPARISON OF SIMILARITY AND EXECUTION TIME

Cosine Similarity	Levenshtein Distance	Rabin-Karp		
		2	4	6
48.80 %	20.61 %	86.21 %	47.13 %	31.40 %
0.45 s	0.22 s	20.74 s	39.15 s	47.52 s

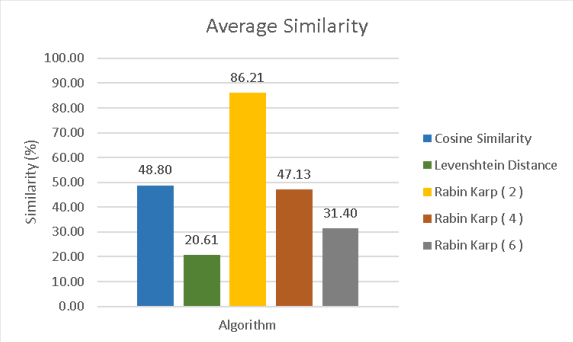


Fig. 7. Comparison diagram of similarity

To facilitate the analysis process, the test results are also presented in the form of diagrams. The diagram comparing the average similarity values can be seen in Fig. 7, and the diagram comparing the average execution time can be seen in Fig. 8. Based on the test results conducted on the same test document, it shows that the Rabin-Karp method with 2-grams has a greater similarity. This is because the test document used is an abstract document that has a word length of between 150 and 350 words. When compared to the time requirements, Rabin-Karp with 2-grams is faster than 3-grams and 4-grams. This is because Rabin-Karp 2-grams only has 2 characters, resulting in the time required to calculate each hash much faster than 3-grams and 4-grams.

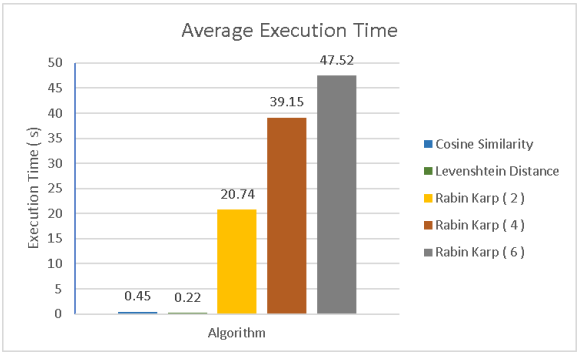


Fig. 8. Comparison diagram of execution time

Cosine Similarity demonstrated the second-best performance following Rabin-Karp 2-gram, with lower computational time. Although the time required for Cosine Similarity is less than that of Rabin-Karp 2-gram—attributable to Cosine Similarity's use of vector space models for computing document similarity rather than substring rolling hash—the time requirement for Cosine Similarity is greater than that of Levenshtein Distance. This is due to the fact that Levenshtein Distance compares two relatively short strings, whereas Cosine Similarity involves comparing two high-dimensional vectors representing entire documents.

B. Discussion

In Cosine Similarity algorithm, the text is represented as a vector, and the difference between words is represented as angles. This approach only considers the frequency of same words, without taking into account the order of these words. Therefore, the similarity score is only influenced by the frequency of shared words, while the word arrangement does not affect the similarity score at all.

In the Rabin-Karp algorithm, this approach involves comparing patterns (substrings). The pattern will be formed based on the value of ***k***-gram, which is influenced by the word order. Therefore, the similarity score is influenced by ***k***-gram and word arrangement. The smaller the value of ***k***-gram, the more patterns will be formed, potentially resulting in higher similarity.

This study adds to the previous research [11] that besides sentence arrangement, the similarity score in the Rabin-Karp algorithm is also influenced by the word arrangement. Responding to the study [12], it is observed that the same sentence with different word order can yield different similarity scores. Thus, if the order is altered during comparison, the similarity score may also change.

In the Levenshtein Distance algorithm, the distance calculation is performed by comparing characters. However, in this study, it has been successfully implemented by computing the distance based on word comparisons. This research addresses the issues in the previous study [26] since the Levenshtein Distance algorithm can perform well on long documents with unstructured sentence positions. This approach compares words at the same position, so even if two sentences have all the same words, the resulting similarity score will be 0% if all their positions are different. Therefore, the similarity score is highly influenced by word order. The more same words are found in the same order, the higher the resulting similarity score will be.

In each algorithm, the execution time is influenced by the number of steps each algorithm has to perform. The more steps that need to be performed, the longer the execution time required. In the Rabin-Karp algorithm, the execution time is also influenced by the value of k-gram, the larger the value of ***k***-gram, the more steps need to be performed.

In the Levenshtein Distance algorithm, the distance calculation method used [25], which directly calculates the word differences using Equation (4), can improve the execution time compared to creating a matrix. This

also makes the Levenshtein Distance algorithm faster than Cosine Similarity.

The test results indicate that the Rabin-Karp algorithm with 2-grams produces the highest similarity score with an average of 83.76%. However, based on the research conducted by [11], [13] on testing the Rabin-Karp algorithm with ***k***-grams from 1 to 10, it was found that ***k***-gram 3 achieved the highest accuracy because it produced very close similarity scores on each test data compared to other ***k***-grams. Hence, 4-gram is more suitable for plagiarism detection. Therefore, in this study, to compare the performance of the Rabin-Karp algorithm with other algorithms, the data used will be the test results with 4-gram. Therefore, the Cosine Similarity algorithm is superior in detecting similarity with an average similarity score of 48.80%, followed by the Rabin-Karp Algorithm with 4-grams at an average of 47.13%, and the Levenshtein Distance algorithm with an average of 20.61%. In terms of execution time, the Levenshtein Distance algorithm is superior as it shows the fastest performance with an average execution time of 0.22 seconds, followed by the Cosine Similarity algorithm with an average of 0.45 seconds, and the Rabin-Karp algorithm with 4-grams with an average of 39.15 seconds.

The same test sentence can yield the highest similarity score with different sentences in the comparison document, depending on the algorithm used in the testing. Thus, out of the 40 test documents, the Cosine Similarity algorithm produces the highest average similarity score in 26 documents, while the Rabin-Karp algorithm with 4-grams yields the highest average similarity score in 14 documents. Conversely, the Levenshtein Distance algorithm consistently produces lower average similarity scores compared to the other algorithms in all test documents.

Based on the performance of each algorithm, the selection of an algorithm for detecting plagiarism in documents can be based on the specific needs and objectives of the application. The Cosine Similarity algorithm is superior if the main priority is the similarity level of documents, without considering the word arrangement, and with relatively fast execution time. The Levenshtein Distance algorithm may be a more suitable choice if the application requires faster execution time while considering the word order. Meanwhile, the Rabin-Karp algorithm with 4-grams could be a better option for applications that emphasize high similarity levels while considering the word arrangement, despite longer execution time considerations.

IV. CONCLUSIONS

Based on the experimental results conducted, it was obtained that the best performance for detecting plagiarism was Cosine Similarity, followed by Rabin-Karp 4-gram and Levenshtein Distance. The plagiarism detection performance of Cosine Similarity, Rabin-Karp 4-gram, and Levenshtein Distance are 48.80%, 47.13%, and 20.61%, respectively. In contrast, the best time requirements are Levenshtein Distance, Cosine Similarity, and Rabin-Karp 4-gram, with time requirements of 0.22s, 0.45s, and 39.15s, respectively.

Plagiarism detection performance can be affected by the length of the words in the document, the arrangement of the words, and the number of the same words tested. Cosine Similarity in calculating the similarity between documents uses a vector space, while Rabin-Karp uses a substring rolling hash influenced by *k*-grams. Levenshtein compares two strings with a relatively short length while Cosine Similarity compares two vectors with a length that represents the entire document.

The selection of an algorithm for detecting plagiarism in documents should be guided by the specific requirements and objectives of the application. The Cosine Similarity algorithm is preferable when the primary focus is on the overall similarity between documents, regardless of word order, and when a relatively fast execution time is desired. In contrast, the Levenshtein Distance algorithm may be more appropriate for applications that prioritize both execution speed and consideration of word order. Meanwhile, the Rabin-Karp algorithm using may be the optimal choice for applications that emphasize achieving high similarity detection with attention to word order, even though it generally requires longer processing time.

REFERENCES

[1] Chowdhury, H.A., & Bhattacharyya, D. K. (2018). Plagiarism: Taxonomy, Tools and Detection Techniques. arXiv:1801.06323, Information Retrieval, doi: 10.48550/arXiv.1801.06323

[2] Velmurugan, V.S. (2024). Types and Definitions of Plagiarism: An Overview. Global Research Journal of Social Sciences and Management, vol. 2 (1), ISSN: 2583-858X, doi: 10.55306/GRJSSM.2024.2102.

[3] Bhavana, M., Rao, K. S., Koduru, G. K., Vatsal, K. V. K., Parvathi, S., & Sravani, M. (2024). Plagiarism Detection and Similarity Checking Program using Machine Learning and String Matching Algorithm. 2024 9th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2024, pp. 2032-2036, doi: 10.1109/ICCES63552.2024.10859471

[4] Perkins, M., Gezgin, U.B. & Roe, J. (2020). Reducing plagiarism through academic misconduct education. Int J Educ Integr, vol. 16 (3), doi: 10.1007/s40979-020-00052-8

[5] Republic of Indonesia, "Regulation of the Minister of Education, Culture, Research, and Technology of the Republic of Indonesia Number 39 of 2021 concerning Academic Integrity in Producing Scientific Works." Jakarta, 2021.

[6] Kumar, P., Gupta, M. K., Rao, C. R. S., Bhavsingh, M., & Srilakshmi, M. (2023). A Comparative Analysis of Collaborative Filtering Similarity Measurements for Recommendation Systems. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 184–192, doi: 10.17762/ijritcc.v11i3s.6180.

[7] Swathi, M., & Selvi, C. (2022). An Improved Similarity Measure Based on Collaborative Filtering for Sparsity Problem in Recommender Systems. In: Bhateja, V., Khin Wee, L., Lin, J.CW., Satapathy, S.C., Rajesh, T.M. (eds) Data Engineering and Intelligent Computing. Lecture Notes in Networks and Systems, vol 446. Springer, Singapore, doi: 10.1007/978-981-19-1559-8_5.

[8] Wahyuningsih, T., Henderi, H., & Winarno, W. (2021). Text Mining an Automatic Short Answer Grading (ASAG), Comparison of Three Methods of Cosine Similarity, Jaccard Similarity and Dice's Coefficient. Journal of Applied Data Sciences, 2(2), doi: 10.47738/jads.v2i2.31.

[9] Alobed, M., Altrad, A. M. M., & Bakar, Z. B. A. (2021). A Comparative Analysis of Euclidean, Jaccard and Cosine Similarity Measure and Arabic Wordnet for Automated Arabic Essay Scoring," 2021 Fifth International Conference on Information Retrieval and Knowledge Management (CAMP), Kuala Lumpur, Malaysia, 2021, pp. 70-74, doi: 10.1109/CAMP51653.2021.9498119.

[10] Setu, D.M., Islam, T., Erfan, M., Dey, S.K., Asif, M.R.A., & Samsuddoha, M. (2025). A comprehensive strategy for identifying plagiarism in academic submissions. J. Umm Al-Qura Univ. Eng. Archit. 16, 310–325, doi: 10.1007/s43995-025-00108-1.

[11] Hartanto, A. D., Syaputra, A., & Pristyanto, Y. (2019). Best Parameter Selection Of Rabin-Karp Algorithm In Detecting Document Similarity. 2019 International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2019, pp. 457-461, doi: 10.1109/ICOIACT46704.2019.8938458.

[12] Astuti, Y., & Wulandari, I., . (2022). An arrangement of the number of K-grams in the performance of Rabin Karp algorithm in text adjustment. Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), p-ISSN: 2502-4752, vol. 26 (3), pp. 1388-1394, doi: 10.11591/ijeecs.v26.i3.pp1388-1394.

[13] Hidayat, W., Utami, E., & Sunyoto, A. (2022). Selection of the Best K-Gram Value on Modified Rabin Karp Algorithm. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), vol. 16 (1), ISSN: 2460-7258, pp. 11-22, doi: 10.22146/ijccs.63686.

[14] Al-Hagree, S., & Al-Gaphari, G. (2022). Arabic Sentiment Analysis on Mobile Applications Using Levenshtein Distance Algorithm and Naive Bayes. 2022 2nd International Conference on Emerging Smart Technologies and Applications (eSmarTA), Ibb, Yemen, 2022, pp. 1-6, doi: 10.1109/eSmarTA56775.2022.9935492.

[15] Nandurkar, D. A., Ujjainkar, P., Miglani, B., & Kanojiya, A. (2023). Plagiarism Checker & Link Advisor using concepts of Levenshtein Distance Algorithm with Google Query Search - An Approach. 2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC), Jeddah, Saudi Arabia, 2023, pp. 1-6, doi: 10.1109/ICAISC56366.2023.10085404.

[16] Alvi, F., Stevenson, M., & Clough, P. (2017). Plagiarism Detection in Texts Obfuscated with Homoglyphs. ECIR 2017: Advances in Information Retrieval. 39th European Conference on Information Retrieval, 08-13 Apr 2017, Aberdeen, Scotland. Lecture Notes in Computer Science. Springer, Cham , pp. 669-675. ISBN 978-3-319-56608-5.

[17] Süzen, N., Gorban, A.N., Levesley, J., & Mirkes, E.M. (2020). Automatic short answer grading and feedback using text mining methods. Procedia Computer Science, vol 169, 2020, pp. 726-743,ISSN 1877-0509, doi: 10.1016/j.procs.2020.02.171

[18] Amalia, E.L., Lestari, V. A., Wijayaningrum, V.N., & Ridla, A. A. (2023). Automatic essay assessment in e-learning using winnowing algorithm. Indonesian Journal of Electrical Engineering and Computer Science, vol. 29 (1), pp. 572-582, ISSN: 2502-4752, doi: 10.11591/ijeecs.v29.i1.pp572-582.

[19] Alfat, L., Faisal, F. M., Negara, K. P. S., Munggaran, M. R., & Ihsan. (2023). Implementing Knuth-Morris-Pratt Algorithm in Detecting The Plagiarism of Document. 2023 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), Semarang, Indonesia, 2023, pp. 54-58, doi: 10.1109/ICITACEE58587.2023.10276453.

[20] Nalawati, R.E., & Yuntari, A.D. (2021). Ratcliff/Obershelp Algorithm as An Automatic Assessment on E-Learning. 2021 4th International Conference of Computer and Informatics Engineering (IC2IE), Depok, Indonesia, 2021, pp. 244-248, doi: 10.1109/IC2IE53219.2021.9649217.

[21] Barut, Z. & Altuntas, V. (2023). Applied Comparison of String Matching Algorithms. Gaziosmanpasa Journal of Scientific Research, vol. 12 (1), ISSN: 2146-8168.

[22] Madhan, N., Dheva Rajan, S., Jain, M. (2025). Directing Natural Language Processing Text Similarity Challenges in Social Media with AI Techniques. In: Goar, V., Kuri, M., Kumar, R., Senjyu, T. (eds) Advances in Information Communication Technology and Computing. AICTC 2024. Lecture Notes in Networks and Systems, vol 1075. Springer, Singapore, doi: 10.1007/978-981-97-6106-7_26.

[23] Leonardo, B. & Hansun, S. (2017). Text Documents Plagiarism Detection using Rabin-Karp and Jaro-Winkler Distance Algorithms. Indonesian Journal of Electrical Engineering and Computer Science, vol. 5 (2), pp. 462- 471, doi: 10.11591/ijeecs.v5.i2.pp462-471

[24] Khan, Z.A. & Pateriya, R.K. (2012). Multiple Pattern String Matching Methodologies: A Comparative Analysis. International Journal of Scientific and Research Publications, vol. 2 (7), ISSN 2250-3153.

[25] Coates, P. & Breitinger, F. (2023). Identifying document similarity using a fast estimation of the Levenshtein Distance based on compression and signatures. arXiv: Information Retrieval, doi: 10.48550/arXiv.2307.11496.

[26] Rao, P.J., Rao, K.N., Gokuruboyina, S., & Neeraja, K.N. (2024). An Efficient Methodology for Identifying the Similarity Between Languages with Levenshtein Distance. In: Kumar, A., Mozar, S. (eds) Proceedings of the 6th International Conference on Communications and Cyber Physical Engineering. ICCCE 2024. Lecture Notes in Electrical Engineering, vol. 1096. Springer, Singapore, doi: 10.1007/978-981-99-7137-4_15.

[27] Kalbaliyev, E., Rustamov, S. (2021). Text Similarity Detection Using Machine Learning Algorithms with Character-Based Similarity Measures. In: Biele, C., Kacprzyk, J., Owsiński, J.W., Romanowski, A., Sikorski, M. (eds) Digital Interaction and Machine Intelligence. MIDI 2020. Advances in Intelligent Systems and Computing, vol 1376. Springer, Cham, doi: 10.1007/978-3-030-74728-2_2.

[28] Finansyah, A. Y. W., Afiahayati, & Sutanto, V.M. (2022). Performance Comparison of Similarity Measure Algorithm as Data Preprocessing Stage: Text Normalization in Bahasa Indonesia. Scientific Journal of Informatics, vol. 9 (1), ISSN 2407-7658, pp. 1-7, doi: 10.15294/sji.v9i1.30052.

[29] Usino, W., Prabuwono, A.S., Allehaibi, K.H.S., Bramantoro, A., Hasniaty A, & Amaldi, W. (2019). Document Similarity Detection using K-Means and Cosine Distance. (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 10 (2), pp. 165-170, doi: 10.14569/IJACSA.2019.0100222

ISSN 2085-4552

A Comparative Study : Predicting Customer Churn in Banking Using Logistic Regression & Random Forest

Mhd Basri[1], Muhammad Iqbal Pradipta[2], Krisna Aditya[3]
[1,2,3] Faculty of Computer Science and Information Technology,
Muhammadiyah University of North Sumatra, Medan, Indonesia
[1]mhd.basri@umsu.ac.id , [2]muhammad.iqbalpradipta@umsu.ac.id, [3]ozwellespencer49@gmail.com

Accepted 20 May 2025
Approved 13 June 2025

***Abstract*— Customer churn prediction is crucial in the banking industry, where retaining existing customers is often more cost-effective than acquiring new ones. Understanding the factors that lead customers to leave can help banks implement proactive retention strategies, improve customer satisfaction, and maintain long-term profitability. This research explores the prediction of bank customer churn using machine learning techniques. The dataset used includes various customer features such as demographics, transaction history, and interactions with the bank. After performing exploratory data analysis (EDA) and pre-processing, two machine learning models were applied: Logistic Regression and Random Forest. Logistic Regression was chosen for its simplicity and interpretability, while Random Forest was selected for its ability to handle complex, non-linear relationships in the data. The EDA results showed that factors such as number of transactions, total transaction value, and credit utilization rate were correlated with the likelihood of churn. Pre-processing included handling categorical data, removing irrelevant features, and dividing the data into training and testing sets. The Logistic Regression model achieved 84% accuracy on training data and 83.9% on testing data, but showed poor performance in terms of recall and F1-score for the "Attracted Customer" class. In contrast, the Random Forest model showed excellent performance with 100% accuracy on both datasets, as well as perfect precision, recall, and F1-score values for both classes. In conclusion, the Random Forest model was selected as the best model to predict bank customer churn. These findings can help banks identify customers at risk of churn and develop effective retention strategies.**

Index Terms*— *Customer churn; Attrition; Banking Industry; Machine learning; Logistic Regression; Random Forest.

I. INTRODUCTION

Customer attrition, also known as customer churn, is the phenomenon where customers terminate their relationship with a business or organization [1]. Customer churn has emerged as a pressing challenge in sectors such as banking, telecommunications, music media, and insurance [2]. Customer attrition in the banking industry occurs when consumers stop using the products and services offered by the bank for some time, and then end their relationship with the bank. Therefore, customer retention has become very important in today's highly competitive banking market [1]. The rate at which a company loses clients within a certain period of time is called the churn rate, also known as the customer attrition rate [3]. Although many banks strive to attract and retain customers, customers are often dissatisfied and choose to switch. Factors that can lead to customer loss can include poor service quality, fees that are perceived to be too high, or a lack of innovation in products and services. In addition, having a strong customer base helps attract new consumers by building trust and gaining referrals from existing customers. These factors make reducing customer attrition an important step for banks to take.

The impact of customer churn can be very significant for banks Losing a customer means losing potentially valuable revenue. In addition, the cost of attracting new customers is often higher than retaining existing ones. To find customers who may be lost, customer churn prediction techniques can be used. Then, the results can change the marketing strategy [4]. If churn is not managed well, it can also damage the bank's reputation, making potential customers hesitant to join. Therefore, companies must accurately predict customer churn and take appropriate actions. A high churn rate can create a negative perception in the community. When many customers leave a bank, it can affect the trust of other customers and make them consider switching banks. This domino effect can be very detrimental to the bank in the long run. With the advancement of technology, researchers are starting to look for solutions to this problem.

Customer churn prediction has become an important research focus in various industries, especially in banking, where customer retention directly affects revenue streams and satisfaction levels.

For example, Lalwani et al. [5] conducted churn prediction research in the telecommunications industry using a dataset consisting of approximately 7,000 customer records with 21 features, including numerical and categorical attributes such as tenure, monthly fee, contract type, and internet service usage. The dataset was pre-processed through missing value handling, categorical coding, and balancing using resampling techniques, thus enabling the application of machine learning models for accurate churn classification. Traditional methods like Logistic Regression have shown promising results, achieving accuracy rates of around 80.45%

Verma et al. study conducted research on churn prediction for savings bank customers using transaction and demographic data obtained from a national bank data warehouse in India [6]. The dataset included customers aged 21-50 years featuring 47 variables selected from 66 available variables, such as age, income, occupation, and transaction behavior. The Random Forest algorithm outperformed other models including GLM, ANN, Decision Tree, and XG-Boost, achieving 78% accuracy on test data and 79% on validation data, with an AUC of 0.844. This study highlights Random Forest as a reliable model for practical application in customer retention strategies.

The study presented by Wagh in predicting customer churn used the Telco Customer Churn dataset obtained from Kaggle, which included 7,043 customer records with 21 attributes (16 categorical and 5 numerical). The dataset labeled 26.53% of customers as churners. The Random Forest algorithm was used for classification, and outperformed the other models tested. Before up-sampling, the model achieved 98.91% accuracy, with 99% precision and recall. After applying SMOTE and ENN techniques to address class imbalance, the model performance further improved to reach 99.09% accuracy [7]. The results of this study conclude that Random Forest offers a robust solution for churn prediction in the telecommunications sector and can be effectively integrated with retention strategies such as survival analysis and Cox proportional hazard modeling.

Another example, research conducted by Prabadevi by testing several machine learning algorithms such as K-Nearest Neighbor (KNN), Logistic Regression (LR), Random Forest (RF) and also Stochastic Gradient Booster (SGB) shows that the Linear Regression algorithm is still superior to the Random Forest algorithm with Classification Score Linear Regression getting Training Score 0.797, Test Performance 0.782, and AUROC 0.826 [10]. 797, Test Performance of 0.782 and AUROC of 0.826. while the Random forest algorithm gets a Training Score of 0.803, Test Performance of 0.787 and AUROC of 0.829. meaning that the Random Forest model is claimed to be better than Linear Regression [8].

Similarly, Dhangar also tested with the Logistic Regression algorithm, GaussianNB, SVM and Random Forest. the results show that SVM and Random forest get the advantage to predict Customer Churn with the Area under the curve (AUC) score of SVM 0.921 and Random forest 0.945. on Logistic Regression Getting AUC of 0.908. shows that SVM and Random forest are better at predicting Customer churn [9].

Advanced techniques, including hybrid approaches such as BiLSTM-CNN models, have emerged as powerful tools for capturing sequential patterns in churn prediction [10]. The banking sector has particularly benefited from these developments, with studies showing that proper handling of class imbalance through under-sampling techniques can significantly improve model performance [6]. Recent research has also explored the effectiveness of ensemble methods and hybrid algorithms, with some studies indicating that traditional approaches like Logistic Regression can outperform more complex models in certain scenarios [5], [9].

In an effort to address customer churn in the banking sector, deep learning approaches offer a promising solution with its ability to comprehensively analyze complex patterns from customer data. Bank churn prediction is made to find out how likely clients are to switch from one bank to another [11]. In this research, the aim is to analyze the bank's data and predict which users are likely to stop using the bank's services and turn into paying customers. Analyzing this data will help banks identify trends and try to retain customers who are on the verge of attrition.

II. METHODOLOGY

The machine learning project used data-driven methodologies in an organized manner to tackle the specified challenge. The first phase was gathering and prepping the data, which involved cleaning it after obtaining it from a reliable source to guarantee its quality and dependability. This procedure involved resolving any discrepancies, standardizing the data, and managing missing numbers. After preprocessing, the dataset's distribution, correlations, and possible patterns were examined using exploratory data analysis (EDA), which influenced the modeling approaches selected.

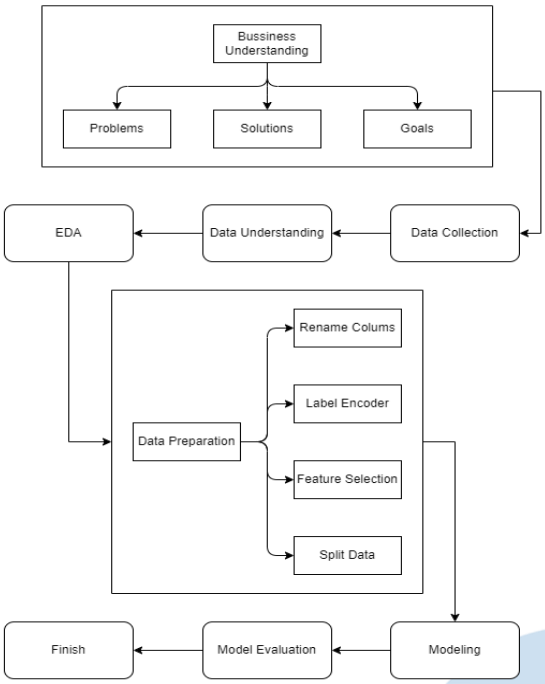


Fig. 1 Research Methods

A. *Bussiness Understanding*

This research focuses on the problem of customer attrition (churn) in the banking industry. Attrition occurs when customers stop using banking products and services, which negatively impacts the bank's revenue. In a highly competitive industry, customer retention is important to maintain business stability and reduce the cost of acquiring new customers. By having historical data from customers, we can use machine learning to predict customers who are at risk of quitting.

The data pre-processing steps that need to be performed to make the data ready to be used for training the model include various steps, such as data cleaning, handling missing values, normalization, and dividing the data into subsets suitable for training and testing.

The solution that can be done is to create a deep learning model and then evaluate the model using the Classification Report by looking at important metrics such as precision, recall, F1-score, and support, which are very helpful in assessing the performance of the model in classifying data and also creating heatmap graphs to analyze the correlation between existing features. Pre-processing the data properly so that it can be used, creating deep learning models to predict or identify customers who are at risk of churn and also knowing what features are most influential in customers who churn or not.

B. *Data Collection*

In the data collection stage, a dataset titled “Bank Customer Churn Prediction” sourced from the Hugging Face Datasets repository was used, and originally published through Analyttica's business analytics platform TreasureHunt. The platform is known for providing real datasets that are used to solve practical business problems. The dataset consists of 10,000 customer records from a private sector bank focused on credit card services. Each customer data includes 14 attributes, including: Customer ID, credit score, location, gender, age, tenure, account balance, number of products used, credit card ownership, active membership status, estimated income, and churn indicators. These attributes provide a comprehensive view of each customer's profile, enabling the construction of predictive models to identify customers at risk of churning. One important characteristic of this dataset is the presence of class imbalance, where only about 16.07% of customers are recorded to have churned. This poses a challenge in the model training process as it requires an appropriate strategy to accurately predict the minority class.

C. *Data Understanding*

The dataset that has been downloaded through Hugging face, then the dataset will be saved to google drive. In statistical analysis, this is very important as it enables proper handling of data, helps in model selection, and improves model performance and accuracy [12]. The dataset is accessed by mounting by specifying the file path to be accessed using Google Colab, then reading the file and adjusting it to pandas Dataframe and finally displaying information from the dataset which consists of 10,127 rows and 23 columns. In the context of customer churn prediction analysis in the banking industry, this stage is very important to identify patterns, trends, and relationships between variables that can affect churn.

D. *Exploratory Data Analysis*

Exploratory data analysis (EDA) process shows that the majority of customers are in the age range of 40-52 years, with some outliers whose age is above 70 years old as shown in the following figure:

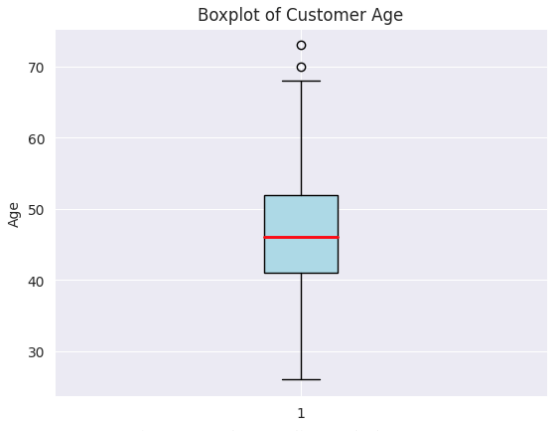


Fig. 2 Boxplot Credit_card_dataset

The figure above shows a boxplot of the age distribution of customers in the dataset used. The red line inside this visualization box shows that the median age of customers is about 46 years old. Most customers are between 35 and 55 years old, or the first and third quartiles. Of the few points above the age of 70, values outside this range are considered outliers. An initial overview of the age distribution of customers is given by this boxplot. This is important to analyze further as age is one of the factors that influence the likelihood of customer churn. In addition, the relatively symmetrical distribution indicates that the age variable does not suffer from significant skewness.

E. Data Preparation

In this section, we will explain in detail about the data cleaning and preprocessing processes carried out to prepare the dataset for analysis and modeling. This process is very important because good data quality will directly contribute to the accuracy and effectiveness of the machine learning models built.

To make the dataset easier to understand and use, column names that were too long were changed to shorter ones. To rename a column, a function with the parameter inplace=True is used, which indicates that the change is applied directly to the DataFrame. To ensure that the change is successful, the renamed column name is displayed.

The use of Label Encoder aims to convert categorical data into numerical format. This function allows categorical fields in a dataset to be processed by machine learning models by converting the category values into numbers, so that the model can learn patterns in the data. Label coding improves regression accuracy by enabling binary classification algorithms, improving error correction, and supporting end-to-end training across multiple regression tasks [13] . Feature importance analysis is essential to understand which input variables have the most significant impact on model predictions [14]. Removing columns that are irrelevant or unrelated to the model makes it easier for the model to process data so that it can focus on relevant features and improve its performance. By removing unrelated features, the model becomes simpler and more efficient.

The purpose of Split data is to divide the dataset into two parts training set and testing set, where there are independent features or X to predict the target and Y as the dependent variable or target label. in this case 20% of the total dataset will be used as testing data while 80% will be used as training data. Details of the data pre-processing used in this study can be seen in Table 1.

Table 1 Data Preprocessing Steps

Step	Description	Explanation
Column Renaming	Rename columns that are too long and change them to shorter versions.	use the rename() function with inplace = True to apply the changes directly to the DataFrame.
Label Encoding	converts categorical variables into a numerical format so that machine learning algorithms can process them effectively.	Each categorical feature is transformed using LabelEncoder().
Feature Removal	Irrelevant and insignificant columns were removed to improve model performance and reduce noise in the data.	This deletion uses the credit.card.drop() function
Data Splitting	The dataset was split into a training set and a testing set using an 80:20 ratio.	X represents the independent features, and y represents the dependent target variable Attrition_Flag. A fixed random_state ensures reproducibility.

F. Model Selection

In the Modeling stage, it discusses the use of machine learning algorithms to solve the churn prediction problem faced in the project. In this case, two models are used, namely Logistic Regression and Random Forest, to classify whether a customer will churn or not. Logistic Regression is a model of the Generalized Linear Model (GLM) where the response variable is a binary number (0 or 1) and follows a binomial distribution [15]. This model works by using a linear combination of input features to predict the probability that the data belongs to a particular category. It is easy to understand because it has a linear coefficient for each feature, which indicates how strongly each feature affects the prediction. This model is also fast and requires less computational resources. If there is a linear relationship between the features and the target variable, logistic regression works well.

Random forest is a machine learning algorithm that uses an ensemble of decision trees to make predictions [16]. Random Forest helps reduce overfitting and often produces more accurate predictions than single tree models. When compared to a single decision tree, this model is more stable against overfitting. Random Forest has complex interactions between features and can handle non-linear data. It can provide metrics to show which aspects affect the prediction the most.

G. Model Evaluation

This research uses a classification-type deep learning model which means that if it is close to 100% accuracy, the performance is good, while if it is below 75%, the performance is poor. The number of metrics that indicate different aspects of classification model performance is very important to consider when developing and evaluating models [13]. The following evaluation metrics and confusion matrix are results obtained from the best-performing model, which in this case is Random Forest :

- **Accuracy**

Accuracy is a measure of how many predictions are correct compared to all predictions made by the model.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Based on the confusion matrix results, the model produces True Positive (TP) = 1699, True Negative (TN) = 327, and there are no False Positive (FP) or False Negative (FN). Then, the accuracy of the model is :

$$\frac{1669+327}{1699+327+0+0} = \frac{2026}{2026} = 100\%$$

- **Confusion Matrix**

Confusion matrix describes the results of model prediction against test data with the following details:

Table. 2 Confusion Matrix

	Negative Prediction (0)	Positive Prediction (1)
Actual Negative (0)	327	0
Actual Positive (1)	0	1699

- **Precision**

Precision is a measure of how many true positive predictions (True Positives, TP) are correct compared to all positive predictions made by the model (including False Positives, FP).

$$Precision = \frac{TP}{TP+FN} \quad (2)$$

With TP = 1699 and FP = 0, the precision is:

$$\frac{1699}{1699+0} = 1.0$$

- **Recall**

One of the evaluation metrics in classification is recall, which measures how many positive examples the model managed to find out of all the truly positive data.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Since FN = 0, recall:

$$\frac{1699}{1699+0} = 1.0$$

- **F1 – Score**

F1-Score, a metric that combines Precision and Recall in a single value, provides a more complete picture of model performance, especially in cases where Precision and Recall are not balanced.

$$F1-Score = \frac{2 \times Precision \times Recall}{Precision \times Recall} \quad (4)$$

Since Precision and Recall are both 1.0, then:

$$F1 = 2 \times \frac{1.0 \times 1.0}{1.0 \times 1.0} = 1.0$$

III. RESULT

In the methods section of this study after collecting and preparing the data, which involves cleaning the data after obtaining it from reliable sources to ensure its quality and reliability. This procedure involves resolving any discrepancies, standardizing data, and managing missing numbers. After preprocessing, the distribution, correlation, and possible patterns of the data set are examined using exploratory data analysis (EDA), which influences the modeling approach chosen. After the modeling stage, the algorithms used are evaluated and compared using performance metrics such as accuracy, precision, recall, and F1-score, to ensure the selection of the most effective model. Hyperparameter tuning is performed to further optimize model performance, using techniques such as grid search or random search.

The methodology also emphasizes model validation through techniques such as cross-validation, which helps assess the model's resistance to overfitting. Finally, the best-performing model is deployed and tested with unseen data to ensure its generalizability. Throughout the project, tools such as Python, Scikit-learn, and Matplotlib were used for implementation and visualization purposes.

Figure 3 below shows the distribution of the length of time a customer (in months) has been registered with a bank. The histogram shows the distribution of the original data, while the Kernel Density Estimation (KDE) adds a smooth curve that estimates the probability distribution of the data. This combination of histogram and KDE is useful for understanding how the data is distributed in more detail.

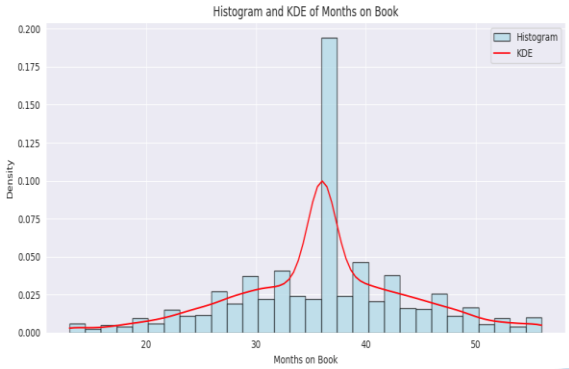


Fig. 3 Months_of_Book Relations

A. *Data Distribution*

The graph shows a distribution that is slightly skewed to the right. This means that there are fewer customers who have been customers for longer, and most customers have a subscription time of less than 50 months. There are some customers who have longer subscription times (more than 50 months), but the number is very small compared to the group around 36 months.

B. *Business Interpretation*

Banks may want to focus on customers with subscription times around 36 months as this is the largest group. This could be a potential target for increasing engagement or offering new products. Customers who have been subscribed for longer (more than 50 months) are a smaller group that may require special attention, especially if there are concerns about churn or customer loyalty.

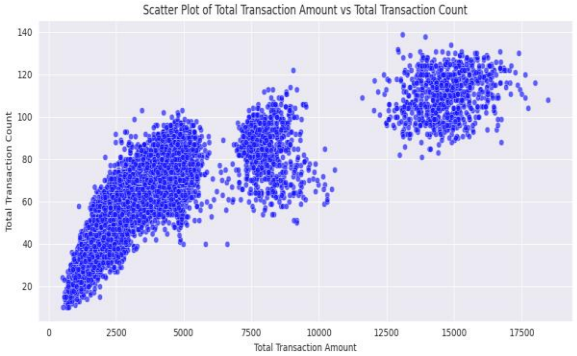


Fig. 4 Total Amount and Count Relationship in Segmentation

Customer Behavior Segmentation can be divided into two, namely High Frequency, Low Value Customers The group on the right shows customers who are actively transacting but with a small transaction value. They may be users who make routine or daily purchases, such as daily shopping. Low Frequency, High Value Customers The group on the left shows customers who rarely transact but when they do, the transaction value is high. They may be users who purchase expensive items regularly as shown in figure 4.

Visualization of correlation between Multivariate features using heatmap it will remove some features that are not needed, shown in Figure 5:

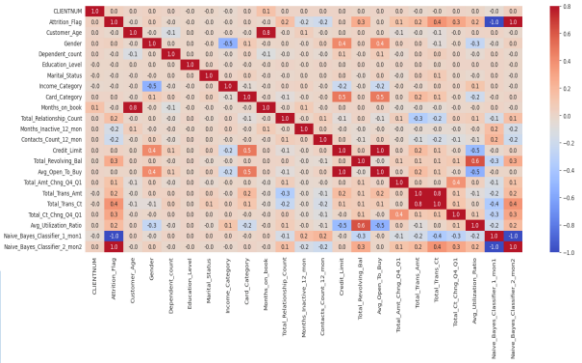


Fig. 5 Overall relationship between variables.

Total_Trans_Ct (Total Number of Transactions) There is a correlation of 0.4 between Attrition_Flag and Total_Trans_Ct, which shows a moderate positive relationship, indicating that the more transactions a customer makes, the less likely they are to churn and the more active customers tend to stay. Total_Trans_Amt (Total Amount of All Transactions) There is a correlation of 0.2 with Attrition_Flag, which indicates a positive relationship. Customers who spend more money through transactions are less likely to leave the service. Total_Relationship_Count has a correlation of 0.2 with Attrition_Flag, which indicates that the more interaction or relationship a customer has with the company (e.g. through products, services, or other contacts), the less likely they will churn.

The data analysis conducted on existing and attracted customers based on gender and marital status provides important insights into customer behavior and preferences. The following is the business interpretation generated from the following figure 6:

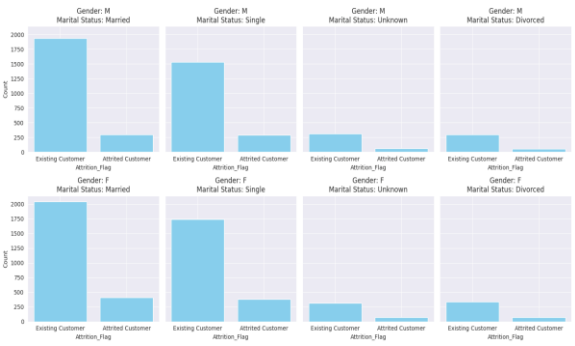


Fig. 6 Relationship between Attrition_Flag and marital status

Churn rate based on Marriage status

It can be seen that both married men and women have lower churn (attrited) rates than customers who are single or whose marital status is unknown. This suggests that married customers tend to be more stable in their relationship with the company, which can be interpreted as higher loyalty. Specifically in the "Married" category, there is a very high proportion of existing customers compared to attracted customers. This indicates that marital status plays a positive role in customer retention. While "Single" customers also show a similar pattern, their loyalty level is slightly lower.

The effect of marital status on churn

Interestingly, customers with "Unknown" and "Divorced" marital statuses tend to have higher churn rates, almost equal to the originality rate of existing customers in a given category. The results suggest that marital status uncertainty can be an important factor in customer retention. When customers' marital status changes, such as divorce, they may experience emotional and financial instability, which may affect their decision to stay in touch with the company.

Opportunities for customer segmentation

From the analysis conducted, it can be concluded that married customers show higher loyalty compared to other customer categories. This creates an opportunity for companies to better segment based on marital status. By understanding that married customers are more likely to remain loyal, companies can design more relevant products and services, specifically aimed at this segment.

The analysis of the distribution of existing and attrited (churned) customers by revenue category in Figure 7, provides valuable insights into the customer profile and strategies that the company can adopt. The following discussion outlines the key findings as well as the opportunities and challenges that arise from the data.

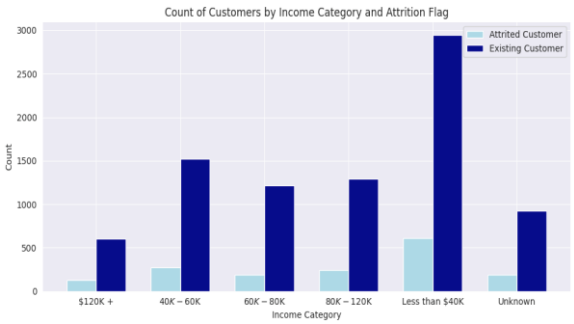


Fig. 7 income_category_with_attrition

Existing Customer Distribution

The data shows that customers with an income of less than 40K have the highest proportion in the existing customer category, much more than the other income categories. This indicates that the company has a strong customer base in the low-income segment, which may be a result of offering affordable products and services. The income category between 40K to 60K also shows a significant number of existing customers, although it is still below the low-income segment.

On the other, customers in the 120K+ and 80K to 120K income categories appear to be quite few. This indicates that the company has not been successful in attracting customers in the high-income segment, which could be an indication of an opportunity to improve marketing or products geared towards this segment.

Distribution of Attrited Customers (Churn)

Although the low-income segment has a high number of customers, the churn rate remains a concern, and companies need to develop strategies to retain customers in this category. Churned customers also tend to come from the "Unknown" category and the less than 40k income category. In contrast, higher income classes, such as 120k+ and 80k to 120k, show very low churn rates. This may indicate higher stability among high-income customers, or perhaps the company has not actively explored this market.

Opportunities in the Low-income Segment

The company can make cheap products and implement a strong loyalty program since the majority of its customers are from the low-income segment. This method has the potential to increase retention of existing customers and simultaneously attract new customers from comparable industries. Offering attractive discounts, referral programs, and manufacturing products that meet consumer needs and preferences are some of the strategies that can be implemented.

Challenges in the High-income Segment

The company struggles to attract customers from the high-income segment which has a very small existing and interested customer base. Developing premium products and services tailored to the needs of this segment can help expand penetration in this market. To discover what can attract high-income customers, more research needs to be done on their behavior and preferences.

Category "Unknown"

The presence of a customer in the "Unknown" category indicates that there is no information available about the customer's revenue. Identifying and collecting information about the revenue of these customers can help build better retention and acquisition strategies. Companies can improve the accuracy of their marketing and make offers that better suit the segment in question by collecting more complete information about their customers.

To make the analysis clearer and more concise, the following table summarizes the correlation between customer attributes and churn rates, along with corresponding insights and business implications.

Table. 3 Marital Status vs Churn

Marital Status	Churn Rate	Bussiness Insight
Married	Low	High loyalty, good retention opportunity
Single	Moderate	Slightly less loyal than married
Divorced / Unknown	High Churn	Potential emotional/financial instability

Table. 4 Income Category vs Churn

Income Category	Churn Rate	Bussiness Insight
< 40K	High existing, high churn	Price-sensitive, needs loyalty programs.
40K – 60K	Moderate existing, moderate churn	Growing segment, needs targeted offers.
80K – 120K / 120K+	Low churn, Low existing	Untapped premium market, opportunity to expand.
Unknown	Moderate Churn	Lack of data, needs better profiling.

Then we can also see in the card category where users are mostly in the Blue card type both in terms of churn and non churn which is found in Figure 8 :

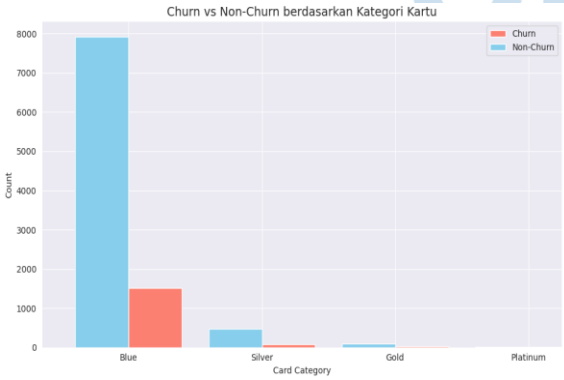


Fig. 8 Card Category

Comparison between churn and non-churn customers for various card categories. From the data shown, the Blue card has the highest number of customers with around 7,800 non-churn customers and around 1,500 churn customers. The Silver category comes in second with a much smaller number of about 500 non-churn customers and very few churn customers. Meanwhile, the Gold category has the least number of customers with only a few hundred non-churn customers and almost no churn. For the Platinum category, there is no significant customer data visible in the graph.

Interestingly, the data shows that the higher the card category, the lower the churn rate, which may indicate that higher category cardholders tend to be more loyal to the service. The Blue card, despite having the largest customer base, also shows a relatively higher churn rate compared to other categories.

IV. DISCUSSION

The performance analysis of the classification models used in this project is presented in this report. The three main metrics analyzed are Accuracy, Precision, and Recall. These metrics provide an in-depth look at the model's ability to correctly identify positive and negative classes.

A. Logistic Regression

In the precision section, it can be seen that Attirted Customer out of 327 data that is predicted by the prediction model, 100% is predicted by Attirted Customer. Existing Customer from 1699 data that is predicted by the prediction model, 84% is predicted by Existing Customer. In the Recall section, In the Recall section, Attirted Customer from 0% predicted Attirted Customer, 0% resulted correct. Existing Customer from 100% predicted Existing Customer, 100% resulted correct, but for Recall and F1 Score it is not good, so from the results it can be concluded that the model created is Not Good Fit. From both categories it can be seen that the resulting model performs poorly, especially in the Attirted Customer label Conclusion From the table results we can see that the model gets good performance with Accuracy on training data 0.8410072830514751 Accuracy Score 0.839091806515301 the full explanation is in table 5.

Table. 5 Classification Report Logistic Regression

	Precision	Recall	F1-Score	Support
Attrited Customer	1.00	0.00	0.01	327
Existing Customer	0.84	1.00	0.91	1699
Accuracy			0.84	2026
Macro avg	0.92	0.50	0.46	2026
Weighted Avg	0.86	0.84	0.77	2026

B. Random Forest

Based on the prediction results against the six labels listed in table 6, the model performed well with

all metrics exceeding 70% and testing accuracy of 90%, in accordance with the subsequent analysis, which concentrated on Precision, Recall, and F1 scores for the two main categories of customers: 'Interested Customers' and 'Existing Customers'. In the precision section of 327 data predicted as Attrited Customer, the model successfully predicted 100% as Attrited, this shows that there are no errors in determining attrited customers. Of the 1699 data predicted as Existing Customer, the model also successfully predicted one hundred percent with the same accuracy. This indicates that all customers identified as existing customers are existing customers.

In the Recall section the model managed to make 100% of all predictions made for the targeted customers, indicating that every targeted customer actually fell into the category, demonstrating the model's ability to find all positive cases. In addition, the model also successfully identified 100% of the predicted customers as existing customers, and all predictions in this category also proved to be correct, demonstrating the model's exceptional accuracy. F1-Score metric results show that the model performs well. This score, which combines Precision and Recall, gives an overall picture of the balance between these two metrics, and since all predictions are correct, this score will also show an excellent value.

Based on the analysis conducted, the model evaluation results are as follows Training Data Accuracy gets 1.0 (100%) while on Score Accuracy gets 1.0 (100%) From these results, it can be concluded that the tested model has excellent performance and high quality, and all metrics have perfect results. Therefore, the model created by the Random Forest algorithm is good fit.

Table. 6 Classification Report Random Forest

	Precision	Recall	F1-Score	Support
Attrited Customer	1.00	1.00	1.00	327
Existing Customer	1.00	1.00	1.00	1699
Accuracy			1.00	2026
Macro avg	1.00	1.00	1.00	2026
Weighted Avg	1.00	1.00	1.00	2026

However, this research has some limitations, including the datasets used may not be fully representative of real-world scenarios, as they are pre-processed and collected from specific sources. The memory and computation time requirements of the random forest model can be quite large especially if a large number of datasets are used to store the training data. Future research can overcome these limitations by studying larger datasets, improving model parameters, or integrating more machine learning algorithms for a more thorough analysis.

V. CONCLUSIONS

From Figure 5, it can be concluded that the influential features to create a model are the features Total_Trans_Ct (Total Number of Transactions):, Total_Trans_Amt (Total Amount of All Transactions):, Total_Relationship_Count (Total Number of Relationships with the Company):, Avg_Utilization_Ratio (Average Credit Usage):, Total_Revolving_Bal (Total Revolving Balance):, Months_Inactive_12_mon (Inactive Months in the Last 12 Months):, Contacts_Count_12_mon (Number of Contacts with the Company in the Last 12 Months):

From the Logistic Regression results we can see that the model performs well in train with 84% accuracy and 83% validation. From the test data, we can create a classification report in table 3. From table 3 we can see the score of the model against the 3 metrics generated by the classification report, judging from the results it can be concluded that the model created does not have a good value especially in the Recall and F1 Score sections.

On the other hand, from the Random Forest results, we can see that the model gets good performance in training with 100% accuracy and 100% validation, this good performance is also proven by evaluating the model using the prepared test data. From the test data, a classification report can be made in table 4. From table 4 we can see the score of the model against the 3 metrics generated by the classification report, judging from the results it can be concluded that the model made is Good Fit. So that the model used is the result of Random Forest.

REFERENCES

[1] Singh, P. P., Anik, F. I., Senapati, R., Sinha, A., Sakib, N., & Hossain, E. (2024). Investigating customer churn in banking: A machine learning approach and visualization app for data science and management. Data Science and Management, 7(1), 7–16. Https://doi.org/10.1016/j.dsm.2023.09.002

[2] Zhang, H., & Zhang, W. (2024). Application of GWO-attention-convlstm model in customer churn prediction and satisfaction analysis in customer relationship management. Heliyon, 10(17). Https://doi.org/10.1016/j.heliyon.2024.e37229

[3] Ahmad, A. K., Jafar, A., & Aljoumaa, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, *6*(1). Https://doi.org/10.1186/s40537-019-0191-6

[4] Xiahou, X., & Harada, Y. (2022). B2C E-Commerce Customer Churn Prediction Based on K-Means and SVM. Journal of Theoretical and Applied Electronic Commerce Research, 17(2), 458–475. Https://doi.org/10.3390/jtaer17020024

[5] Lalwani, P., Mishra, M. K., Chadha, J. S., & Sethi, P. (2022). Customer churn prediction system: a machine learning approach. Computing, 104(2), 271–294. Https://doi.org/10.1007/s00607-021-00908-y

[6] Verma, P. (2020). Churn prediction for savings bank customers: A machine learning approach. Journal of Statistics

Applications and Probability, 9(3), 535–547. Https://doi.org/10.18576/JSAP/090310

[7] Wagh, S. K., Andhale, A. A., Wagh, K. S., Pansare, J. R., Ambadekar, S. P., & Gawande, S. H. (2024). Customer churn prediction in telecom sector using machine learning techniques. Results in Control and Optimization, 14. Https://doi.org/10.1016/j.rico.2023.100342

[8] Prabadevi, B., Shalini, R., & Kavitha, B. R. (2023). Customer churning analysis using machine learning algorithms. International Journal of Intelligent Networks, 4, 145–154. https://doi.org/10.1016/j.ijin.2023.05.005

[9] Dhangar, K. (2021). A review on customer churn prediction using machine learning approach. In novateur publications international journal of innovations in engineering research and technology (Vol. 8, Issue 5).

[10] Khattak, A., Mehak, Z., Ahmad, H., Asghar, M. U., Asghar, M. Z., & Khan, A. (2023). Customer churn prediction using composite deep learning technique. Scientific Reports, 13(1). Https://doi.org/10.1038/s41598-023-44396-w

[11] AL-Najjar, D., Al-Rousan, N., & AL-Najjar, H. (2022). Machine Learning to Develop Credit Card Customer Churn Prediction. Journal of Theoretical and Applied Electronic Commerce Research, 17(4), 1529–1542. Https://doi.org/10.3390/jtaer17040077

[12] Afolabi, S., Ajadi, N., Jimoh, A., & Adenekan, I. (2025). Predicting diabetes using supervised machine learning algorithms on E-health records. Informatics and Health, 2(1), 9–16. Https://doi.org/10.1016/j.infoh.2024.12.002

[13] Shah, D., Xue, Z. Y., & Aamodt, T. M. (2022). Label Encoding for Regression Networks. Http://arxiv.org/abs/2212.01927

[14] Hamzat, A. K., Salman, U. T., Murad, M. S., Altay, O., Bahceci, E., Asmatulu, E., Bakir, M., & Asmatulu, R. (2025). Development of robust machine learning models for predicting flexural strengths of fiber-reinforced polymeric composites. Hybrid Advances, 8. Https://doi.org/10.1016/j.hybadv.2025.100385

[15] Setiawan, E., Azis Suprayogi, M., Kurnia, A., Setiawan, E., Suprayogi, M. A., & Kurnia, A. (2025). BAREKENG: Journal of Mathematics and Its Applications a comparison of logistic regression, mixed logistic regression, and geographically weighted logistic regression on public health development in java a comparison of logistic regression, mixed logistic regression and …. Issue 1 j. Math. & app, 19(1), 129–0140. Https://doi.org/10.30598/barekengvol19iss1pp0129-0140

[16] Chen, P., Xiong, H., Cao, J., Cui, M., Hou, J., & Guo, Z. (2025). Predicting postoperative adhesive small bowel obstruction in infants under 3 months with intestinal malrotation: a random forest approach. *Jornal de Pediatria*. Https://doi.org/10.1016/j.jped.2024.11.011

ISSN 2085-4552

Enhancing Intelligent Tutoring Systems through SVM-Based Academic Performance Classification and Rule-Based Question Recommendation

*[1]Fenina Adline Twince Tobing, [2]Toto Haryanto
[*1,2]School of Data Science, Mathematics and Informatics, IPB University,
Bogor, West Java, 16680, Indonesia
[1]Informatics Department, Faculty of Engineering and Informatics,
Universitas Multimedia Nusantara, Tangerang, Indonesia
email: *[1]feninatobing@apps.ipb.ac.id, [1]fenina.tobing@umn.ac.id, [2]totoharyanto@apps.ipb.ac.id

Accepted 04 June 2025
Approved 09 June 2025

***Abstract*— This research aims to automatically classify students' academic performance levels using Support Vector Machine (SVM) algorithm and automatically recommend questions based on classification results. Dataset consists of six assignment scores per student, averaging students into three performance levels: Beginner, Intermediate, and Advanced. Before training, the data undergoes preprocessing involving normalization with StandardScaler and splitting into training and testing sets. Model is trained using Radial Basis Function (RBF) kernel with hyperparameter tuning to optimize its performance. Evaluation results show that the model achieved an accuracy of 91.67%, with a precision of 93.06%, a recall of 91.67%, and an F1-score of 91.89%. The best performance was found in Intermediate class, the dominant category in dataset, while performance in Advanced category was relatively lower due to limited sample size. Following classification, a rule-based recommendation system is used to suggest questions that match the student's predicted level of competence. This approach supports a more adaptive and personalized learning environment. The findings demonstrate, SVM algorithm effectively supports intelligent learning systems such as Intelligent Tutoring System (ITS). Future work should include data balancing techniques, expansion of dataset size, and comparative analysis with other algorithms such as Random Forest or K-Nearest Neighbors (KNN) to enhance model generalization.**

Index Terms*— *Academic Performance; Intelligent Tutoring System (ITS); Machine Learning; Question Recommendation; Support Vector Machine (SVM).

I. INTRODUCTION

Technological advancements in education have significantly contributed to the emergence of Intelligent Tutoring Systems (ITS). AI driven platforms designed to deliver tailored instruction by analyzing students' learning behaviors and academic performance. ITS enables more adaptive and individualized learning processes, enhancing students' engagement and understanding [1]. One of the main challenges in ITS development is the system's capability to accurately assess and classify student academic performance. Traditional methods often rely on average test scores or assignment grades, which may not fully represent a student's learning trajectory or personal characteristics [2]. Such static assessments are less effective in dynamic learning environments that require timely and responsive instructional feedback.

Affective Intelligent Tutoring Systems (Affective ITS) have emerged as an advanced approach to personalize learning by incorporating students' emotional responses. As noted by Fernández-Herrero [3], such systems utilize emotion recognition to adapt feedback and content, thereby improving engagement and learning outcomes in real-time educational settings. Liu, Latif, and Zhai [4] conducted a systematic review highlighting recent developments in Intelligent and Robot Tutoring Systems. Their findings emphasize that AI-driven tutoring technologies enhance adaptability and student engagement, while also noting ongoing challenges in scalability, ethics, and cognitive modeling.

Machine Learning (ML) has emerged as a promising solution to this issue by offering objective and data-driven methods for performance analysis. Among various ML algorithms, Support Vector Machine (SVM) is particularly well-regarded for its accuracy and effectiveness in classification tasks. Prior studies [5], have shown that SVM models can accurately predict academic ability based on demographic data and learning styles. Likewise, research [6] demonstrated the application of hybrid AI models combining SVM and Decision Trees for real-time content recommendations in ITS, especially in STEM education. In addition to its widespread application in

educational contexts, the Support Vector Machine (SVM) algorithm has also shown strong performance in text classification and sentiment analysis, [7] successfully implemented SVM combined with Chi-Square feature selection to categorize user feedback into sentiment classes, demonstrating the algorithm's effectiveness in handling large-scale data classification with a reported accuracy of 77%. This reinforces the suitability of SVM for high-dimensional and complex classification tasks, such as predicting student performance levels in adaptive learning environments like Intelligent Tutoring Systems (ITS).

Classify students' academic performance using assignment score data. Students are categorized into three performance levels—Beginner, Intermediate, and Advanced—based on the average of six assignments [8]. Dataset is preprocess through standardization, and model performance is evaluated using metrics such as accuracy, precision, recall, and F1-score [9] indicating strong performance in identifying student categories, particularly within the Intermediate group.

Following classification, the system employs a rule-based recommendation approach to suggest practice questions from a structured question bank. Questions are organized by difficulty level, allowing the system to match students with materials appropriate to their skill level. This integration of classification and recommendation supports personalized learning, enabling ITS to deliver more effective and targeted instruction. In summary, this research presents a hybrid framework combining SVM-based classification with rule-based question recommendation, supporting the broader goal of developing intelligent and adaptive educational systems. The proposed approach not only enhances the diagnostic capabilities of ITS but also improves relevance and impact of the learning content provided to students.

II. METHODOLOGY

This research employs SVM algorithm to perform academic performance classification. [10] explored the use of SVM to classify student learning abilities in system modeling and simulation courses. Their findings demonstrate that SVM-based classification supports effective personalization by aligning instructional content with learners' ability levels, thereby enhancing educational outcomes in technical learning environments.

The methodological workflow begins with data collection and preprocessing, which includes data cleaning, normalization, and splitting the dataset into training and testing subsets. SVM model is trained through a hyperparameter tuning process to optimize its performance. Model evaluation is conducted using a confusion matrix, along with performance metrics such as precision, recall, and F1-score, to assess the model's accuracy and classification effectiveness.

A. Methodological Flowchart

The overall process of the study is summarized in the methodological flowchart shown in figure 1, which outlines main steps from data preparation to personalized question recommendation based on classification outcomes

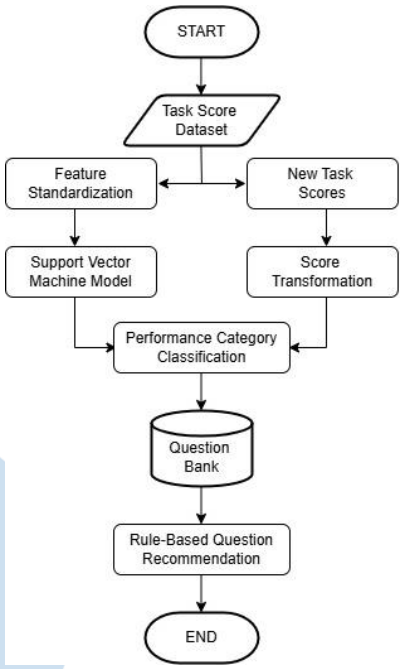


Fig1. Research Flowchart

To visualize the end-to-end process, a flowchart is presented in Figure 1, outlining each major stage in the research methodology—from data processing to the automated recommendation of questions based on classified student performance. The stages are as follows:

1. Start → The system begins with the initialization of the classification and recommendation modules.
2. Assignment Score Dataset → The system receives input in the form of student assignment scores, consisting of six recorded task values for each student.
3. Feature Standardization → These scores are standardized using the StandardScaler method to ensure consistent feature scaling, preventing any single feature from dominating the model training process.
4. SVM Model Training → The classification model is developed using SVM algorithm with a Radial Basis Function (RBF) kernel. The model is trained on the standardized data, with tuning applied to parameters C and gamma to enhance accuracy.
5. New Assignment Scores → The system accepts new input data representing assignment scores of unclassified students.

6. Score Transformation → These new scores undergo the same standardization process as the training data to ensure consistency during prediction.
7. Performance Classification → The standardized data is then classified by the trained SVM model into one of three academic performance levels: Beginner, Intermediate, or Advanced.
8. Question Bank → The system maintains a curated question bank categorized by difficulty levels and subject topics.
9. Rule-Based Question Recommendation → Based on the classification results, the system recommends questions aligned with the student's predicted performance level using a Rule-Based Recommendation strategy that maps student categories to question difficulty.
10. End → The process concludes once the student receives questions tailored to their level, supporting the implementation of a technology-driven adaptive learning system.

B. RBF Kernel in SVM

SVM is a machine learning algorithm that operates within a hypothesis space to identify optimal decision boundaries for classification tasks. SVM is widely used for both binary and multi-class classification problems, where it aims to separate data into distinct classes using a decision boundary known as a hyperplane. If the data is linearly separable, the hyperplane appears as a straight line; otherwise, for non-linear cases, the boundary is curved and more complex [11].

SVM relies on several key parameters that significantly influence the performance of the classification model, namely gamma, cost (C), and kernel function [12].

- Gamma parameter defines how far the influence of a single training example reaches. A low gamma value means that the influence extends far, producing smoother decision boundaries, while a high gamma indicates a more localized influence, which may capture more complex patterns but risks overfitting.
- Cost parameter (C) controls the trade-off between achieving a low error on training data and maintaining a simple decision boundary. A larger C places more emphasis on correctly classifying training examples, potentially at the cost of model generalization.
- Kernel function is responsible for transforming the input data into a higher-dimensional space where a linear separator may be found more easily. This process enables SVM to handle complex classification problems, even when the data is not linearly separable in its original space.

Several types of kernels are commonly used in SVM [13]:

1. Linear Kernel is the simplest form and is used when the dataset is linearly separable. It is computationally efficient and suitable for problems with high-dimensional but sparse features.
2. Polynomial Kernel is applied when the decision boundary is non-linear. It maps the original data into a higher-dimensional space using polynomial functions. This kernel includes a degree parameter (d), which controls the flexibility of the model. However, using a higher degree may lead to less stable performance.
3. RBF Kernel is particularly effective for non-linear datasets. It maps data into an infinite-dimensional space, allowing for complex decision boundaries. The gamma parameter in RBF plays a crucial role in defining the influence of each data point. Compared to other kernels, RBF tends to provide lower classification errors and better generalization in many applications.

RBF kernel is employed in this study due to its effectiveness in handling non-linearly separable data. In the context of academic performance classification, the relationship between features (such as assignment scores and other characteristics) and student performance categories is not always linear. The RBF kernel addresses this challenge by projecting input data into a higher-dimensional space where linear separation is more feasible. Moreover, the RBF kernel is known for its flexibility and empirical performance, making it a reliable choice for classification tasks. It includes the gamma parameter, which controls the influence of individual training samples and helps capture complex local patterns in the dataset. Given these advantages, the RBF kernel was selected to ensure optimal accuracy and efficiency in classifying student academic performance using the SVM model.

C. Confusion Matrix

Confusion Matrix is a common tool for evaluating performance machine learning models and applied binary as well as multiclass classification problems. As noted [14], it provides a tabular summary of prediction outcomes across four key categories as in Figure 2.

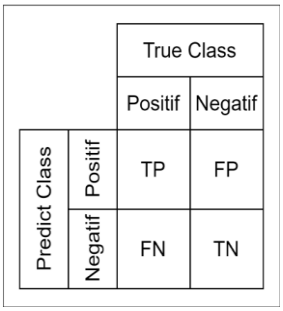


Fig 2. Confusion Matrix

Confusion Matrix provides a structured overview of a model's prediction performance by categorizing outcomes into four components: True Positive (TP) – correctly predicted positive instances; True Negative (TN) – correctly predicted negative instances; False Positive (FP) – incorrect positive predictions; and False Negative (FN) – incorrect negative predictions [15].

Based on these components, several evaluation metrics can be derived to assess classification performance:

1. Accuracy represents the overall proportion of correct predictions and is calculated using the formula:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

2. Precision measures the proportion of positive identifications that were actually correct, defined as:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

3. Recall (also known as sensitivity) reflects the model's ability to retrieve all relevant instances and is given by:

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

4. F1-Score is the harmonic mean of precision and recall, useful for evaluating models with imbalanced classes:

$$F1score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

These metrics provide a comprehensive understanding of how well a classification model performs, especially when dealing with imbalanced datasets.

III. Result And Discussion

Based on training results SVM model, classification performance on test dataset was found to be satisfactory. Model was evaluated using a confusion matrix along with performance metrics such as precision, recall, and F1-score. The results indicate that the model is capable of achieving high accuracy, particularly in the majority class. The obtained F1-score reflects a balanced trade-off between precision and recall, suggesting the model is not only accurate but also reliable in identifying the target categories.

Hyperparameter tuning process played a crucial role in enhancing the model's overall performance, especially in optimizing choice of kernel and regularization parameters. The best classification results were achieved using the Radial Basis Function (RBF) kernel with optimized values of C and gamma, underscoring the importance of careful parameter selection in improving the effectiveness of the SVM algorithm.

Table I delineates the distribution of subjects classified into distinct performance categories as determined by the applied classification methodology. This table offer structured summary of categorization outcomes utilized in analyses.

TABLE I. Category Distribution Table

Category	Count	Percentage
Intermediate	62	51.67%
Beginner	40	33.33%
Advanced	18	15.0%

Dataset utilized in this study is categorized into three levels of student performance: Beginner, Intermediate, and Advanced. The distribution of the data reveals that the Intermediate category is the most dominant, comprising 62 records (51.67%), followed by Beginner with 40 records (33.33%), and Advanced with only 18 records (15%). This class imbalance warrants attention, as it may affect the model's ability to accurately learn and identify instances from minority class—particularly the Advanced category, which contains the fewest data samples.

RBF kernel is widely used in SVM to address non-linear classification problems. Unlike linear kernels, RBF kernel maps data into a higher-dimensional space, allowing classes to be separated by a hyperplane. This transformation enables the capture of complex data patterns, which is critical in applications such as image recognition and medical diagnostics. RBF kernel computes the similarity between data points using a Gaussian function, where the γ (gamma) parameter controls the range of influence of each point. A higher γ results in narrower influence, while a lower γ leads to broader influence across the decision boundary. RBF kernel facilitates efficient computation without explicitly projecting data into the higher-dimensional space. This property reduces computational complexity and improves generalization to unseen data, thereby mitigating overfitting. Due to its flexibility and efficiency, the RBF kernel remains a preferred choice for SVM-based classification tasks involving complex and non-linear relationships.

RBF kernel is specifically for training the SVM model due to its effectiveness in dealing with non-linear data.

1. Feature Normalization:
   The begins with normalizing features using StandardScaler. This step is essential because SVMs are sensitive to the scale of the data. By applying StandardScaler, the data is transformed so that it has a mean of 0 and a standard deviation of 1. This ensures that all features contribute equally to the model training and that the RBF kernel performs optimally without being biased by differences in the scale of features.

2. Using RBF Kernel:
   a. SVC (Support Vector Classification) model is being created with the RBF kernel (svm_model = SVC(kernel='rbf')).
   b. RBF kernel is chosen because it excels at handling complex, non-linear relationships between data points. Unlike linear kernels, which can only separate data that is linearly separable, the RBF kernel maps the data into a higher-dimensional space, allowing SVM to find a hyperplane that can separate data points more effectively.

3. Why RBF Kernel:
   a. RBF kernel is especially suitable when dealing with complex patterns in the data that cannot be separated with a straight line or plane. It is also able to handle high-dimensional data, which is often the case in real-world scenarios. The kernel trick allows the SVM to compute the optimal hyperplane in this higher-dimensional space without explicitly transforming the data, making the process computationally efficient.
   b. In this context, using the RBF kernel allows model to create flexible decision boundaries, ensuring that it can classify the data points in a way that linear kernels would not be able to do effectively. This is why the RBF kernel is preferred in this example.

4. Training the SVM Model:
   Once data is scaled, SVM model is trained using RBF kernel. svm_model.fit(X_train_scaled, y_train) line indicates the model is learning from the training data (X_train_scaled) and their corresponding labels (y_train), with the RBF kernel ensuring that the decision boundaries can adapt to complex data patterns

The model becomes more flexible and can handle non-linear relationships between the data points, providing better performance on real-world datasets that often exhibit complex patterns. Table II presents the classification performance metrics used to evaluate the effectiveness of the proposed model across different learner categories.

TABLE II. TABLE CLASSIFICATION REPORT

Category	Precision	Recall	F1-Score	Support
Advanced	0.67	1.00	0.80	2
Beginner	1.00	0.83	0.91	6
Intermediate	0.94	0.94	0.94	16
accuracy	0.9306	0.9167	0.9189	24
macro avg	0.87	0.92	0.88	24
weighted avg	0.93	0.92	0.92	24

Trained SVM model was evaluated using a test set comprising 24 samples, including 16 Intermediate, 6 Beginner, and 2 Advanced instances. Evaluation results indicate that the model achieved an overall accuracy of 91.67%, with precision of 93.06%, recall of 91.67%, and an F1-score of 91.89%. Per-class classification performance reveals that the model performed best on the Intermediate category, achieving precision, recall, and F1-score values of 0.94. For the Beginner category, the model attained perfect precision (1.00) but a lower recall of 0.83, suggesting that one Beginner instance was misclassified. In the Advanced category, the model achieved a recall of 1.00, indicating it correctly identified all Advanced instances; however, the precision was only 0.67, suggesting that some samples from other categories were incorrectly label as Advanced. Overall, the model yielded a weighted average F1-score of 0.92, reflecting strong general classification performance despite some imbalance across class predictions. The high performance in the Intermediate class aligns with its dominance in the dataset.

Confusion matrix in Figure 3 further supports the evaluation findings by illustrating the distribution of predictions across the actual class labels

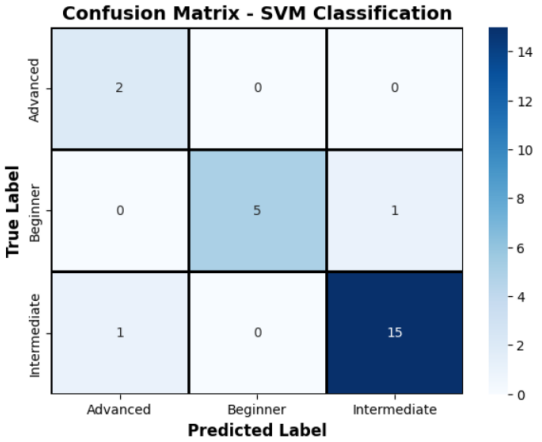


Fig 3. Confusion Matrix Analysis

All Advanced instances (2 samples) were correctly classified with no errors. In the Beginner class, 5 out of 6 samples were accurately identified, while 1 sample was misclassified as Intermediate, accounting for the imperfect recall value. In the Intermediate class, 15 out of 16 samples were correctly predicted, with 1 sample misclassified as Advanced. This indicates that the model occasionally confuses Intermediate with Advanced instances, although its overall performance in the dominant class remains high.

```
import pandas as pd
import numpy as np

# Buat new_data sebagai DataFrame dengan nama kolom yang sama
new_data = pd.DataFrame(np.array([[10, 80, 75, 90, 88, 70]]), columns=X_train.columns)

# Transformasi dengan StandardScaler tanpa peringatan
new_data_scaled = scaler.transform(new_data)
kategori_prediksi = svm_model.predict(new_data_scaled)

print("\n✅ Kategori Prediksi untuk Data Baru:", kategori_prediksi[0])
```

```
✅ Kategori Prediksi untuk Data Baru: Intermediate
```

Fig 4. Student Performance Category Prediction using SVM

Figure 4 illustrates the prediction process of a new student's academic performance category based on their assignment scores. The input data is provided as a numerical array representing six assignment scores (e.g., 100, 80, 75, 90, 88, 70). These values are then converted into a Data Frame with the same structure and feature columns as the training dataset (X_train) to ensure compatibility with the model.

Prior to prediction, the input data is transformed using the pre-fitted StandardScaler to match the scale of the training data. This step ensures that feature distributions remain consistent between training and inference phases. The pre-trained SVM model (svm_model) then performs the classification on the standardized input.

The prediction result places the new student in the Intermediate category, indicating that based on the pattern of their assignment scores, the student is estimated to have a moderate academic performance level. This process demonstrates the practical application of classification models in automatically and objectively categorizing new students based on historical performance data.

Figure 5 showcases a selection of five randomly chosen questions from each academic performance level: Beginner, Intermediate, and Advanced. These questions are retrieved from a curated question bank that categorizes items based on difficulty and topic, such as HTML, CSS, JavaScript, and PHP. The sampling illustrates how the rule-based recommendation system works in practice, delivering learning materials aligned with the student's predicted skill level. By doing so, the system helps support more personalized and adaptive learning, ensuring that students engage with content that matches their current abilities.

```
Level yang tersedia: ['Beginner' 'Intermediate' 'Advanced']

🧠 Soal untuk Level 'Beginner':
    ID                                              Soal          Topik  \
8  Q18  Bagaimana cara membuat daftar tidak terurut (u...           HTML
1  Q02              Buatlah struktur dasar dokumen HTML.  HTML Structure
5  Q15  Apa itu tag <a> dan bagaimana cara membuat tau...           HTML
0  Q01  Apa itu HTML dan apa fungsinya dalam web devel...           HTML
7  Q17  Tulis kode HTML untuk menampilkan gambar denga...           HTML

       Tipe Soal     Level
8   Isian Singkat  Beginner
1  Praktik Coding  Beginner
5   Isian Singkat  Beginner
0   Isian Singkat  Beginner
7  Praktik Coding  Beginner

🧠 Soal untuk Level 'Intermediate':
    ID                                              Soal          Topik  \
18  Q23  Buat program JavaScript yang menampilkan alert...     JavaScript
11  Q07  Buatlah layout grid sederhana menggunakan Flex...     CSS Layout
15  Q20    Jelaskan konsep event handling pada JavaScript.     JavaScript
10  Q06     Apa itu responsive design dan mengapa penting?  CSS Responsive
17  Q22  Apa itu method 'querySelector()' dan bedanya d...  JavaScript DOM

       Tipe Soal        Level
18  Praktik Coding  Intermediate
11  Praktik Coding  Intermediate
15           Essay  Intermediate
10           Essay  Intermediate
17   Isian Singkat  Intermediate

🧠 Soal untuk Level 'Advanced':
    ID                                              Soal          Topik  \
28  Q29  Buat fungsi pencarian real-time dengan JavaScr...      Full Stack
21  Q12  Jelaskan bagaimana cara kerja AJAX dan manfaat...  JavaScript/AJAX
25  Q26  Jelaskan konsep asynchronous dalam JavaScript ...  JavaScript Async
20  Q11  Buat aplikasi sederhana (To-Do List) menggunak...     JavaScript
27  Q28  Implementasikan pagination pada daftar produk ...            PHP

       Tipe Soal     Level
28   Project Mini  Advanced
21           Essay  Advanced
25           Essay  Advanced
20   Project Mini  Advanced
27  Praktik Coding  Advanced
```

Fig 5. Question Bank Categorized by Difficulty Level

The question bank has been systematically categorized based on difficulty level to support the implementation of a recommendation system aligned with the students' classified performance. By leveraging the results from the SVM-based classification, the system is able to recommend questions that match the learner's current academic level Beginner, Intermediate, or Advanced thereby promoting a more adaptive and personalized learning experience

Each question entry includes several key attributes, such as:

- ID: A unique identifier assigned to each question.
- Question: The actual prompt or instruction that students must respond to or complete.
- Topic: The subject matter covered by the question, such as HTML, CSS, or HTML Structure.
- Question Type: The format in which the question is presented, including short answer, coding exercises, or multiple-choice.
- Level: The difficulty level of the question. In this sample, all items are label as Beginner, Intermediate, or Advanced.

Figure 6 implementation of automatic question recommendation system that operates following the classification of students' academic performance

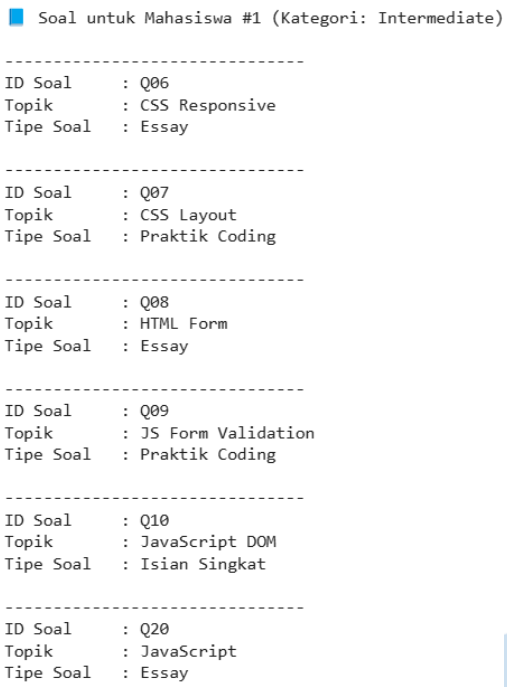


Fig 6. Automatically Recommended Questions Based on Student Classification

This recommendation process based on Rule-Based Recommendation approach, which aligns the performance level predicted by the SVM model with the corresponding difficulty level of questions from the pre-defined question bank.

For example, Student #1, who was classified as Intermediate by SVM model, receives a tailored set of questions that match their predicted performance level. Each recommended question includes key attributes such as the Question ID, Topic, and Question Type. The topics span various areas relevant to the Web Programming course, including CSS Responsive Design, CSS Layout, HTML Forms, JavaScript Form Validation, and JavaScript DOM Manipulation. The types of questions recommended include Essay, Coding Practice, and Short Answer, aimed at evaluating both conceptual understanding and practical application skills.

This rule-based matching system ensures that each student receives questions aligned with their performance category Beginner, Intermediate, or Advanced thereby fostering a more adaptive and personalized learning process. The goal of this approach is to enhance learning effectiveness while maintaining student engagement and motivation by delivering appropriately challenging content.

While using SVM and rule-based recommendation is not entirely new, this study brings a fresh perspective by focusing on real assignment scores as the primary features for classification, something not commonly used in similar works that often rely on broader data like demographics or exam scores. The system then connects the classification results directly to a structured question bank, offering students questions that truly match their learning level. What makes this approach different is its simple yet effective integration of machine learning with real-time instructional support, making it practical for classrooms or online learning environments. By combining these two components into one seamless process, students not only get categorized accurately but also receive personalized learning materials instantly. This tight coupling between prediction and recommendation adds real value to intelligent tutoring systems and helps move them closer to being truly adaptive and supportive in day to day learning.

IV. CONCLUSIONS

This research demonstrates SVM algorithm with RBF kernel can be effectively applied to classify students' academic performance into three categories: Beginner, Intermediate, and Advanced.

Preprocessing procedures most notably data normalization and dataset partitioning, contributed substantially to the improved performance of classification model. Evaluation results indicate the implemented Support Vector Machine (SVM) classifier achieved robust performance, with accuracy of 91.67%, precision of 93.06%, recall of 91.67%, and an F1-score of 91.89%. Among the classified categories, the model exhibited its highest performance in identifying the Intermediate group, which constituted the majority of instances in dataset. Conversely, lower performance was observed the Advanced category, potentially attributable to limited representation of samples in class.

Subsequent the classification process, SVM outputs were integrated into a Rule-based Question Recommendation mechanism. This component leveraged predicted performance categories to guide the selection of learning materials tailored to each learner's proficiency level. The findings collectively demonstrate the efficacy of combining machine learning-based classification with rule-based instructional support, underscoring the model's potential to enhance personalized learning within intelligent educational systems, particularly Intelligent Tutoring Systems (ITS).

For future work, several improvements are recommended:

- Data balancing techniques such as SMOTE (Synthetic Minority Over-sampling Technique) or undersampling can be applied to address class imbalance and improve performance on underrepresented categories.
- Alternative classification models such as Random Forest, K-Nearest Neighbors (KNN), and Gradient Boosting should be explored for comparative analysis to better understand their effectiveness on the same dataset.
- Expanding dataset size, especially for minority classes like Advanced, is crucial to

help the model learn better feature representations and improve generalization.

- Real-world application of the model in adaptive learning systems should also be considered, where the classified student level can guide the delivery of customized learning materials or assignments tailored to individual needs.
- Moreover, advanced evaluation strategies such as the Three-Way Confusion Matrix [14] may be incorporated to manage uncertain or borderline predictions. This would enhance the reliability of classification results and support more personalized decision-making within intelligent tutoring systems.

REFERENCES

[1] "AI in Education: The Rise of Intelligent Tutoring Systems | Park University." Accessed: Mar. 12, 2025. [Online]. Available: https://www.park.edu/blog/ai-in-education-the-rise-of-intelligent-tutoring-systems/

[2] C. C. Lin, A. Y. Q. Huang, and O. H. T. Lu, "Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review," *Smart Learn. Environ.*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/S40561-023-00260-Y.

[3] J. Fernández-Herrero, "Evaluating Recent Advances in Affective Intelligent Tutoring Systems: A Scoping Review of Educational Impacts and Future Prospects," 2024, doi: 10.3390/educsci14080839.

[4] V. Liu, E. Latif, and X. Zhai, "Advancing Education through Tutoring Systems: A Systematic Literature Review," Mar. 2025, Accessed: Apr. 22, 2025. [Online]. Available: https://arxiv.org/abs/2503.09748v1

[5] B. I. Nugroho, N. A. Santoso, and A. A. Murtopo, "Prediksi Kemampuan Akademik Mahasiswa dengan Metode Support Vector Machine," *remik*, vol. 7, no. 1, pp. 177–188, Jan. 2023, doi: 10.33395/REMIK.V7I1.12010.

[6] B. H. Rao, M. Rengarajan, R. Changala, K. Santosh, M. Aarif, and M. Ponni Valavan, "Hybrid AI Approach Combining Decision Trees and SVM for Intelligent Tutoring Systems in STEM Education," *10th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2024*, pp. 574–579, 2024, doi: 10.1109/ICACCS60874.2024.10716888.

[7] L. Luthfiana, J. C. Young, and A. Rusli, "Implementasi Algoritma Support Vector Machine dan Chi Square untuk Analisis Sentimen User Feedback Aplikasi," *Ultim. J. Tek. Inform.*, vol. 12, no. 2, pp. 125–126, 2020, doi: 10.31937/ti.v12i2.1828.

[8] S. Bulut, "A Comparison of the First Three Flute Lessons with Beginner and Intermediate-advanced Level Flute Students," *Procedia - Soc. Behav. Sci.*, vol. 177, pp. 229–234, Apr. 2015, doi: 10.1016/J.SBSPRO.2015.02.397.

[9] M. Armanda, F. Adline, T. Tobing, and U. M. Nusantara, "Implementation of Support Vector Machine Method for Twitter Sentiment Analysis Related to Cancellation of u-20 World Cup in Indonesia," vol. 11, no. 1, 2024.

[10] C. Liu, S. Yang, and M. Engineering, "Personalized Learning Ability Classification Using SVM for Enhanced Education in System Modeling and Simulation Courses," vol. 1, no. 4, pp. 295–307, 2024.

[11] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," vol. 6, no. 3, 2021, Accessed: Mar. 12, 2025. [Online]. Available: http://situs.com

[12] A. Zaiem Praghakusma and N. Charibaldi, "Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi)," vol. 9, no. 2, pp. 33–42, 2021, doi: 10.12928/jstie.v8i3.xxx.

[13] "Apa itu Kernel Trick? Pengertian dan Jenis-jenis Fungsi Kernel SVM - Trivusi." Accessed: Mar. 12, 2025. [Online]. Available: https://www.trivusi.web.id/2022/04/fungsi-kernel-svm.html

[14] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Inf. Sci. (Ny).*, vol. 507, pp. 772–794, Jan. 2020, doi: 10.1016/J.INS.2019.06.064.

[15] S. Sathyanarayanan and B. R. Tantri, "Confusion Matrix-Based Performance Evaluation Metrics," no. November, 2024, doi: 10.53555/AJBR.v27i4S.4345

ISSN 2085-4552

A Recommendation System for Prewedding Location Selection using Count Vectorization and Cosine Similarity

Dede Kurniadi[1], Ilham Ahmad Maulana[2]
[1,2] Department of Computer Science, Institut Teknologi Garut, Garut, Indonesia
[1] dede.kurniadi@itg.ac.id, [2]2006118@itg.ac.id

Accepted 28 May 2025
Approved 04 June 2025

***Abstract*— Choosing the right pre-wedding location is a concern for many couples due to the many options available, which often causes confusion and is time-consuming in decision-making. Therefore, a recommendation system is needed to assist couples in determining the prewedding location that suits their preferences. This research aims to provide alternative recommendations for prewedding locations and simplify the process of selecting a suitable location. This system integrates the Content-based Filtering method by applying Count Vectorization and Cosine Similarity to calculate the similarity between locations based on features in the dataset. These algorithms were chosen for their effectiveness in handling categorical text data and identifying similar items, making them suitable for generating relevant prewedding location recommendations. The dataset used consists of 75 entries, each representing a unique location with the following variables: venue, address, distance (in kilometers), price, theme, and type. Based on the similarity scores, the system generates a ranked list of 5 recommended locations that closely match the selected venue. The Rapid Throwaway Prototyping method ensures the system development is done iteratively and involves direct feedback from users. The recommendation system is evaluated using the Mean Reciprocal Rank (MRR) metric to measure the effectiveness of the recommendations provided by the system. The results show that the developed prewedding location recommendation system can provide relevant location recommendations with good performance, as evidenced by the Mean Reciprocal Rank (MRR) value of 0.88, which indicates that the system is effective in placing the most relevant locations at the top of the recommendation list. The high MRR value shows the system's effectiveness in providing relevant recommendations, improving customer experience, and supporting the company's competitiveness in the prewedding documentation industry.**

Index Terms*— *Content-based Filtering; Cosine Similarity; Count Vectorization; Prewedding; Recommendation System.

I. INTRODUCTION

Marriage is a sacred event of binding the marriage vows made by a couple to formalize the marriage bond based on applicable religious, legal, and social norms [1]. In addition, there is a procession that is usually carried out before the wedding, namely a prewedding photo session[2]. Photo sessions at the moment before and during the wedding for most are important things that must be done because these moments may not be repeated[3]. As a moment that precedes the wedding day, these sessions are often used to express the unique character, theme, or love story of each couple[4]. In the documentation of these moments, especially prewedding, location selection is a crucial factor that can affect the results of both photo and video documentation[5]. The right location can not only enhance visual aesthetics but also add meaningful value to a moment[6].

Hazen Pictures, a company focusing on photography and videography services for event documentation, understands the importance of choosing the right venue for client satisfaction. Hazen Pictures, located on the island of Batam with its abundant natural beauty, is committed to providing exemplary service to couples by providing aesthetically pleasing venues that match the preferences and desires of couples in their pre-wedding moments. In addition, Hazen Pictures strives to be a reliable partner for couples to create meaningful prewedding documentation for each couple. But prewedding location selection is still challenging for every couple and Hazen Pictures. Most of Hazen Pictures' clients experience confusion due to the variety of location options. From exotic beaches with beautiful white sands, green parks, urban locations, or charming cities, each location offers a different atmosphere and aesthetic. Therefore, a recommendation system is needed to assist in making decisions about a prewedding location and provide a list of recommendations based on each couple's preferences.

The correct algorithm is needed in making the recommendation system [7].

A study by [8] focused on building a restaurant recommendation system to help users identify dining places that suit their preferences, such as type of cuisine and proximity. The method used in this study is content-based filtering, supported by cosine similarity to measure the closeness between restaurant attributes like address and category. This technique allows the system to provide tailored recommendations for each user. To support the system, data was collected through web scraping from popular platforms such as TripAdvisor and Google Maps, resulting in a diverse and representative dataset of restaurants. The system was developed as a web application and evaluated through trials involving 30 participants. The testing showed that the system achieved an accuracy of 88%, proving its capability to deliver relevant and useful restaurant recommendations based on user preferences.

The second research conducted by [9] focusing on developing a recipe recommendation system aimed at assisting housewives in deciding what to cook based on available ingredients. The researcher implemented a content-based filtering approach using Term Frequency-Inverse Document Frequency (TF-IDF) and cosine similarity in this study. The TF-IDF method was applied to numerate ingredient data, while cosine similarity was used to measure the similarity between user-input ingredients and existing recipes. The dataset used in this study consisted of 30 recipes obtained from the website makapahariini.com. The system is designed to recommend recipes most relevant to the ingredients users enter through a search form. Based on the Root Mean Square Error (RMSE) evaluation, the system scored 0.356, indicating that the recommendation results were fairly accurate and relevant in suggesting suitable recipes.

The third research was conducted by [10] which focused on developing a web-based application called Nongkies to recommend coffee shops and restaurants as hangout spots, primarily targeting university students in Bandung. It used a content-based filtering approach with cosine similarity to match user preferences with place attributes. The dataset included 55 coffee shop entries from the Trakter.id website. Results showed the system provided accurate and relevant recommendations, with 52.3% of students hanging out once a week and an average satisfaction score of 4.153 out of 5. The system proved effective in helping users discover suitable locations, improving their decision-making and social experiences.

The fourth research conducted by [11] he made a recommendation system for choosing an electric car in his research. This research uses content-based filtering by comparing the TF-IDF method with Count Vectorization. Both methods have the same two functions, namely obtaining the number of occurrences of words in the document. The difference is that TF-IDF, in addition to calculating the occurrence of words this method performs weighting on each word that appears in the document to obtain its level of importance. The attributes include model, price, manufacturer, range, max-speed, horsepower, car type, drive type, release year, and manufactured. The results of this study show that the system can provide electric car recommendations for users. In terms of accuracy in this study using a confusion matrix with the results for TF-IDF, the accuracy value is 64%, while for Count Vectorization, the accuracy value is 75%.

The fifth research conducted by [12] in his research, he made a recommendation system for choosing a movie. This study uses two methods, namely the first method of content-based filtering, which consists of count vectorizer and cosine similarity. The second method is collaborative filtering. Both methods are methods used for recommendation systems. Content-based filtering is used to generate similar recommendation items, and then collaborative filtering is used to generate recommendations that can reduce errors and be more personalized. The attributes used in this research regarding movie recommendation systems include genres, cast, keywords, directors, titles, users, and ratings. This research shows that the hybrid filtering method produces a small error value of 0.68 and can present relevant user recommendations.

Based on research references, this research will create a recommendation system for prewedding location selection at Hazen Pictures with novelty from research on prewedding location selection before, so using the Count Vectorization algorithm is an option because based on research [11], in his research comparing the TF-IDF and Count Vectorization methods to calculate the occurrence of words in a document, he found that the Count Vectorization method is better than the TF-IDF method in terms of accuracy. After encoding the count vectorization method, calculations with cosine similarity are performed to find similarities in the data in the system. Data similarity is obtained from the calculation of the closest value distance [12]. This study adopts the Rapid Throwaway Prototyping method, which facilitates the iterative development of the recommendation system through early prototyping, user feedback evaluation, and continuous improvement [13]. The data used includes prewedding locations around Batam Island with attributes such as venue, address, distance (in kilometers), price, theme, and type, all sourced from Hazen Pictures. The recommendation system will be implemented using a content-based filtering approach incorporating Count Vectorization and Cosine Similarity. System evaluation will be carried out using Mean Reciprocal Rank (MRR), chosen for its effectiveness in assessing the top-ranked relevance of the recommendations provided [14]. This research aims to develop a recommendation system that provides relevant prewedding location suggestions based on client preferences, supports Hazen Pictures and their clients in making faster and more informed decisions,

and streamlines the selection process by offering ranked recommendations based on similarity calculations.

II. METHODOLOGY

The Rapid Throwaway Prototyping Model method was used in this study, which is one variant of the prototyping model [15]. This method allows rapid prototyping to provide a visual picture of how the system works [16]. The stages involved in this method include formulating outline requirements, developing a prototype, evaluating a prototype, specifying the system, developing software, and validating the system.

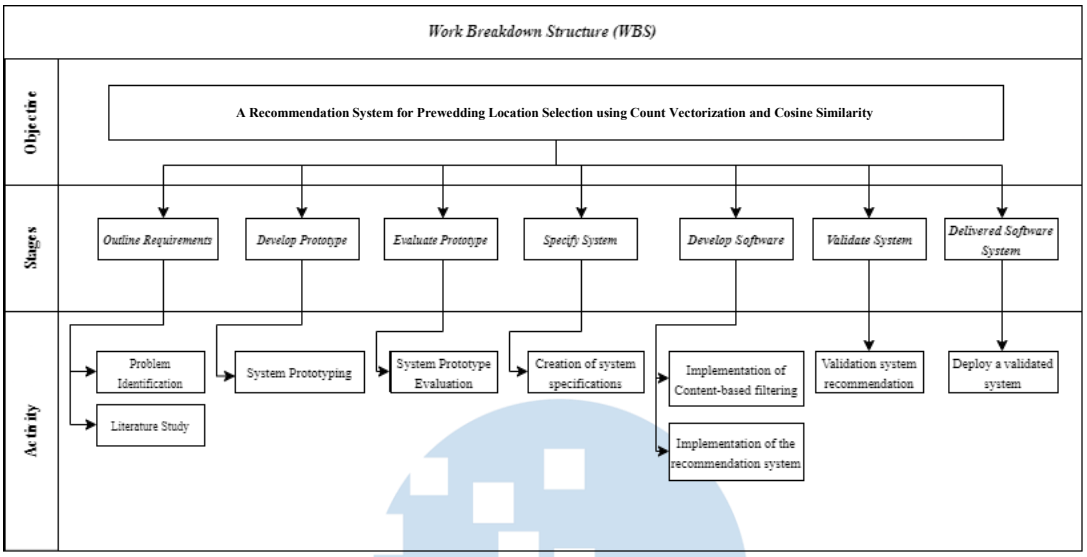


Fig. 1. Research Stages

Based on Figure 1, the research stages can be explained as follows:

1) Outline Requirements, at this stage, the identification process is carried out by understanding the problems at Hazen Pictures regarding selecting prewedding locations by conducting interviews. In addition, literature studies from journals, books or other related support are also carried out.
2) Develop a Prototype. At this stage, an initial prototype is made to produce a system prototype based on the problems identified.
3) Evaluate Prototype, this stage evaluates the system prototype with Hazen Pictures to get feedback on the suitability of the design to the needs.
4) Specify System, at this stage, specifications are made for the system to be built based on the evaluation stages carried out on the earlier prototype.
5) Develop Software. At this stage, the activity of implementing Content-based Filtering by implementing Count Vectorization to convert text into numeric by counting the occurrence of words in the text [11]. The count vectorization formula is presented below.

$$Count\ Vectorizer = V(i,j) \quad (1)$$

With V=number of frequencies; i=word; and j=document.

Furthermore, the results of Count Vectorization are calculated using Cosine Similarity to calculate the similarity between locations [17]. The cosine similarity formula is presented as follows.

$$A \cdot B = A[0] \times B[0] + \cdots + A[n] \times B[n] \quad (2)$$

$$\sqrt{A} = \sqrt{(A[0]^2 + A[1]^2 + \cdots + A[n]^2)} \quad (3)$$

$$\sqrt{B} = \sqrt{(B[0]^2 + B[1]^2 + \cdots + B[n]^2)} \quad (4)$$

$$\frac{A \cdot B}{\sqrt{A} \times \sqrt{B}} \quad (5)$$

A and B are two vectors or documents. Next, the recommendation system is built based on the system specification and the evaluated prototype.

6) Validate System: This stage is carried out to validate the system that has been built with Hazen Pictures to find out whether the recommendation system has met the needs and specifications of the system. And measuring performance on recommendation results using the Mean Reciprocal Rank metric. The formula for Mean Reciprocal Rank is presented as follows:

$$RR = \frac{1}{First\ relevan\ position} \quad (6)$$

$$MRR = \frac{1}{N} \sum_{i=1}^{N} RR_i \quad (7)$$

7) Delivered Software System, at this stage, the application is deployed and submitted to Hazen Pictures.

III. Result And Discussion

A. RESULTS

1) *Outline Requirements*

The first stage is outline requirements, where two activities are carried out: problem identification and literature review.

a) Problem Identification: At this stage, it was discovered that although Batam has many attractive prewedding locations, the abundance of choices confused clients when choosing the most suitable location. Hazen Pictures plans to add a prewedding location recommendation feature on its website to assist clients in choosing a location based on the desired theme. This feature is expected to simplify the selection process, although the final decision remains in the hands of the client. Some of the simple system requirements identified include:

- The system needs to display information about each location.
- The system can help provide alternative locations according to preference.

b) Literature Study: A literature study was conducted to gain in-depth insight into relevant technologies and methods in developing the recommendation system. This activity includes reading and reviewing related literature to understand the problem and reviewing approaches used in previous research to address similar problems. Count Vectorization allows the system to measure the occurrence of words in location descriptions, while Cosine Similarity calculates the similarity between client preferences and available locations. In addition, the Mean Reciprocal Rank (MRR) evaluation metric was chosen for its ability to measure the performance of the recommendation system, especially in assessing the relevance of the generated location list and ensuring that the most relevant recommendations are at the top.

2) *Develop Prototype*

This stage is carried out by making a prototype of the system that has been planned previously. This step is carried out based on the need to visualize and test how the recommendation system's flow will work. The prototype was created using Figma design software, and the following is an iteration process in prototyping, which includes iterative stages to adjust the design to the needs, as seen in Figure 2.

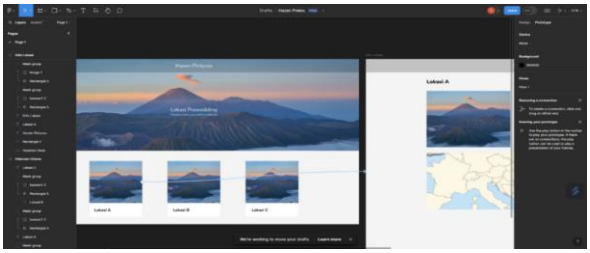


Fig. 2. Prototyping Process

Figure 2 shows the prototyping process, which was done in two iterations. Once the prototype was completed and evaluated based on feedback from users and Hazen, the evaluation was used to refine the second prototype. The prototype focused on testing the basic concepts of the system, mainly navigation, and functionality, with the results shown in Figure 3 and Figure 4.

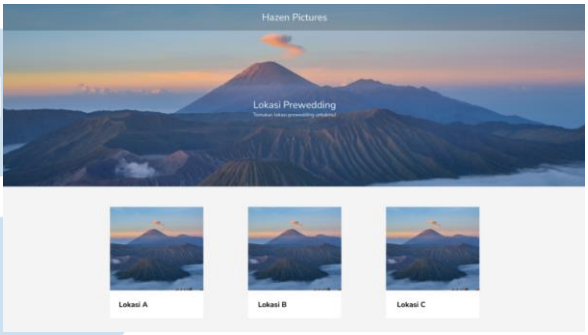


Fig. 3. First prototype of the home page

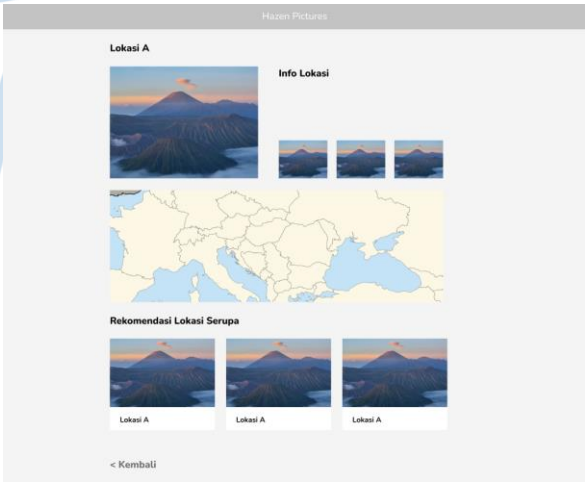


Fig. 4. Prototype of the location page

Figures 3 and 4 show the prototype which includes two main pages: a home page with a list of prewedding locations and a location page that displays information and recommendations of similar locations. Based on the evaluation and feedback, a second prototype was developed to correct shortcomings and improve features, with the improvements shown in Figures 5 and 6.

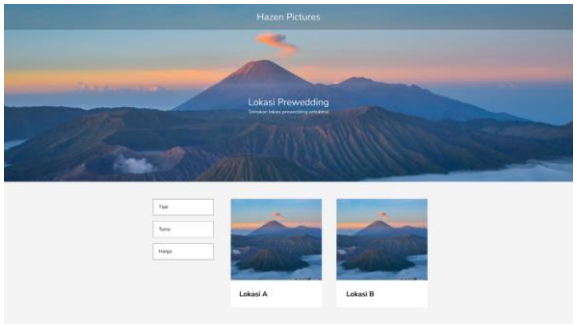


Fig. 5. Second prototype of the home page

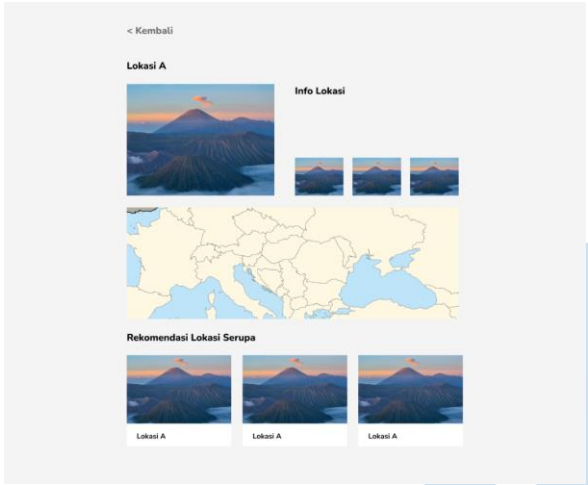


Fig. 6. Second prototype of the location page

The system prototype created in Figures 5 and 6 can be accessed through the link https://s.id/hazenlocprototipe. Based on the prototype results, the prototype consists of two pages. The first page displays all prewedding location queries, and users can view all prewedding locations and sort them by theme, type, and price based on user preferences. The second page displays info from the prewedding location and 5 recommendations that are similar to the location the user views.

3) *Evaluate Prototype*

The evaluation prototype stage evaluates the previous prototype by researchers and Hazen Pictures. This evaluation stage aims to update and get input from users to proceed to the next stage. The prototype evaluation process is carried out online as shown in Figure 7.

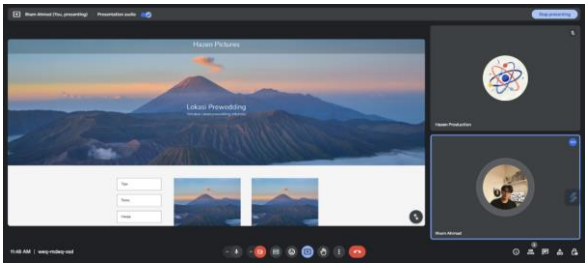


Fig. 7. Prototype Evaluation

Figure 7 shows an evaluation of the prototype conducted online via Google Meet between the researcher and Hazen Pictures. Discussions covered the functionality and design of the prototype, with user feedback used for improvement. The evaluation was conducted twice. The following input is based on the first evaluation stage, including:

- Added dropdowns to sort locations by theme, type, and price.
- Remove the navbar from the location page.
- Moved the position of the back button on the location page.

The prototype was improved according to the feedback and re-evaluated. The results of the second evaluation showed that the prototype had met the criteria and was ready to proceed to the Specify System stage.

4) *Specify System*

The specified system stage is when researchers compile a recommendation system specification based on prototype evaluation and user needs analysis. This specification guides the development of a recommendation system ready to be implemented. The results of this stage are functional requirements, which include:

- The system can display all prewedding locations
- When a location card is clicked, the location info and a list of recommendations should be displayed at the bottom.
- The recommendation system displays a list of alternative recommendations for a maximum of 5 prewedding locations.
- The system can sort prewedding locations by type, theme, and price.

In addition, non-functional requirements are also determined, namely, the user interface must be easy to use by users.

5) *Develop Software*

The development software stage transforms the prototype and specifications into a functional recommendation system. Researchers implement the algorithm along with the features that have been planned previously. The following are the steps taken at the development software stage.

a) Implementation of Content-based Filtering (CBF): Content-based filtering (CBF) is a recommendation method that provides suggestions based on the similarity of attributes or content between items that users and other items in the dataset have selected [18]. The application of the Content-based Filtering method in the

prewedding location recommendation system is used to provide recommendation results that have similarities in items [19]. The following are the implementation steps of Content-based Filtering presented in Figure 8.

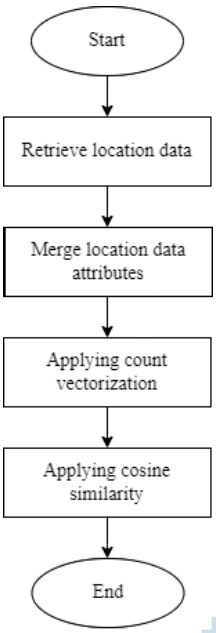


Fig. 8. Implementation steps

Based on Figure 8. The process starts with retrieving location data. After that, the attributes in the data are combined into one text. Next, the merged text data is converted into a numerical representation by counting the number of word occurrences [20] using Count Vectorization. After the numerical representation is determined, the data's similarity is calculated using Cosine Similarity. The following is a complete explanation of the implementation of content-based filtering.

1. Retrieve and merge location data
   Each attribute of the location data is merged into one text. This merging aims to represent each location comprehensively in text form. The following is a sample of data combined in Table I.

TABLE I. SAMPLE DATA FOR COUNT VECTORIZATION

No	Document
1	Viovio Beach Outdoor Beach Sijantung Galang Batam City Riau Islands
2	Melur Beach Outdoor Beach Sijantung Galang Island Batam City Riau Islands

Table I presents sample data from two documents, which is then used as input for the count vectorizer process.

2. Applying Count Vectorizer
   The merged text is converted into a numerical representation using Count Vectorization. This method counts the number of word occurrences in each document (location) and produces a word-document matrix representing the frequency of word occurrences in each location. The following are the results of the application of Count Vectorization presented in Table II.

TABLE II. COUNT VECTORIZATION RESULTS

No	Word	Document	
		1	2
1	Batam	1	1
2	City	1	1
3	Galang	1	1
4	Islands	1	1
5	Melur	0	1
6	Outdoor	1	1
7	Pantai	2	2
8	Pulau	0	1
9	Riau	1	1
10	Sijantung	1	1
11	Viovio	1	0

Table II displays the results of count vectorization, where each unique word that appears in the document is given the value of the frequency of occurrence in each document. From the table, the vectors formed are, $D1 = [1,1,1,1,0,1,2,0,1,1,1]$ and $D2 = [1,1,1,1,1,1,2,1,1,1,0]$.

3. Applying Cosine Similarity
   After the Count Vectorization results are prepared, the next step is to calculate the similarity between the two documents using equations (2), (3), (4), and (5). The calculated similarity of 0.88 indicates a fairly high level of similarity between the two documents or items being compared.

After knowing how the similarity between locations is calculated, the system process for providing prewedding location recommendations is shown in Figure 9.

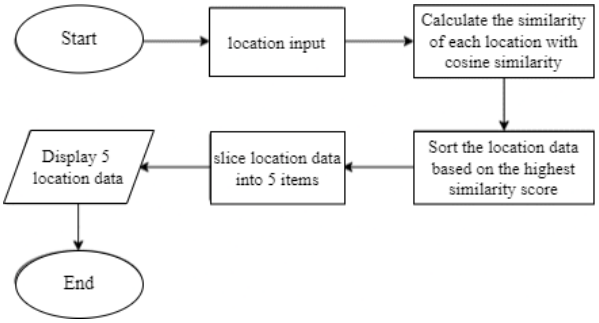


Fig. 9. Recommendation system flow

Based on Figure 9, the system flow provides the following recommendations:

- The system receives input in the form of a prewedding location
- The system looks for a similarity score against all other locations in the data.
- The similarity score calculation uses Cosine Similarity.
- The list of prewedding location recommendations will be sorted based on the highest or most similar similarity score.
- The system will set restrictions on the list of prewedding locations. Only five alternative prewedding locations are displayed.
- The system will present feedback on five alternative prewedding location recommendations.

b) Implementation of the recommendation application: At this stage, coding activities are carried out, which include implementing Content-Based Filtering using Count Vectorization, calculating Cosine Similarity, and creating functions to generate recommendations. The following is the result of the recommendation system presented in Figure 10.

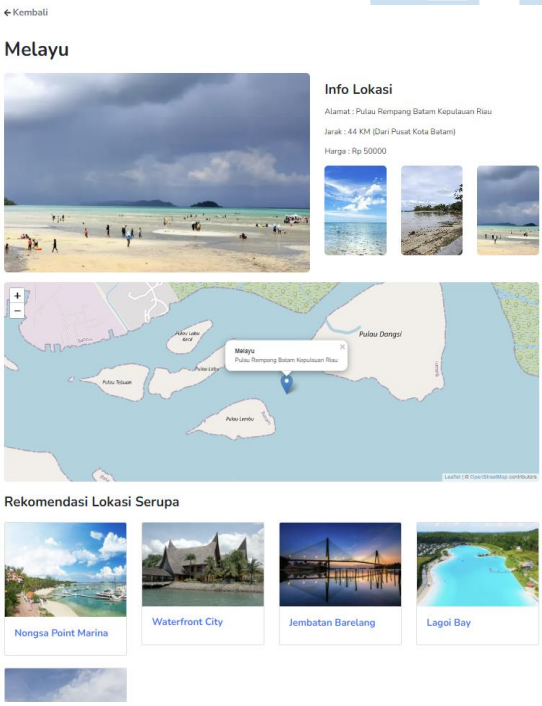


Fig. 10. Result of the recommendation system

On this page, the system will display information about the location after the user selects and enters one of the prewedding location pages. In addition, to enhance the user experience, the system will also display a list of similar prewedding location recommendations. In this way, users can easily find other relevant options that match their preferences, assisting them in making more informed decisions and expanding their choices.

6) *Validate System*

The validate system stage is where researchers test the recommendation system that has been developed to evaluate whether the system functions as expected and meets user needs. The company conducted experiments on the recommendation system that had been built. The results of the test can be seen in Table III.

TABLE III. TEST RESULTS

No.	Test Activity	Test Result	Status
1	Access the home page	All prewedding locations are displayed	Successfully
2	Click on a prewedding location	The prewedding location info is displayed, and a list of 5 similar prewedding locations is shown	Successfully
3	Click the dropdown to select a location prewedding by theme	Prewedding locations by theme theme are displayed	Successfully
4	Click the dropdown and select prewedding location by type	Prewedding location by type displayed	Successfully
5	Click the dropdown and select prewedding locations by price	Prewedding locations by price from lowest to highest and vice versa are displayed	Successfully
6	Interface navigation test	Users can access all pages without difficulty.	Successfully

The table above shows that the prewedding location recommendation system functions as expected is easy to use and meets the functional and non-functional requirements that have been set.

Testing the performance of the recommendation system in producing a list of alternative recommendations in the first relevant order [14] using Mean Reciprocal Rank in equations (6) and (7). The results of the MRR calculation are presented in Table IV.

TABLE IV. MRR CALCULATION RESULTS

No.	Location	RR
1	Nongsa Point Marina	0.2
2	Ocarina Park	1
3	Pantai Viovio	1
4	Mangrove Pandang Tak Jemu	0
5	Waterfront City	1
6	Pantai Melayu	1
7	Jembatan Barelang	1
8	Lagoi Bay	0.5

No.	Location	RR
9	Laguna	1
10	Ria Bintan	0.5
…	…	…
75	Jembatan Barelang 6	1
Mean Reciprocal Rank		**0.88**

The test results using the Mean Reciprocal Rank metric are known to be 0.88, which shows that, on average, relevant recommendation results are found at the top of the recommendation list. This high MRR metric system can provide appropriate recommendations [21] satisfy user needs, and ensure a better user experience in choosing a prewedding location.

7) *Delivered Software System*

The validated recommendation system is handed over to the company for operational use. The system is uploaded to the chosen host platform, and access is granted to the company so they can use it directly. After the system is handed over, initial monitoring is done to ensure the system is functioning correctly and delivering the results expected by the company. The recommendation system can be accessed through https://s.id/hazenpreloc.

B. DISCUSSION

This research successfully developed a recommendation system for prewedding location selection using the Content-based Filtering recommendation method by applying Count Vectorization and Cosine Similarity calculations. These two things are used to convert and calculate the similarity between prewedding locations based on location features in the dataset. Using the prewedding location data from Hazen Pictures, the recommendation system can provide recommendations and compile a list of relevant locations. The implementation of Count Vectorization and Cosine Similarity worked quite well in measuring the similarity between different locations.

This research aligns with the challenges identified in the background, where prewedding location selection can be difficult and time-consuming due to the many available options. This research is consistent with the findings of [3], who noted challenges in location selection but did not use clear evaluation metrics. This study makes up for that shortcoming by applying Count Vectorization, Cosine Similarity, and Mean Reciprocal Rank. In addition, this research differs from research by [2], which uses the Vikor and Borda methods but is not implemented in a user-accessible system. This research develops an algorithm-based model implemented in a practical system, provides solutions that can be used immediately, and offers improvements and additions in prewedding location recommendation technology.

Based on the test results with the company, the recommendation system functions by expectations and meets the needs that have been set. The evaluation results using the MRR metric of 0.88 show that, on average, relevant recommendation results are found at the top of the list. This high MRR metric allows the system to provide recommendations that match and satisfy user needs and ensure a better user experience when choosing a prewedding location.

With the results of this research, the recommendation application is expected to help brides-to-be choose a prewedding location that suits their wishes, reduce confusion, and speed up the decision-making process.

IV. CONCLUSIONS

Based on the research results, the prewedding location recommendation system implemented with the Content-based Filtering method using Count Vectorization and Cosine Similarity effectively provides relevant recommendations. The system successfully converts and calculates similarities between prewedding locations based on features in the Hazen Pictures dataset, resulting in recommendations that match user preferences. Evaluation using the Mean Reciprocal Rank (MRR) metric with a value of 0.88 indicates that the recommended locations are high on the list, which makes it easier for brides-to-be to choose suitable prewedding locations and speeds up the decision-making process. For future research and development of recommendation systems, it is recommended that the dataset be expanded by adding more location features, number of locations, and user reviews. This will help improve the relevance and personalization of recommendations. In addition, combining multiple recommendation methods can improve system performance and provide results better suited to individual user needs.

REFERENCES

[1] E. M. Rosa, "Analisis Fenomena Budaya Foto Pre-Wedding Di Masyarakat: Studi Takhrij dan Syarah Hadis," *J. Ris. Agama*, vol. 1, no. 1, pp. 223–232, 2021, doi: 10.15575/jra.v1i1.14405.

[2] A. Syahputra, "Sistem Pendukung Keputusan Pemilihan Lokasi Pre-Wedding di Kota Medan dengan Menggunakan Metode VIKOR dan BORDA," *J. Sist. Komput. dan Inform.*, vol. 1, no. 3, p. 207, 2020, doi: 10.30865/json.v1i3.2159.

[3] M. R. Syahputra, R. Winanjaya, and H. Okprana, "Sistem Pendukung Keputusan Pemilihan Lokasi Prewedding Menggunakan Metode Weight Product," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 695–701, 2019, doi: 10.30865/komik.v3i1.1680.

[4] M. R. Maraghi, U. Nadhiroh, and U. Jatmiko, "Pengaruh Lokasi, Harga, Dan Digital Marketing Terhadap Minat Sewa Jasa Yunida Wedding Organizer," *MUSYTARI Neraca Manajemen, Akunt. Ekon.*, vol. 1, no. 2, pp. 1–13, 2023.

[5] J. Abidin and P. Kasih, "Simple Additive Weighting dalam Sistem Bantu Pemilihan Paket Foto Prewedding di JF Creative Media," *Univ. Nusant. PGRI Kediri. Kediri*, vol. 1, p. 180, 2022.

[6] A. Amri, "Prewedding Photo Procession And The Role Of The Family In Them," vol. 10, no. 2, pp. 246–264, 2020.

[7] R. T. Yogatama Wijaya and Y. S. Nugroho, "Sistem Rekomendasi Pemilihan Paket Foto Pernikahan Di Scaper Cinema Photo Mengguna," pp. 1–17.

[8] F. Christyawan, A. N. Rohman, and A. D. Hartanto, "Application of Content-Based Filtering Method Using Cosine Similarity in Restaurant Selection Recommendation System," vol. 6, no. 3, pp. 1559–1576, 2024, doi: 10.51519/journalisi.v6i3.806.

[9] N. N. K. Sari, R. Priskila, and P. B. A. A. Putra, "Implementasicontent-Based Filtering Menggunakan Tf-Idf And Cosine Similarity Untuk Sistem Rekomendasi Resep Masakan," *J. Teknol. Inf.*, vol. 18, no. 1, pp. 43–51, 2024.

[10] A. Raihan, A. Ibrahim, and A. Ghozali, "Hangout Places Recommendation System Using Content-Based Filtering and Cosine Similarity Methods Abdul," *J. Dinda*, vol. 4, no. 2, pp. 35–42, 2024.

[11] M. Zaynurroyhan, A. Pambudi, and S. Artikel, "Perbandingan TF-IDF dengan Count Vectorization Dalam Content-Based Filtering Rekomendasi Mobil Listrik," *J. Keilmuan dan Apl. Tek. Inform.*, vol. 5, no. 36, pp. 8–15, 2023.

[12] A. Al Amin, A. Sunyoto, and H. Al Fatta, "Improve Quality of Recommendation System Using Hybrid Filtering Approach," *Proc. Int. Conf. Innov. Sci. Technol. (ICIST 2020)*, vol. 208, no. Icist 2020, pp. 103–106, 2021, [Online]. Available: https://doi.org/10.2991/aer.k.211129.023

[13] S. Jacobs and D. Eck, "Prototyping: Making design concepts real and testable," *Hear. Voice Cust.*, pp. 113–130, 2003.

[14] G. Yulianta, Y. Setiawan, and D. Puspitaningrum, "Information retrieval system pada pencarian dokumen digital pada smartphone menggunakan algoritma boyer moore," vol. 4, no. 2, pp. 236–244, 2020.

[15] R. S. Pressman and B. R. Maxim, *Software Engineering : A Practitioner's Approach*. 2001.

[16] D. Kurniadi, A. Sutedi, D. Nursyaban, A. Mulyani, K. Garut, and P. Korespondensi, "Sistem Rekomendasi Pemilihan Pengepul Limbah Di Pt. Pituku Cordova International Menggunakan Algoritma Haversine," vol. 11, no. 1, pp. 155–166, 2024, doi: 10.25126/jtiik.20241117694.

[17] M. Yusuf and A. Cherid, "Implementasi Algoritma Cosine Similarity Dan Metode TF-IDF Berbasis PHP Untuk Menghasilkan Rekomendasi Seminar," *J. Ilm. Fak. Ilmu Komput.*, vol. 9, no. 1, pp. 8–16, 2020.

[18] S. A. Gunarto, E. S. Honggara, and D. D. Purwanto, "Website Sistem Rekomendasi dengan Content Based Filtering pada Produk Perawatan Kulit," *J. Sist. dan Teknol. Inf.*, vol. 11, no. 3, p. 399, 2023, doi: 10.26418/justin.v11i3.59049.

[19] S. Scratch, "Step-by-Step Guide to Building Content-Based Filtering." [Online]. Available: https://www.stratascratch.com/blog/step-by-step-guide-to-building-content-based-filtering/

[20] D. A. Pratiwi and A. Qoiriah, "Sistem Rekomendasi Wedding Organizer Menggunakan Metode Content-Based Filtering Dengan Algoritma Random Forest Regression," *J. Informatics Comput. Sci.*, vol. 3, no. 03, pp. 231–239, 2021, doi: 10.26740/jinacs.v3n03.p231-239.

[21] T. I. Ramadhan, A. Supriatman, and T. R. Kurniawan, "Evaluasi dan Implementasi Indobert Question Answering (QA) pada Domain Spesifik Menggunakan Mean Reciprocal Rank," pp. 180–188, 2024, doi: 10.33364/algoritma/v.21-1.1542.

ISSN 2085-4552

Identifying Student Competency Patterns in Informatics and Computer Engineering Education at Universitas Sebelas Maret using K-Means Clustering for Academic Guidance

Ratih Friska Dwi Andini[1], Febri Liantoni[2], Aris Budianto[3]
[1,2,3]Informatics Engineering and Computer Education, Sebelas Maret University
[1]ratihfrisska@gmail.com, [2]febri.liantoni@gmail.com, [3]arisbudianto@staff.uns.ac.id

Accepted 25 June 2025
Approved 30 June 2025

***Abstract*— This study evaluates the effectiveness of the K-Means algorithm in clustering student competencies using course score data from students in the Informatics and Computer Engineering Education program at an Indonesian public university. The K-Means algorithm was applied to group students into five distinct competency clusters based on their academic grade patterns. The model's performance was measured using the Silhouette Score, which resulted in a value of 0.3489, indicating a moderate quality of cluster separation. The results suggest potential applications for a student recommendation system for choosing elective courses and as an evaluation tool for the study program. Key limitations of this approach include the algorithm's sensitivity to the initial placement of cluster centers and the dependency on selecting the appropriate number of clusters (k) for optimal results.**

***Index Terms*— Academic performance; Clustering; Data mining; K-Means; Student competence.**

I. INTRODUCTION

Students of the Informatics and Computer Engineering Education study program at Sebelas Maret University will be faced with choosing a course of interest or concentration in the sixth semester. The available areas of concentration include desktop programming, web programming, video game software development, network administrator, and multimedia. Each student has different abilities and competencies in certain fields. However, based on observations, some students do not fully understand their strengths or weaknesses in these areas. This often causes them difficulty in determining the choice of study field, interest, or career path that suits their potential.

At the tertiary level, knowledge of areas of interest, expertise or competence is important in the student learning process. However, low awareness of self-potential often hinders students in determining the right actions to achieve academic and career success. Based on Masturina's research, students who have known their career interests early on find it easier to determine the actions to be taken to develop expertise and skills that are relevant to the needs of the world of work [1].

The abundant academic data of students holds great potential to be explored to help the decision-making process. In this context, data mining technology and machine learning algorithms, such as K-Means, can be utilized to analyze data and identify student competency patterns. The K-Means algorithm is known to be able to group data into certain groups based on similar characteristics. The use of clustering algorithms such as K-Means in education has shown significant results. Previous studies Mohd Talib et al., and Tuyishimire et al., show that this algorithm can identify student performance patterns based on academic data, thus helping educational institutions better understand student needs and potential [2] [3]. This study aims to apply the K-Means algorithm in the process of grouping students based on their academic data patterns. This data-based approach is expected to provide insight to study programs in supporting more appropriate decision making, as well as helping students understand their own competencies to develop their potential optimally.

II. METHODOLOGY

In this study, we have determined the object of research, namely students of the Informatics and Computer Engineering Education study program at a public university in Indonesia. This study uses an exploratory quantitative approach to analyze student competency patterns with the K-Means algorithm.

A. Research Stages

This research was carried out in several stages:

1) Analysis: At the initial stage, research data needs were identified.

2) Design: The process flow for clustering with the K-Means algorithm was designed.
3) Development: A clustering model was created using the Python programming language by leveraging the Scikit-learn library.
4) Determining the Number of Clusters (k): This stage aimed to determine the optimal number of clusters (k) before the final model was implemented. This process involved two methods:
    a. Elbow Method: This method was used to find a potential number of clusters by plotting the explained variance as a function of the number of clusters, then identifying the "elbow" point where adding more clusters no longer yields significant returns
    b. Validation with Silhouette Score: The number of clusters suggested by the Elbow Method was then validated using the Silhouette Score. This metric provides a quantitative measure of clustering quality, where a higher score indicates better-defined clusters
5) Implementation: The results from the formed clustering model were then applied in a decision support system
6) Evaluation: The final stage is to evaluate the quality of the resulting clustering model. This evaluation uses the Silhouette Score metric to measure how well-defined the final clusters are.

B. Data Source

At this stage, student data of the PTIK Study Program at University X was collected, including course values representing various fields of competence, from semester 1 to semester 5, with a total of 72 students. The data contains attributes such as student identification numbers, student names, and grades for courses that include theoretical and practicum aspects, such as Basic Multimedia, Structured Programming, Web Programming, 2D Animation Engineering, and Computer Network Security.

C. Selection

The selection stage aims to choose the most relevant and unique data for analysis. The first step in this stage is to identify and remove records that appear repeatedly to ensure the uniqueness of each entry.

Additionally, selection involves considering the features or attributes of the data to be used. The selection of relevant features is crucial as it can enhance model performance, reduce computational complexity, and improve interpretability [4][5]. By focusing on appropriate attributes and removing irrelevant or redundant data, simpler and more accurate models can be constructed, which enhances learning accuracy and reduces computation time [6]. After this selection stage, the dataset, containing unique records and relevant features, is ready for the subsequent cleaning process.

D. Cleansing

A procedure was implemented to handle incomplete data. Upon inspection, any student record found with missing grades in several courses was subsequently removed. The deletion method was chosen because the students had either officially withdrawn or had not enrolled in the respective courses, making data imputation irrelevant as there was no underlying academic performance to estimate. This approach is consistent with statistical data preparation practices where the cause of missing data guides the handling strategy.

E. Transformation

The Interquartile Range (IQR) method is a statistical technique used to detect and handle outliers in data analysis. Outliers, which are data points that deviate significantly from others, can distort statistical estimates such as the mean and standard deviation, reducing the reliability of analysis results. In predictive modeling, outliers can affect model stability and accuracy, making their detection and treatment crucial [7].

The IQR method identifies outliers by calculating the first quartile (Q1) and third quartile (Q3) and determining the IQR as the difference between Q3 and Q1. Data points falling below Q1 - 1.5IQR or above Q3 + 1.5IQR are classified as outliers. These outliers can either be removed or normalized, depending on the dataset size and analysis needs. Normalization adjusts extreme values to fit within an acceptable range, ensuring the dataset remains reliable without unnecessary data loss [8].

F. Data Processing

1) Data preparation

Data normalization is a crucial preprocessing step in clustering analysis, as it ensures that features contribute equally to the distance calculations used in clustering algorithms. Normalization can significantly influence the results of clustering by affecting the detection of cluster centers and the overall clustering structure.

Normalization helps in achieving more accurate clustering results by ensuring that no single feature dominates due to its scale. This is particularly important in distance-based clustering methods, where features with larger scales can disproportionately affect the clustering outcome [14] [15].

2) Clustering with K-Means

Apply the K-Means algorithm to form a more homogeneous cluster and determine the number of clusters using the Elbow or Silhouette Score method. The Elbow Method

involves plotting the explained variance as a function of the number of clusters and identifying the "elbow" point where adding more clusters yields diminishing returns in variance explained. This method is straightforward but can sometimes be subjective due to its graphical nature [16] [17] [18] [19]. Some studies suggest that the Elbow Method can be automated to reduce subjectivity, enhancing its effectiveness across various datasets [13].\

Some studies suggest that the Elbow Method can be automated to reduce subjectivity, enhancing its effectiveness across various datasets [18]. The Silhouette Score measures how similar an object is to its own cluster compared to other clusters. A higher silhouette score indicates better-defined clusters [16] [20] [21].

This method is often used to validate the number of clusters suggested by the Elbow Method, providing a quantitative measure of cluster quality[20] [22].

3) Analysis of clustering results

After the clusters are formed using the K-Means algorithm, the next stage is to thoroughly analyze and interpret the characteristics of each cluster to understand the students' competency patterns. This analysis process will involve several steps:

a. Descriptive Statistical Analysis: For each cluster, a descriptive statistical analysis will be performed. This includes calculating the mean and standard deviation (SD) for the course grades that constitute the clustering features. This step aims to build a quantitative and measurable academic profile for each cluster, thereby avoiding general descriptions such as "stable value" or "highest performance" that lack a statistical basis.

b. Comparative Analysis Between Clusters: The statistical profiles of each cluster will then be systematically compared with one another. A comparison of mean scores between clusters will be used to objectively identify in which competency areas a cluster excels or is weaker compared to the others.

c. Interpretation and Relation to Concentration Areas: Next, these quantitative profiles will be interpreted and linked to the concentration areas available in the study program. This analysis aims to map each cluster profile to the most relevant area of expertise or career interest. For instance, a cluster with high mean scores in 'Web Programming' and 'Multimedia Basics' courses would be interpreted as having a strong inclination towards the web and multimedia development concentrations.

d. Data Visualization: To facilitate understanding and presentation of the findings, the analysis will be supported by data visualizations. Diagrams such as bar charts or radar charts will be used to clearly and visually depict the competency profile of each cluster

III. RESULT AND DISCUSSION

A. Student Data

Student data that will be analyzed by the researcher includes the attributes of course values from 1st semester to 5th semester.

Table 1 Example of a research dataset

NIM	Multimedia dasar	Pemrograman Terstruktur	Komunikasi Data dan Jaringan	Fotografi	Desain Web	Alghoritma dan struktur data	Administrasi jaringan komputer	Desain Grafis Percetakan
1	80,75	90,5	89,45	81,5	88,02	82	92	8
2	85,25	96,5	94,45	81,4	86,73	92	92,5	8
3	84,75	94	92,5	85,8	82,895	87	84	82,1
4	81	85,5	81,4	85,25	82,665	91	91,5	84,0
5	84	92,5	87,5	84	83,955	81,5	90,5	81,7
7	79,75	83	69,45	82,3	81,32	62,5	85,5	8
8	81,75	91	84,45	83,35	76,955	74	79	83,6
9	83,5	92	90	85,2	78,68	81,5	90	83,8
10	84,5	93,5	87,5	81,45	79,055	76	86	83,8
11	83,75	94	94,45	84,3	86,255	91,5	90,5	85,0
12	85	96,5	76,4	83,45	100	90,5	91,5	85,0
13	84,5	95	87,5	84,25	84,195	83,5	89	84,2
14	82,75	60,5	95	80,8	84,245	86,5	74	80,7
15	78,75	92	89,45	82,6	89,05	75,5	89	83,2
16	85	92	100	85,55	87,115	87,5	87	84,3
19	81,75	89	76,4	85,1	80,625	85,5	82,5	83,
20	82,25	96	86,95	82,05	84,745	92	92	84,0
22	85	93,5	88,9	80,95	81,345	81,5	90,5	82,3
23	83,5	96,5	82,5	83	86,23	90,5	85	81,9
24	81,25	93,5	88,9	83,75	87,645	96	88	85,0

B. Data Cleaning

In the data cleaning process, duplicate data and empty data have been cleaned. From the results of the check, there are 12 student data that do not have grades in several courses. This condition occurs because the student has resigned or did not take the course in question in that period.

C. Data Transformation

The process of cleaning data from the oulier using the IQR method found that there were 7 outliner data, then deleted.

```
# Menghitung IQR
Q1 = df[average_columns].quantile(0.25)
Q3 = df[average_columns].quantile(0.75)
IQR = Q3 - Q1

# Menentukan batas bawah dan atas untuk outlier
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

# Menghapus outlier menggunakan IQR
df_clean_iqr = df[~((df[average_columns] < lower_bound) | (df[average_columns] > upper_bound)).any
```

Figure 1 IQR Implementastion Code

D. K-Means Clustering

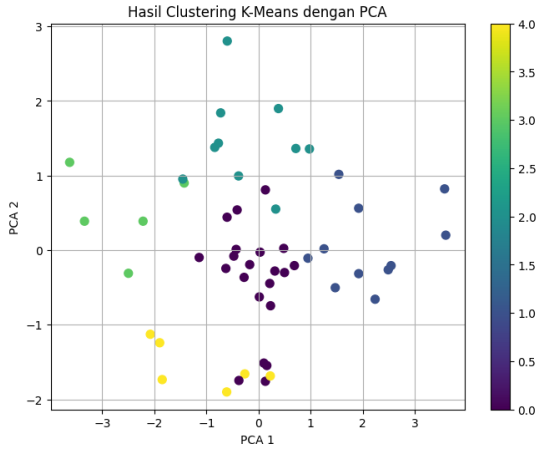


Figure 2 Visualization of Clustering Result with PCA

E. Cluster Interpretation

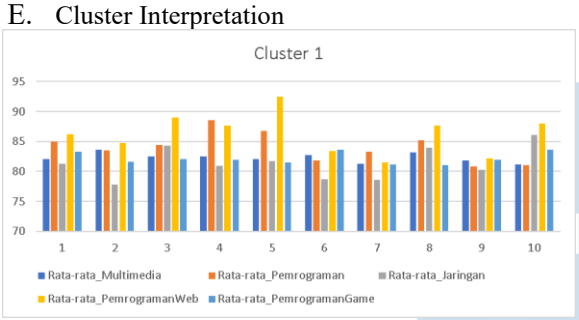


Figure 3 Clustered Column Chart cluster 1

Cluster 1 consists of students with balanced performance. This cluster consists of 10 students with fairly stable performance in all courses. There is no field that is very prominent or very low.

The calculated mean and standard deviation (SD) for each competency area within this cluster are as follows:

Table 2 Mean and Standard Deviation of cluster1

Competency Area	Mean	Standard Deviation (SD)
Web Programming	86.28	3.55
Programming	84.05	2.45
Multimedia	82.30	0.82
Game Programming	82.16	0.96
Networking	81.38	2.58

The qualitative description of "balanced performance" is well-supported, as all mean scores are high (above 81). The provided bar chart also visually confirms this, showing that for each of the 10 students, all five competency scores are consistently proficient.

Although no field shows notably low scores, the statistical analysis reveals that *Web Programming* and *Programming* have the highest average scores, indicating a tendency toward specialization in these areas rather than an evenly distributed skill set.

The concept of stable performance is further supported by the very low standard deviations in *Multimedia* (0.82) and *Game Programming* (0.96), suggesting that students within this cluster not only perform well but also exhibit a high degree of consistency in these specific domains.

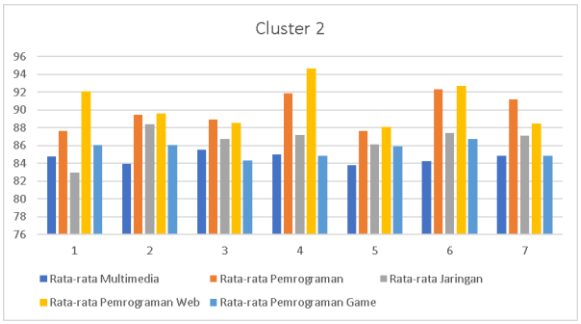


Figure 4 Clustered Column Chart cluster 2

Cluster 2 comprises students with consistently outstanding performance across all courses, particularly in *Programming* and *Web Programming*, where the average scores approach 90. This initial qualitative observation is substantiated by a detailed statistical analysis that quantifies their academic excellence.

Table 3 Mean and Standard Deviation of cluster 2

Competency Area	Mean	Standard Deviation (SD)
Web Programming	90.60	2.53
Programming	89.87	1.92
Networking	86.58	1.68
Game Programming	85.55	0.86
Multimedia	84.59	0.61

The calculated means and standard deviations for each competency area reveal a distinct pattern. *Web Programming* (mean = 90.60, SD = 2.53) and *Programming* (mean = 89.87, SD = 1.92) stand out, supporting the assertion that students in this cluster achieve near-perfect scores in these domains. The consistently high averages across the remaining areas—*Networking* (86.58), *Game Programming* (85.55), and *Multimedia* (84.59)—further confirm that these students excel across the board.

This interpretation is visually supported by the clustered column chart, where *Programming* and *Web Programming* are consistently represented as the highest-performing areas, often exceeding the 90-point threshold. The chart also highlights the elevated baseline across all competencies.

Moreover, the low standard deviations observed in *Multimedia* (0.61) and *Game Programming* (0.86) point to a high degree of consistency within the cluster, indicating that the outstanding performance is not limited to a few individuals but is uniformly distributed among all students in the group.

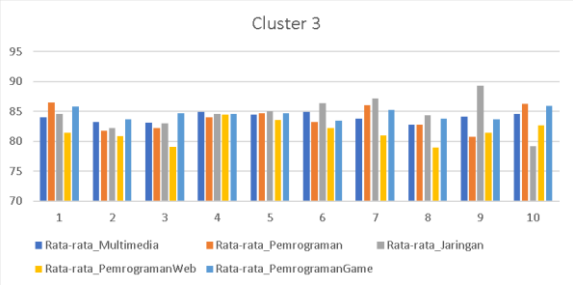


Figure 5 Clustered Column Chart cluster 3

The initial description characterizes Cluster 3 as consisting of 10 students with moderate and balanced performance. It is noted that their overall performance tends to be lower than other clusters and that their scores in the Web Programming course are slightly lower than in other subjects. A detailed statistical analysis provides a quantitative perspective on this profile

Based on the provided data for the 10 records in Cluster 3, the calculated mean and standard deviation (SD) are as follows:

Table 4 Mean and Standard Deviation of cluster 3

Competency Area	Mean	Standard Deviation (SD)
Networking	84.57	2.79
Game Programming	84.55	0.88
Multimedia	83.98	0.69
Programming	83.80	1.89
Web Programming	81.54	1.81

The statistical findings reinforce and elaborate upon the initial qualitative characterization of this cluster.

The average scores across all competency areas fall within a relatively narrow range in the low-to-mid 80s, supporting the classification of this group as exhibiting moderate performance compared to clusters with higher achievement levels. The data also substantiates the observation that *Web Programming* represents the weakest area within the group, with a mean score of 81.54—the lowest among all subjects.

While the performance profile appears balanced, with four out of five means ranging closely between 83.80 and 84.57, subtle variations are present. The highest levels of achievement are observed in *Networking* and *Game Programming*, suggesting these areas as relative strengths.

Notably, the group demonstrates exceptional consistency in *Multimedia* (SD = 0.69) and *Game Programming* (SD = 0.88), indicating a high degree of uniformity in student performance within these subjects. Conversely, the greatest variation is found in *Networking* (SD = 2.79), suggesting a wider disparity in skill levels in this domain.

These patterns are visually reflected in the clustered column chart, which illustrates the close grouping of most scores within the 80–85 range—highlighting the balanced nature of the group's performance. The comparatively lower scores in *Web Programming* are also clearly depicted in the chart, consistent with the statistical summary.

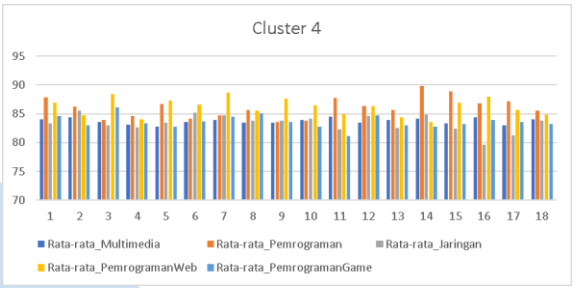


Figure 6 Clustered Column Chart cluster 4

Cluster 4 students with above average performance. This cluster consists of 18 students with slightly higher than average performance, especially in Programming courses. This cluster is the cluster with the largest number of studentsThe initial description characterizes Cluster 4 as the largest group, consisting of 18 students with "above average" performance. A particular strength in Programming courses is also highlighted. A detailed statistical analysis provides quantitative evidence to support and elaborate on this profile.

Based on the provided data for the 18 records in Cluster 4, the calculated mean and standard deviation (SD) are as follows:

Table 5 Mean and Standard Deviation of cluster4

Competency Area	Mean	Standard Deviation (SD)
Web Programming	86.16	1.57
Programming	86.03	1.84
Multimedia	83.71	0.52
Game Programming	83.50	1.06
Networking	83.44	1.38

The statistical analysis offers a deeper insight into the "above average" profile initially associated with this cluster.

The observed strengths are clearly supported by the data, with *Programming* (mean = 86.03) and *Web Programming* (mean = 86.16) emerging as the highest-

performing areas. These figures confirm the group's strong inclination toward development-oriented competencies.

An important characteristic of this relatively large cluster is its high level of internal consistency. The standard deviation for *Multimedia* is notably low at 0.52, indicating that performance in this subject is nearly identical among all 18 students. Similarly, the low variability in other subjects suggests a generally stable and uniform skill profile across the group.

This interpretation is further supported by the clustered column chart, which illustrates the consistency of student performance. Most bars fall within the 80–89 range, reinforcing the "above average" classification. In addition, the prominence of *Programming* and *Web Programming* scores—often among the highest for each student—visually aligns with the statistical summary.

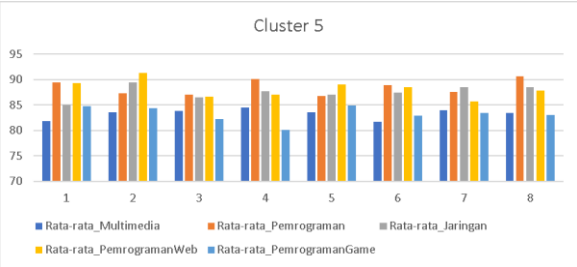


Figure 7 Clustered Column Chart cluster 5

The initial description characterizes Cluster 5 as a group of 8 students with strengths in specific areas. It highlights that this cluster excels in Programming and Networking courses but is slightly lower in Game Programming courses. A descriptive statistical analysis provides quantitative details that confirm and refine this specialized profile.

Based on the provided data for the 8 records in Cluster 5, the calculated mean and standard deviation (SD) are as follows:

Table 6 Mean and Standard Deviation of cluster 5

Competency Area	Mean	Standard Deviation (SD)
Programming	88.45	1.42
Web Programming	88.17	1.77
Networking	87.52	1.35
Multimedia	83.29	0.94
Game Programming	83.20	1.49

The statistical findings offer strong support for the initial qualitative assessment of this cluster.

The group's strengths are clearly reflected in the mean scores for *Programming* (88.45), *Networking* (87.52), and *Web Programming* (88.17), which are among the highest in the cluster. These results confirm the group's solid proficiency in development-related areas, as initially described.

The comparatively lower performance in *Game Programming*, with a mean score of 83.20, substantiates the earlier observation that this area represents a relative weakness within the cluster.

Overall, the standard deviations across subjects are relatively low, particularly in *Multimedia* (SD = 0.94), suggesting that the students in this cluster not only share similar strengths but also exhibit consistent patterns of performance, including their weaker areas.

This interpretation is further supported by the clustered column chart, which illustrates consistently high scores in *Programming*, *Networking*, and *Web Programming* for the majority of the 8 students. The slightly lower bars for *Game Programming* are also clearly visible, aligning with the quantitative analysis.

F. Clustering Evaluation

In this study, accuracy evaluation was conducted to assess the effectiveness of the K-Means and Naïve Bayes algorithms in grouping and classifying student competencies. The clustering results using K-Means were evaluated through the Silhouette Score, which produced a value of 0.34889702766857306, indicating that the clustering results were in the moderate category.

A moderate Silhouette Score likely reflects a combination of factors related to the data's nature and the methodology used. A primary reason could be the data's inherent structure, where the groups naturally overlap rather than being distinct and well-separated. Additionally, the chosen clustering algorithm, such as K-Means, may be ill-suited if the data's true clusters are non-spherical or have varying densities. Finally, the result may also be due to a suboptimal choice for the number of clusters (k), as this parameter heavily influences the clustering outcome.

```
# Menghitung Silhouette Score
labels=kmeans.labels_
silhouette_avg = silhouette_score(X, labels)
print(f"Silhouette Score: {silhouette_avg}")

Silhouette Score: 0.34889702766857306
```

Figure 8 Silhoutte Score

The silhouette score is used to evaluate the quality of data grouping in a cluster. If the silhouette score is close to +1, the data is considered to be in the correct or well-identified cluster. Conversely, if the score is close to 0, the data is likely to be on the border between two clusters, so there is potential for misplacement of the cluster. Meanwhile, a silhouette score close to -1 indicates that the data is in the wrong cluster, because it is closer to another cluster than the cluster where it is currently located [15].

The clustering results obtained show that the data has a fairly clear separation between one cluster and another, although there are still some data points that are close to the boundaries between clusters. This shows that although the clustering has been done well, there is still room to improve the quality of the separation between clusters to make it more optimal.

Given the moderate performance and the inherent limitations of the primary algorithm, it is prudent to discuss alternative methods that could be explored in future work. For instance, Hierarchical Clustering offers the advantage of not requiring the number of clusters to be pre-specified, instead building a tree-like structure of nested clusters (a dendrogram) that can be insightful for understanding relationships at different levels of granularity, which is particularly useful if the data contains meaningful sub-groups. Alternatively, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) can identify arbitrarily shaped clusters and is robust to outliers, which it automatically flags as noise. This makes DBSCAN a strong candidate if the underlying data patterns are not spherical or if the dataset is known to contain anomalous data points.

IV. CONCLUSIONS

This study successfully implemented the K-Means algorithm to identify five distinct student competency patterns within the Informatics and Computer Engineering Education program at Universitas Sebelas Maret. The analysis revealed diverse student profiles, including a group of 'Elite High-Achievers' excelling in all areas, a specialized cohort of 'Development and Networking Specialists', and a large, stable group of 'Consistent Moderates'. These findings demonstrate that K-Means is effective for uncovering meaningful, data-driven structures in student academic records without requiring initial labels. The practical implications of these findings are significant. The identified cluster profiles can form the foundation for a data-driven academic guidance system to help students make more informed decisions when choosing their area of concentration. Furthermore, these insights provide the study program with a valuable tool for curriculum evaluation and for better understanding the competency landscape of its student body.

However, the study acknowledges key limitations. The clustering quality, as measured by the Silhouette Score, was moderate at 0.3489. This result suggests that while the clusters are distinct, they are not perfectly separated, likely due to the overlapping nature of student skills and the inherent tendency of K-Means to prefer spherical clusters, which may not fully represent the complex distribution of academic competencies. The algorithm's effectiveness is also highly dependent on the selection of an optimal number of clusters (k) and its sensitivity to initial cluster centers.

Based on these limitations, future research should explore the application of alternative clustering algorithms. Methods such as Hierarchical Clustering or DBSCAN could provide different perspectives, potentially identifying non-spherical patterns or nested sub-groups within the data. Further work could also involve enriching the dataset with non-academic features to create more holistic student profiles and developing the proposed recommendation system into a fully functional tool for academic guidance.

REFERENCES

[1] D. Masturina, "Pengaruh Kompetensi diri terhadap perencanaan karir," *Psikoborneo*, vol. 6, no. 2, pp. 198–205, 2018, doi: https://doi.org/10.30872/psikoborneo.v6i2.4558.

[2] N. I. Mohd Talib, N. A. Abd Majid, and S. Sahran, "Identification of Student Behavioral Patterns in Higher Education Using K-Means Clustering and Support Vector Machine," *Applied Sciences (Switzerland)*, vol. 13, no. 5, Mar. 2023, doi: 10.3390/app13053267.

[3] E. Tuyishimire, W. Mabuto, P. Gatabazi, and S. Bayisingize, "Detecting Learning Patterns in Tertiary Education Using K-Means Clustering," *Information (Switzerland)*, vol. 13, no. 2, Feb. 2022, doi: 10.3390/info13020094.

[4] M. Rong, D. Gong, and X.-Z. Gao, "Feature Selection and Its Use in Big Data: Challenges, Methods, and Trends," *IEEE Access*, vol. 7, pp. 19709–19725, 2019, doi: 10.1109/ACCESS.2019.2894366.

[5] J. Li *et al.*, "Feature Selection," *ACM Computing Surveys (CSUR)*, vol. 50, pp. 1–45, 2016, doi: 10.1145/3136625.

[6] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, 2018, doi: 10.1016/j.neucom.2017.11.077.

[7] T. Nyitrai and M. Virág, "The effects of handling outliers on the performance of bankruptcy prediction models," *Socioecon Plann Sci*, p., 2019, doi: 10.1016/J.SEPS.2018.08.004.

[8] H. Aguinis, R. Gottfredson, and H. Joo, "Best-Practice Recommendations for Defining, Identifying, and Handling Outliers," *Organ Res Methods*, vol. 16, pp. 270–301, 2013, doi: 10.1177/1094428112470848.

[9] M. Suárez-Álvarez, D. Pham, M. Prostov, and Y. Prostov, "Statistical approach to normalization of feature vectors and clustering of mixed datasets," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 468, pp. 2630–2651, 2012, doi: 10.1098/rspa.2011.0704.

[10] V. Starovoitov and Y. Golub, "Data normalization in machine learning," *Informatics*, p., 2021, doi: 10.37661/1816-0301-2021-18-3-83-96.

[11] H. Humaira and R. Rasyidah, "Determining The Appropiate Cluster Number Using Elbow Method for K-Means Algorithm," *Proceedings of the Proceedings of the 2nd Workshop on Multidisciplinary and Applications (WMA) 2018, 24-25 January 2018, Padang, Indonesia*, p., 2020, doi: 10.4108/eai.24-1-2018.2292388.

[12] D. Saputra, D. Saputra, and L. Oswari, "Effect of Distance Metrics in Determining K-Value in K-Means Clustering Using Elbow and Silhouette Method," *Proceedings of the Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019)*, p., 2020, doi: 10.2991/aisr.k.200424.051

[13] A. Onumanyi, D. N. Molokomme, S. Isaac, and A. Abu-Mahfouz, "AutoElbow: An Automatic Elbow Detection Method for Estimating the Number of Clusters in a Dataset," *Applied Sciences*, p., 2022, doi: 10.3390/app12157515.

[14] M. Syakur, B. Khotimah, E. Rochman, and B. D. Satoto, "Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster," *IOP Conf Ser Mater Sci Eng*, vol. 336, p., 2018, doi:

10.1088/1757-899X/336/1/012017

[15] K. Shahapure and C. Nicholas, "Cluster Quality Analysis Using Silhouette Score," *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 747–748, 2020, doi: 10.1109/DSAA49011.2020.00096.

[16] C. Yuan and H. Yang, "Research on K-Value Selection Method of K-Means Clustering Algorithm," *J (Basel)*, p., 2019, doi: 10.3390/J2020016.

[17] N. Sagala and A. Gunawan, "Discovering the Optimal Number of Crime Cluster Using Elbow, Silhouette, Gap Statistics, and NbClust Methods," *ComTech: Computer, Mathematics and Engineering Applications*, p., 2022, doi: 10.21512/comtech.v13i1.7270.

ISSN 2085-4552

Web-Based Online Learning Platform with RAD and laravel Methods

Mohammad Reza Fahlevi[1], Saeful Mu'minin[2]
[1,2] Department of Informatics, Nahdlatul Ulama Indonesia University, Jakarta, Indonesia
[1] rezafah@unusia.ac.id , [2] saeful2026027@gmail.com

Accepted 25 June 2025
Approved 01 July 2025

Abstract*— The rapid advancement of technology has significantly impacted the field of education, particularly in providing broader and more efficient access to learning. One of the main problems in conventional learning systems is the limited access to learning materials, along with rigid schedules and physical constraints. This research was conducted to address these issues by developing a web-based online learning system that offers greater flexibility and accessibility for students at Universitas Nahdlatul Ulama Indonesia (UNUSIA). The e-learning system was developed using the Laravel framework due to its stability and scalability. Key features include user authentication, class selection, access to learning materials, exams, payments, and certification. The system development followed the Rapid Application Development (RAD) methodology, which consists of requirement planning, system design using UML, iterative development with user feedback, and testing using black box and User Acceptance Testing (UAT) methods. The implementation resulted in a functioning system with main modules such as login, dashboard, class access, exams, and certificate pages. Based on the UAT results, the system achieved a user satisfaction score of 95.6%, indicating strong acceptance and effectiveness in enhancing students' learning experiences. This web-based learning system is expected to support a more interactive, adaptive, and effective learning process in today's digital era.

Index Terms*— E-learning; User Acceptance Testing; Laravel; RAD; learning.

I. INTRODUCTION

Technological developments have brought significant changes to various fields, one of which is in the field of education [1][2]. The use of technology in education can be an important factor in improving student learning outcomes. Technology helps students understand subject matter more easily and also improve their skills [3][4]. Therefore, a learning system is needed that allows easy and flexible access from any location and at any time. In addition, this system should allow the selection of classes that suit the needs of each user[5]. In developing and building an e-learning system, it is important to adopt an effective and efficient approach to ensure the system functions optimally and according to the needs of its users.

E-learning aims to create an electronic classroom that is equivalent to classes in formal education. The goal is to distribute learning materials in real-time, allowing students to immediately access the materials as they are uploaded [6]. Universitas Nahdlatul Ulama Indonesia (UNUSIA) can utilize e-course-based e-learning to provide materials that are more specific and focused on the skills that students want to improve. By using e-course, students can choose materials that suit their own needs. Many universities have implemented various programs to assist students in acquiring their desired skills, including UNUSIA. As a lecturer and student at UNUSIA, the researcher is very enthusiastic in utilizing this technology and makes this university a case study in research. UNUSIA is a university located in Jakarta and currently continues to improve the quality of its educational services through an e-campus-based e-learning platform.

However, many students still face difficulties in accessing learning materials due to the limitations of traditional learning systems, which are often inflexible and constrained by time and location. The lack of integration between technology and students' learning needs presents a major challenge in delivering effective and accessible education. Therefore, there is a need for an online learning system that can overcome these barriers and provide a more adaptive, flexible, and efficient learning experience.

The e-learning system developed using the Laravel framework is expected to make an important contribution to education by improving their learning experience and helping students acquire the skills they need. Many students experience limited access to materials due to the existing curriculum, so they cannot achieve the desired abilities. Therefore, technology is needed that can provide a better learning experience for all students. Laravel framework is one of the reliable solutions to achieve this goal.

The utilization of Laravel technology in the development of e-learning sites with RAD (Rapid Application Development) approach can accelerate the process of e-course system design. The RAD method is a model of the system development life cycle (SDLC) that emphasizes software development through shorter cycles [7]. This method is very helpful in shortening the

possibility of errors at the beginning in the creation and development of an e-course website system that uses the Laravel framework.

Previous research on the creation of e-learning systems revealed a number of shortcomings. The first research conducted by Pamungkas & Raharja [8] showed that the method used was less flexible and time-consuming because it prioritized the interface and the limited technology used. The second study by Prianto & Septanto [9] used the waterfall method , but the results were ineffective because the technology chosen was old technology that was no longer relevant to the latest developments. The third study conducted by Widyawati, et al [10] also showed weaknesses, where the technology used was a less recent version and the method applied was the waterfall method, which is considered less efficient because it is an outdated software development method.

This research aims to design, implement, and evaluate a web-based e-learning system by applying the Rapid Application Development (RAD) method and the Laravel 10 framework, in order to enhance learning flexibility and accessibility for students at Universitas Nahdlatul Ulama Indonesia (UNUSIA). The system includes various features such as user authentication (login and registration), transaction processing, class access and management, user account management, online exams, certificate generation, and support for multiple payment methods.

The development of this e-learning platform is essential to improve educational accessibility through an online environment. With e-learning, students can study flexibly and independently, overcoming geographic and time-related constraints commonly encountered in traditional education. Furthermore, e-learning enables the delivery of diverse classes and programs tailored to students' individual needs, thereby significantly enhancing the inclusiveness and effectiveness of higher education.

II. METHODOLOGY

The methodology used in the design of a *web-based online* learning *platform* with the RAD method and Laravel 10 using the system development method is applied using the RAD *(Rapid Application Development)* approach. Model of several *System Development Life Cycle* (SDLC) models. RAD is an information system development method that has a short development time, so it is considered suitable to be applied in the development of this system [11] The design of this method involves several stages of system development.

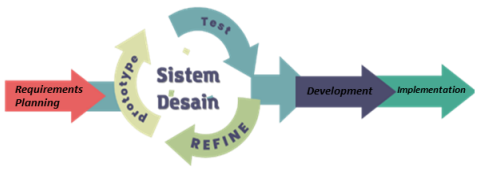


Fig. 1. RAD Development Cycle

A. *Requirements Planning*

This stage is the initial phase of system development, where problems are identified and data collection is conducted to define the system's goals[12]. In this study, data was collected through interviews and questionnaires involving 20 students and 2 academic staff members at Universitas Nahdlatul Ulama Indonesia (UNUSIA). The aim was to gather information regarding current obstacles in learning processes and understand the specific features users expect from an e-learning platform.

The results of this requirements analysis revealed several key needs, including easy login and registration processes, flexible course access, online payments, automated certificate generation, and a user-friendly interface. These identified needs became the foundation for the design and development of the system and were validated during subsequent phases of the RAD methodology.

B. *System Design*

This system design stage utilizes the Unified Modeling Language (UML) to create the proposed design and prototype. This design will later be implemented and used by the decision support team in the system. In this process, programmers and analysts work together to build and display visualizations of user designs and workflows [13]. Some of the diagrams used include case diagrams, activity diagrams, and class diagrams. The following is an explanation of each diagram designed for this system:

The use case diagram illustrates the three main actors in the system—Admin, Author, and User—and describes their interactions with the system. It presents the activities or functionalities that each actor can access, thereby helping to clarify user workflows and the system's primary features. The use case diagram for this system is shown in Figure 2.

Fig. 2. Use Case E-Learning

The activity diagram is used to model the workflow of activities within the system. It provides a visual representation of the process flow from start to finish, including decision points and branching conditions that may occur. This diagram helps in understanding the logical sequence of operations and the interactions that take place at each stage of the system. The activity diagram for this system is presented in Figure 3.

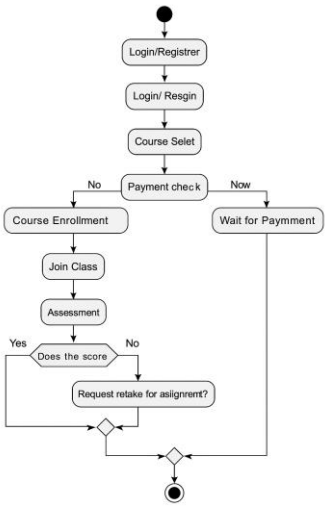


Fig. 3. Activity Diagram E-Learning

The class diagram represents the static structure of the system, including classes along with their respective attributes, methods, and interrelationships. This diagram facilitates a better understanding of the system's data structure and the interactions between its components, serving as a foundation for the development of business logic and data models during the implementation phase. The class diagram for this system is presented in Figure 4.

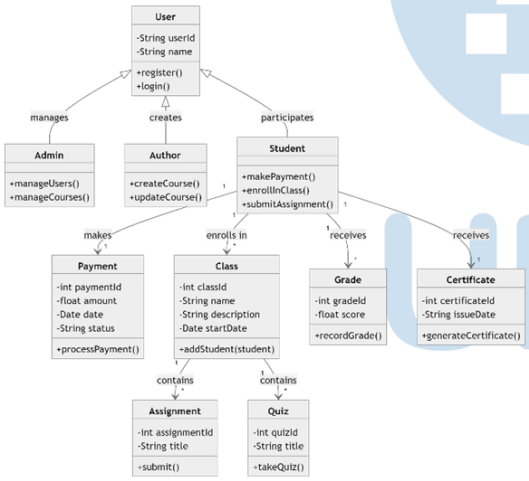


Fig. 4. Class Diagram E-Learning

Figure 4 illustrates the class diagram of the e-learning system, which shows the static structure of the system, including its main classes, attributes, methods, and relationships. At the top of the hierarchy is the User class, which contains general attributes such as userId and name, along with basic methods like register() and login(). This class is inherited by three subclasses: Admin, Author, and Student. These subclasses do not define their own attributes explicitly in the diagram, as they inherit the common attributes from the User superclass. Instead, each subclass introduces specific methods relevant to their respective roles.

The Admin class includes methods to manage users and courses, the Author class contains methods to create and update courses, and the Student class is responsible for handling payments, enrolling in classes, and submitting assignments. The diagram also displays associations between other entities such as Class, Payment, Assignment, Quiz, Grade, and Certificate, each with relevant attributes and functions. Relationships like makes, enrolls in, receives, and contains help illustrate how the classes interact with one another throughout the learning process.

C. Development

The third stage in RAD is evaluating user feedback. This evaluation includes features, functions, visuals, and interfaces of the program being developed by the development team [14]. At this stage the programmer has done the coding for the system that has been designed.

D. Implementation

The last stage is system implementation, including system evaluation to ensure that the system can function properly and meet needs [15]. At this stage, various testing methods are applied, such as black box testing to verify functionality without examining internal structures, and User Acceptance Testing (UAT) to confirm the system's usability and alignment with user expectations. Additionally, this phase may involve deploying the system in a real-world environment, providing user training, and gathering feedback for any necessary refinements.

III. RESULT

At this stage, researchers display the results of the system design in the form of a web page display that includes login, dashboard, main page, course, payment, exam, and certificate features. System testing is carried out using the black box method and User Acceptance Testing to ensure the functionality of each feature. The following is a more detailed explanation of each display.

A. Login Page

The login page is the user's initial access point to the system, where they enter their email and password. This page validates the credentials to ensure they match the stored data in the database. If the login is successful, the user is redirected to the main dashboard, where they can access system features based on their role and permissions. In cases where the credentials are invalid, the system displays an error message, prompting the user to re-enter their details or reset their password. Additional security measures can also be implemented to enhance protection against unauthorized access.

Fig. 5. Login Page Display

B. *Dashboard Page*

The dashboard is the page after login that is designed to provide a summary of important information and quick access to key features. The dashboard displays statistics, notifications, and navigation to various parts of the system such as courses, payments, and exams, making it easy for me to get an overview and quickly access the features I need.

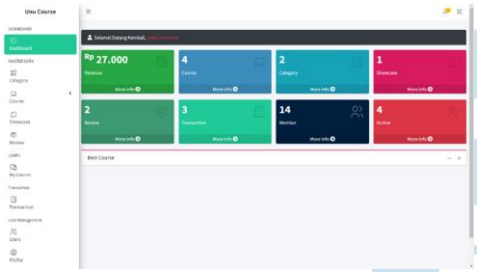


Fig. 6. Dashboard Page Display

C. *Home Page*

The *home* page is a *landing page* that I designed as a starting point for system introduction. This page presents the main information about the system, key features, and easy navigation to other important sections. The goal is to give a good first impression and make it easier for users to explore the various features available.

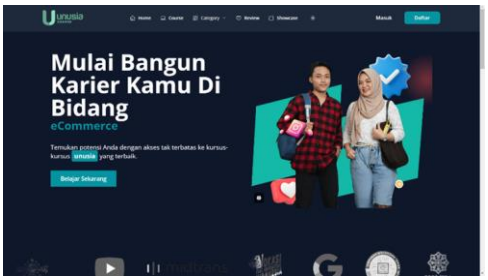


Fig. 7. Home Page Display

D. *Payment Page*

The payment page is the section where users can make transactions and payments for the courses they choose. This page uses Midtrans as a payment tool to facilitate secure and easy transactions. Users can select available payment methods, enter payment details, and complete the transaction with an integrated and secured process.

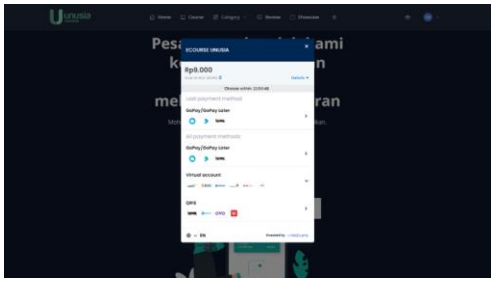


Fig. 8. Payment Page Display Using Midtrans

E. *Learning page*

The learning page allows users to access the course materials and modules they have purchased. Here, I provide navigation to various sessions or course topics, complete with learning materials and interactive tools. This page is designed to make it easier for users to follow and complete the course efficiently.

Fig. 9. Course Learning Page Display

F. *Exam Page*

The exam page is for multiple-choice tests accessible only after all class material is completed. It ensures participants finish all sessions before attempting the exam, reinforcing mastery of the material.

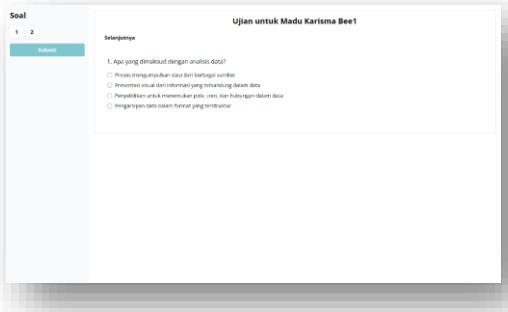


Fig. 10. Exam Page Display

G. *Certificate Page*

The certificate page is where users can view and download certificates after completing the exam with a score of 60 or above. After passing the exam, a

certificate will be generated and displayed on this page as proof of user achievement. This page is designed to make it easy for users to quickly access their certificate and ensure that they can save or print the certificate as official documentation.

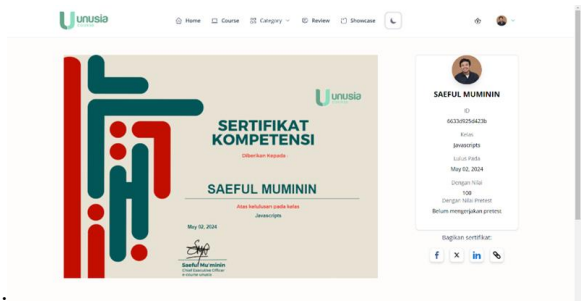


Fig. 11. Certificate Page Display After Exam Completion

H. *Black box testing and user* Acceptance Testing

This software testing focuses on evaluating the functionality of the system by checking whether the system features work as expected. This process is performed based on various predefined input and output scenarios. This approach is known as black box testing, where the test only looks at the result without examining or accessing the source code.

The main goal of this testing is to ensure that every function in the system works according to predefined specifications. It also seeks to identify any bugs or discrepancies that may occur in the system's performance. The validation includes the input process, workflow, and results provided by the system, to ensure compliance with user needs and expectations.

The final stage of testing involves User Acceptance Testing (UAT), where end users test the system to ensure that the software meets their needs and expectations. UAT aims to ensure the system is ready to be used in real conditions, according to the user's desired specifications.

TABLE I Black Box Testing Website E-Learning

No	Testing	Test Case	Expected Results	Test Results	Conclusion
1.	Login Page	Enter email and password	Successfully logs into the system	Matches	Valid
2.	Dashboard View	Verify dashboard display	The dashboard is displayed with all elements	Matches	Valid
3.	Main Page View	Access the main page	The main page displays text and a list of classes	Matches	Valid
4.	Payment Page View	Display payment page	The payment page displays payment method	Matches	Valid
5.	Class Access Page	View purchased classes and gain access	The page shows purchased classes and confirms successful payment	Matches	Valid
6.	Class Exam Page	Display class exams	Shows exams input by admin or authors	Matches	Valid
7.	Certificate Page	Display certificates based on classes	Shows certificates corresponding to classes and user grades	Matches	Valid

Table 1 displays the results of black box testing conducted on the e-course system to verify whether each core feature functions according to its intended specifications. The testing involved simulating typical user actions and observing system responses without examining the internal code. Seven test cases were evaluated, covering the login process, dashboard display, main page access, payment processing, class access, exam functionality, and certificate generation. All features produced the expected results and matched the defined output scenarios, resulting in a conclusion of "Valid" for each case. These results indicate that the main components of the system function correctly and are ready for real-world use.

TABLE II UAT Question Items for Assessing E-Course Usability

No	Question Code	Statement
1	Q1	I find the e-course website easy to use.
2	Q2	I find the interaction system on the e-course website to be simple.
3	Q3	I find that many features on the e-course website are easy to remember.
4	Q4	I find that the e-course website requires a reasonable amount of time to complete tasks.

5	Q5	I find it easy to understand how to use the e-course website.
6	Q6	I find that errors frequently occur while using the e-course website.
7	Q7	I am satisfied with the e-course website's ability to meet my needs.
8	Q8	I find it easy to learn how to use the e-course website.
9	Q9	I find it difficult to use the e-course website.
10	Q10	I am confident that I can use the e-course website effectively.

Table 2 presents the list of statements used in the User Acceptance Testing (UAT) to evaluate the usability and user satisfaction of the e-course website. Each item is designed to measure different aspects of user experience, such as ease of use (Q1, Q5), interface simplicity (Q2), memorability of features (Q3), task efficiency (Q4), and system reliability (Q6). Additionally, the questionnaire assesses user satisfaction (Q7), learnability (Q8), perceived difficulty (Q9), and user confidence (Q10). The combination of positive and negatively worded statements ensures a more balanced evaluation, helping to identify both strengths and potential usability issues in the system from the end-user perspective.

TABLE III E-Learning UAT Calculation Results

No.	Answer Scores from UAT Results					Total Score	Average	| Percentage%
	SS	S	N	TS	STS			
Q1	15	5	0	0	0	95	4.75	95%
Q2	15	5	0	0	0	95	4.75	95%
Q3	16	3	1	0	0	95	4.75	95%
Q4	17	2	1	0	0	96	4.8	96%
Q5	16	4	0	0	0	96	4.8	96%
Q6	17	3	0	0	0	97	4.85	97%
Q7	18	1	1	0	0	97	4.85	97%
Q8	17	2	1	0	0	96	4.8	96%
Q9	16	2	2	0	0	94	4.7	94%
Q10	16	3	1	0	0	95	4.75	95%
Average = (Total Percentage / Total Questions)						95.6%		

The results of the User Acceptance Testing (UAT), as presented in Table 3, indicate a high level of user satisfaction with the e-course system. A total of 20 respondents were involved in the testing process, selected purposively from active students at Universitas Nahdlatul Ulama Indonesia (UNUSIA) who were directly involved in using the system during the trial phase. This number was determined based on time constraints, resource availability, and access to representative end users during development. The UAT consisted of ten statements evaluated on a 5-point Likert scale, ranging from Strongly Disagree (STS) to Strongly Agree (SS). The responses showed consistently high ratings, with most questions receiving scores of 95% or higher. The highest scores were observed in Q6 and Q7 (97%), indicating strong user agreement on system reliability and satisfaction. The lowest, Q9 (94%), still reflects a high acceptance level despite minor difficulties. The overall average satisfaction score reached 95.6%, confirming that the system is user-friendly, functional, and ready to be deployed in a real educational environment

IV. CONCLUSIONS

This study concludes that the implementation of a web-based online learning platform using the Rapid Application Development (RAD) method and the Laravel framework has successfully achieved the research objectives. The system has been implemented with core features such as user authentication, class access, payment, assessment, and certificate generation. Through black box testing and User Acceptance Testing (UAT) with 20 student respondents, the system demonstrated high functional reliability and user satisfaction, with a satisfaction rate of 95.6%. These results indicate that the system is capable of supporting a self-directed, flexible, and accessible learning experience for students at the Nahdlatul Ulama University of Indonesia (UNUSIA). Despite its effectiveness, the current version of the system has some limitations. The system lacks real-time communication features such as discussion forums or video conferencing, and user testing was limited to a relatively small sample. Future work may involve expanding the system with interactive features, broader user testing, integration with mobile platforms,

and the use of AI to provide adaptive learning pathways. These improvements will further enhance the platform's usability, scalability, and relevance in meeting evolving educational needs.

ACKNOWLEDGMENT

The author expresses heartfelt gratitude to Universitas Nahdlatul Ulama Indonesia and the supervising lecturers for their invaluable assistance and unwavering support throughout the preparation of this research article. Their guidance, encouragement, and constructive feedback have been instrumental in ensuring the successful completion of this work. The author also acknowledges the university's contributions in providing the resources and academic environment necessary for conducting the research. Without their support, this study would not have been possible.

REFERENCES

[1] A. Fricticarani, A. Hayati, R. R, I. Hoirunisa, and G. M. Rosdalina, "Strategi Pendidikan Untuk Sukses Di Era Teknologi 5.0," *J. Inov. Pendidik. dan Teknol. Inf.*, vol. 4, no. 1, pp. 56–68, 2023, doi: 10.52060/pti.v4i1.1173.

[2] A. A. Rochim, "Kecerdasan Buatan: Resiko, Tantangan Dan Penggunaan Bijak Pada Dunia Pendidikan," *Antroposen J. Soc. Stud. Hum.*, vol. 3, no. 1, pp. 13–25, 2024, doi: 10.33830/antroposen.v3i1.6780.

[3] S. Suyuti, P. M. Ekasari Wahyuningrum, M. A. Jamil, M. L. Nawawi, D. Aditia, and N. G. Ayu Lia Rusmayani, "Analisis Efektivitas Penggunaan Teknologi dalam Pendidikan Terhadap Peningkatan Hasil Belajar," *J. Educ.*, vol. 6, no. 1, pp. 1–11, 2023, doi: 10.31004/joe.v6i1.2908.

[4] G. Pendidikan and A. Islam, "IMPLEMENTASI KEPEMIMPINAN GURU PAI DALAM MENINGKATKAN MOTIVASI BELAJAR SISWA MELALUI PEMANFAATAN TEKNOLOGI DI ERA DIGITAL Muaddyl Akhyar Universitas Islam Negeri Sjech M . Djamil Djambek Bukittinggi Junaidi Universitas Islam Negeri Sjech M . Djamil Djambek Bukittinggi Supriadi Universitas Islam Negeri Sjech M . Djamil Djambek Bukittinggi Susanda Febriani Universitas Islam Negeri Sjech M . Djamil Djambek Bukittinggi Ramadhoni Aulia Gusli Universitas Islam Negeri Sjech M . Djamil Djambek Bukittinggi Abstrak," vol. 18, no. 6, pp. 4234–4248, 2024.

[5] A. R. Kartiko, A. Zulkarnain, P. Studi, and S. Informasi, "PENGEMBANGAN SISTEM INFORMASI LEARNING MANAGEMENT SYSTEM UNTUK PENINGKATAN PENGALAMAN," pp. 723–732, 2024.

[6] Y. Huda, E. Tasrif, F. M. Saragih, W. Mustakim, and N. Vebriani, "Meta Analisis Pengaruh Media E-Learning terhadap Hasil Belajar pada Pendidikan Kejuruan," *J. Educ.*, vol. 5, no. 2, pp. 2808–2820, 2023, doi: 10.31004/joe.v5i2.929.

[7] N. Hidayat and K. Hati, "Penerapan Metode Rapid Application Development (RAD) dalam Rancang Bangun Sistem Informasi Rapor Online (SIRALINE)," *J. Sist. Inf.*, vol. 10, no. 1, pp. 8–17, 2021, doi: 10.51998/jsi.v10i1.352.

[8] C. A. Pamungkas and P. A. Raharja, "Rancang Bangun Learning Management System Berbasis Code Igniter Menggunakan Metode Prototype," *JSiI (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 215–220, 2022, doi: 10.30656/jsii.v9i2.5276.

[9] Agus Prianto and H. Septanto, "Rancang Bangun Learning Management System, Studi Kasus: Smk Wipama Cikupa," *J. Inform. Dan Tekonologi Komput.*, vol. 3, no. 3, pp. 252–263, 2023, doi: 10.55606/jitek.v3i3.2368.

[10] W. Widyawati, A. Surahmat, E. Nasri, and S. Febriyanto, "Rancang Bangun Aplikasi Learning Management System Dengan Framework Codeigniter Untuk Meningkatkan Efektivitas Pembelajaran Pada Smk Darul Ishlah," *J. Sist. Inf. dan Inform.*, vol. 5, no. 1, pp. 68–77, 2022, doi: 10.47080/simika.v5i1.1702.

[11] A. P. Simanungkalit, N. A. Putri, and V. Tasril, "Rancang Bangun Sistem Informasi Approval Dismentling NTE Telkom Akses dengan Metode RAD (Rapid Application Development)," *Indones. J. Educ. Comput. Sci.*, vol. 1, no. 1, pp. 16–22, 2023, doi: 10.60076/indotech.v1i1.3.

[12] V. No, J. Hal, and V. S. Gunawan, "Implementasi Sistem Informasi Administrasi Pembayaran SPP Pada SDIT Darul Hikmah Metode Rapid Application Development (RAD)," vol. 6, no. 3, pp. 562–568, 2024.

[13] H. Fernandy, I. Ali, and M. P. Juwono, "Rancang Bangun Sistem Tracer study UNUSIA Berbasis Web Menggunakan Metode Rapid Application Development," *J. Ilmu Komput. dan Sist. Inf.*, vol. 6, pp. 171–179, 2023.

[14] M. R. Fahlevi, M. A. Rohidin, and I. P. D. A. S. Prabowo, "Rancang Bangun Sistem Informasi Pendaftaran Santri Baru Berbasis Web Menggunakan Metode Rapid Application Development (RAD)," *Device*, vol. 14, no. 1, pp. 30–37, 2024, doi: 10.32699/device.v14i1.6720.

[15] F. Esya Ardhana and B. Parga Zen, "Perancangan Sistem Informasi Sekolah Dasar Berbasis Website Menggunakan Metode Rapid Application Development," *J. Sist. Inf. Galuh*, vol. 2, no. 1, pp. 71–83, 2024, doi: 10.25157/jsig.v2i1.3729.

ISSN 2085-4552

The relationship between sleep hours and exam scores in MATH001

Bonifasius Martin Wibawa[1], Revan Revian[2], William Lo Channiko[3], Fransiscus Ati Halim[4]
[1,2,3] Faculty of Engineering and Informatics, Informatics Study Program, Universitas Multimedia Nusantara, Tangerang, Indonesia
[4] Faculty of Engineering and Informatics, Information System Program, Universitas Multimedia Nusantara, Tangerang, Indonesia
[1]bonifasius.martin@student.umn.ac.id, [2]revan.revian@student.umn.ac.id, [3]william.lo@student.umn.ac.id, [4]fransiscus.ati@lecturer.umn.ac.id

Accepted 25 June 2025
Approved 01 July 2025

***Abstract*— This study examines the relationship between academic performance and sleep quality among Undergraduate Informatics students from the 2022 class. Examining math001 scores and Pittsburgh Sleep Quality Index (PSQI) data, the analysis reveals a diverse range of performance in math001, with a notable concentration of students reporting good sleep quality. Covariance and correlation matrices suggest an inverse relationship between PSQI scores and exam performance, indicating that better sleep quality may be associated with higher exam scores. Shape measure analysis further emphasizes the prevalence of good sleep quality among students. However, residual tests unveil challenges such as heteroscedasticity and autocorrelation, cautioning against overinterpretation of the regression model. The GLS model reveals a significant negative correlation between PSQI scores and exam performance, offering valuable insights into the potential impact of sleep quality on academic outcomes. This study contributes to understanding the complex dynamics between sleep quality and academic achievement, acknowledging the need for nuanced interpretation and consideration of underlying statistical assumptions.**

Index Terms*— PSQI ; sleep quality; math001 ; exam score, academic performance.

I. INTRODUCTION

Quality sleep is crucial for human health, impacting learning, physical, and mental well-being. [1]. The recommended sleep duration for teenagers and adults is 7 to 9 hours.[2] Sleep benefits neural processing, influencing insight, motor skills, perception, and visualization.[2]. It also preserves emotional well-being and cognitive functions. [3]; [4].

Sleep quality is the individual's satisfaction with aspects like efficiency, latency, duration, and waking conditions. [5]. Influenced by lifestyle, environment, work, health, social life, economy, and stress [6]It's also affected by diet, physical activity, and genetics. [7].

Optimal sleep positively impacts students' neurocognitive abilities, enhancing memory and problem-solving [8]. Sufficient sleep correlates with innovative problem-solving, leading to better academic performance. [9]. Good sleep quality, indicated by PSQI > 5, is linked to positive TEOG exam results in medical students. [10]. Research reveals no significant difference in sleep quality between students with varying academic performance levels. [6]. Similarly, poor sleep behavior doesn't correlate with academic success. [11].

This study aims to prove the following hypothesis: Students with better sleep quality will have higher math001 exam scores than students with poor sleep quality. In other words, sleep quality is directly proportional to math001 exam scores.

Mayor hypothesis

H0: Students' sleep quality does not affect their math001 exam scores.
H1: Students' sleep quality affects their math001 exam scores.

Minor hypothesis 1

H0: Students' average PSQI score is less than 5.
H1: Students' average PSQI score is equal to or greater than 5.

Minor hypothesis 2

H0: Students' average aggregate math001 exam score is less than 55.
H1: Students' average aggregate math001 exam score is greater than or equal to 55.

II. METHODOLOGY

A. Sleep Quality

Sleep quality, measured by variables like sleep efficiency, latency, and wakefulness after sleep onset, is vital for overall health and well-being [5]. In the realm of neurophysiological states, sleep plays a

critical role in learning, memory, and cognitive processing. Studies on sleep deprivation highlight substantial impairments in cognitive processing, impacting the acquisition and integration of information in the brain [12]. Sleep deprivation not only leads to decreased performance in thinking and behavior but also poses health risks, increasing the likelihood of traffic accidents. Furthermore, it diminishes the positive effects of sleep on cognitive processes such as learning and creativity [13].

B. *Exam Scores as a Measure of Performance*

Exams are one of the most commonly used performance measurement methods in education. Exams objectively assess students' understanding, mastery of subject matter, and cognitive abilities [14]. In addition, exams can reflect the extent to which students have understood and mastered the material. This allows instructors and educational institutions to evaluate the effectiveness of teaching methods and curricula. [15].

C. *Relationship between Sleep Quality and Academic Performance*

Several studies emphasize the crucial role of sleep quality in shaping academic and work performance. Poor sleep quality is often linked to subpar academic performance, while good sleep quality, coupled with sufficient duration and consistency, tends to correlate with improved academic outcomes. [6];[16].

Academic performance, measured by grades obtained in lectures, is considered a reflection of sleep quality. Studies propose a direct proportionality between sleep quality and student grades. [16]; [17].

Research on college students consistently shows that better sleep quality, longer duration, and greater consistency contribute to enhanced academic performance. A study at MIT by Okano and colleagues (2019) revealed a positive correlation between improved sleep factors and better grades. Similarly, research conducted at NUS College in Singapore by [18] Demonstrated a significant association between poor sleep quality, insufficient sleep, and lower academic performance.

D. *Pittsburgh Sleep Quality Index*

The Pittsburgh Sleep Quality Index (PSQI) is a self-reported questionnaire that assesses sleep quality and disturbances over time. PSQI is designed to evaluate the overall sleep quality of the subject. Each of the questionnaire's 19 self-reported items is divided into one of seven categories: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of medications, and daytime dysfunction.

The possible total sum of PSQI is 21, with higher results meaning worse sleep quality and vice versa. In other words, better sleep quality is indicated by lower PSQI scores. The general guidelines for PSQI index scores are as follows:

0–4 = Good sleep quality
5–10 = Poor sleep quality

With values over 10 indicating worse sleep quality.

This research utilizes quantitative methods through questionnaires, which will be distributed to the Undergraduate Informatics students, class of 2022. Who are enrolled in a Math001 course during the even semester of 2023

E. *Research Participants*

The participants of this study are limited to the Undergraduate Informatics students class of 2022 who have taken Math001 in their second semester (even semester of 2023). Participation in this study will be conducted through publicly accessible questionnaires, which will be provided in the form of Google Forms and distributed accordingly. The UMN Informatics students' class of 2022 is chosen as the object of this study as they are within close reach of the researchers, thus making it easier to collect data. The population of this research is limited to the Undergraduate Informatics students, class of 2022, which consists of 300 students. In order to satisfy the minimum percentage of samples (30% of the population), researchers must obtain at least 90 responses from the questionnaire.

The data acquired from the questionnaires amounts to 100 respondents, with 94 out of the 100 being valid, processable responses. The number of valid data is sufficient to satisfy the 30% minimum sample requirement, as 94 results in 31,33% sample percentage.

F. *Research Procedure*

Before analyzing the collected data, researchers must perform residual tests using a normality test, a homoscedasticity test, and autocorrelation tests. The normality test uses the Shapiro-Wilk Test. [19]The homoscedasticity test is used for the Constant-error variance test using the Breusch-Pagan Test. [20] , and the autocorrelation test uses the Durbin-Watson Test [21]. All three residual tests will be done in R using the related functions. All residual test results need to show acceptance of H0

The following hypotheses are set for the residual test:

NormalityTest

- H0: Data is distributed normally.
- H1: Data is not distributed normally.

Homoscedasticity Test

- H0: Data is Homoscedastic.
- H1: Data is Heteroscedastic

Autocorrelation test

- H0: Sleep quality and math001 exam scores are not autocorrelated.
- H1: Sleep quality and math001 exam scores are autocorrelated.

The math001 exam scores acquired from the questionnaires are then aggregated into a single numeric index using the following equation:

math001 score index = (mid exam score) * 0.43 + (final exam score) * 0.57 (1)

This equation is by the scoring guidelines of UMN, modified to disregard students' assignment scores

G. Data Collection

The sampling technique researchers use in this research is Simple Random Sampling, which is a method of drawing samples from a population or universe in a particular way so that each member of the population has an equal chance of being selected or taken (Kerlinger, 2006). To implement this technique, the researcher created a survey in the form of a questionnaire using Google Forms, which consists of questions related to the hypothesis that the researchers have set, then distributed the questionnaire to Undergraduate Informatics student class of 2022 through online chat applications, social media, and manually by asking people to scan the QR code that will lead to a google form to fill out the same questionnaire.

H. Data Analysis and Equations

In this research, the mean is used to provide a central measure of UMN Informatics students' math001 exam scores, offering an average performance that reflects the typical score in the class of 2022. The median is employed as a less sensitive measure of central tendency, particularly useful in handling outliers or skewed data. Variance and standard deviation are analyzed to understand the spread and consistency of math001 exam scores, while quartiles help explore the distribution among different levels of academic achievement.

Skewness and kurtosis assess the asymmetry and shape of the exam score distribution, providing insights into systematic biases and concentration around the mean. Covariance and Pearson correlation examine the relationship between sleep quality and exam scores, indicating whether changes in one variable correspond to consistent changes in the other. The Shapiro-Wilk test assesses the normality of sleep quality and exam score distributions, ensuring the validity of subsequent analyses.

The Breusch-Pagan test investigates heteroscedasticity in exam scores based on varying levels of sleep quality, crucial for robust regression analysis. The Durbin-Watson test explores potential autocorrelation in residuals, maintaining the independence assumption of the regression model. T-tests evaluate the significance of differences in mean exam scores between groups with different sleep quality levels.

For the study's regression model, the Generalized Least Squares (GLS) model is chosen due to its suitability in handling heteroscedasticity and correlated errors, accommodating weighted observations, and considering potential correlations between sleep quality and other factors. This makes GLS well-suited for capturing the intricate dynamics influencing academic performance among Undergraduate Informatics students in the class of 2022

I. PSQI

The Pittsburgh Sleep Quality Index (PSQI) is the chosen tool for assessing the sleep quality of Undergraduate Informatics students from the class of 2022. Recognized for its reliability, the PSQI captures various dimensions of sleep patterns through components like subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, and daytime dysfunction, offering a comprehensive overview of sleep quality. Lower PSQI scores indicate better sleep quality. Before the primary data collection, a pre-test of the adapted PSQI questionnaire will be conducted with a small sample of informatics students. This pre-test aims to identify and address any potential clarity, wording, or comprehension issues. Feedback from the pre-test participants will inform refinements to the questionnaire, ensuring its effectiveness and reliability in capturing relevant data. Once finalized, the questionnaire will be uploaded to Google Forms for distribution to the target population

J. Research Limitations

This research may be limited by the self-reported nature of the sleep quality data. Additionally, this research may be limited by the fact that it only includes Undergraduate Informatics students from the class of

2022, who are enrolled in a Math 001 course during the even semester of 2023. This research may also be limited by the fact that the timeframe of this research is limited to 6 weeks, from October and November 2023

K. Research Tools

In this research, researchers used several applications to facilitate the process of collecting, processing, and analyzing sample data, including RStudio, Google Forms, and Microsoft Excel. A questionnaire is created using Google Form and distributed it to obtain data from the respondents, the

data is then downloaded in spreadsheet format so that it can be accessed using the Microsoft Excel software to make it easier for the researchers to process the data, and grouped the data to facilitate the data analysis process in the RStudio application.

Microsoft Excel is a software program that allows users to process and calculate numerical data. After the data was collected, processed, and processed, the researchers analyzed the data to obtain the desired statistical results using the RStudio application. RStudio is a software program used to statistically analyze and display data in the form desired by the user.

III. Result and Discussion

A. Central Tendency and Spread Analysis

The following images are the results of the syntax run in RStudio to calculate central tendency and spread measures for the exam scores.

```
> mean_nilai <- mean(nilai)
> mean_nilai
[1] 65.02333
> median_nilai <- median(nilai)
> median_nilai
[1] 66.705
> getmode <- function(v) {
+   uniqv <- unique(v)
+   uniqv[which.max(tabulate(match(v, uniqv)))]
+ }
> mode_nilai <- getmode(nilai)
> mode_nilai
[1] 62.29
> variance_nilai <- var(nilai)
> variance_nilai
[1] 290.9151
> sd_nilai <- sd(nilai)
> sd_nilai
[1] 17.05623
> quantile_25 <- quantile(nilai, 0.25)
> quantile_25
    25%
54.7725
> quantile_75 <- quantile(nilai, 0.75)
> quantile_75
   75%
78.635
> iqr_nilai <- IQR(nilai)
> iqr_nilai
[1] 23.8625
```

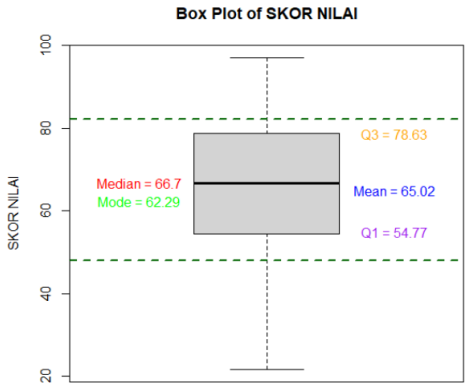


Fig. 1. Central Tendency, Spread, and Boxplot of Math01 Exam Score using R-studio

Undergraduate Informatics students from the class of 2022 had a mean math001 exam score of 65.02, with a slightly higher median of 66.705, indicating a potential negative skewness in the score distribution. The mode was 62.29, suggesting a significant number of students achieved this score. The variance and standard deviation were relatively high at 290.9151 and 17.05623, respectively, indicating notable variability in scores. Examining quartiles revealed a 25% quantile (Q1) of 54.7725, a 75% quantile (Q3) of 78.653, and an interquartile range (IQR) of 23.8625, highlighting a substantial range in students' performance within the middle 50% of the distribution.

```
> mean_psqi <- mean(psqi)
> mean_psqi
[1] 6.478723
> median_psqi <- median(psqi)
> median_psqi
[1] 6
> getmode <- function(v) {
+   uniqv <- unique(v)
+   uniqv[which.max(tabulate(match(v, uniqv)))]
+ }
> mode_psqi <- getmode(psqi)
> mode_psqi
[1] 5
> variance_psqi <- var(psqi)
> variance_psqi
[1] 9.004919
> sd_psqi <- sd(psqi)
> sd_psqi
[1] 3.00082
> quantile_25 <- quantile(psqi, 0.25)
> quantile_25
25%
  4
> quantile_75 <- quantile(psqi, 0.75)
> quantile_75
75%
  8
> iqr_psqi <- IQR(psqi)
> iqr_psqi
[1] 4
```

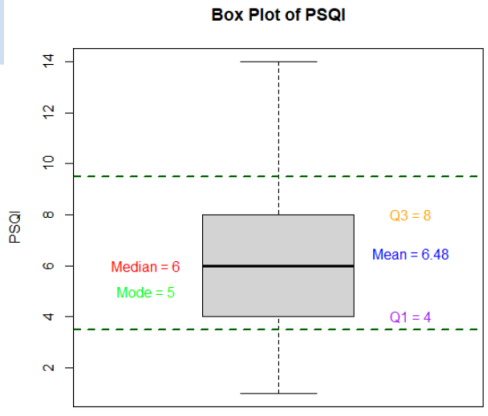


Fig. 2. Central Tendency, Spread, and Boxplot of Math01 PSQI Score using R-studio

The undergraduate Informatics Students class of 2022 had an average Pittsburgh Sleep Quality Index (PSQI) score of approximately 6.48, with a median of 6, suggesting a centered distribution of sleep quality.

The most frequently reported score was 5, indicating a concentration of students at this level. The variance and standard deviation were 9.004919 and 3.00082, respectively, indicating notable variability in sleep quality scores. Quartile analysis revealed a 25% quantile (Q1) of 4 and a 75% quantile (Q3) of 8, with an interquartile range (IQR) of 4, depicting the spread of sleep quality within the central 50% of the distribution.

B. Covariance and Correlation

The presented image displays the covariance matrix and Pearson's correlation coefficient matrix between PSQI scores and students' exam scores. The covariance matrix indicates a high variance in exam scores (approximately 290.9) and moderate variance in PSQI scores (approximately 9.00).

```
Covariance Matrix:
> print(covariance)
                 FINAL PSQI SCORE SKOR NILAI
FINAL PSQI SCORE         9.004919  -16.51296
SKOR NILAI             -16.512961  290.91507
> cat("\nPearson's Correlation Coefficient Matrix:\n")

Pearson's Correlation Coefficient Matrix:
> print(correlation)
                 FINAL PSQI SCORE SKOR NILAI
FINAL PSQI SCORE        1.0000000 -0.3226279
SKOR NILAI             -0.3226279  1.0000000
```

Fig. 3. Covariance and Correlation

The covariance between the two variables is approximately -16.51, suggesting an inverse relationship: lower PSQI scores tend to correspond with higher exam scores. The correlation coefficient matrix reveals a Pearson's correlation coefficient of approximately -0.323, indicating a weak negative linear correlation between sleep quality and exam scores. This implies that as sleep quality improves (lower PSQI), there is a tendency for exam scores to increase

C. Shape Measure Analysis

The following are the results of shape measure analysis, PSQI Histogram, and Overall Score Histogram.

```
> library(moments)
> skewness(data$`FINAL PSQI SCORE`)
[1] 0.7568865
> kurtosis(data$`FINAL PSQI SCORE`)
[1] 3.347485
> skewness(data$`SKOR NILAI`)
[1] -0.3684753
> kurtosis(data$`SKOR NILAI`)
[1] 2.524035
```

Fig. 4. Shape measure analysis using R programming

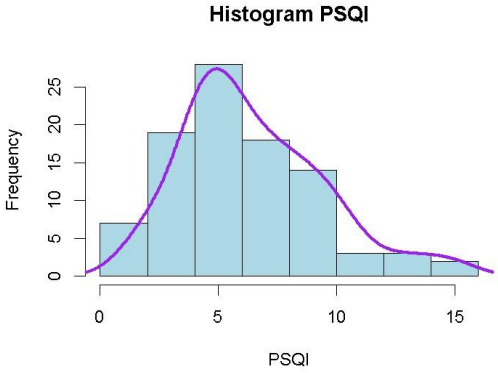


Fig. 5. Histogram PSQI of Math001 Exam Scores

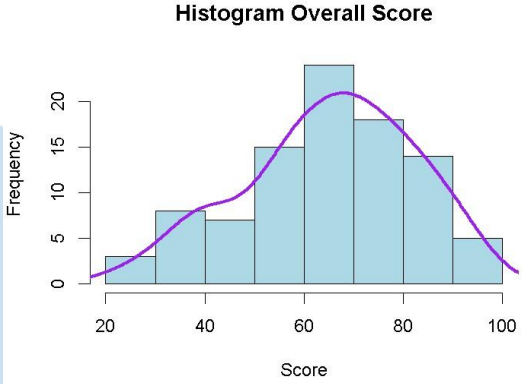


Fig. 6. Histogram of Math001 Exam Overall Score

The PSQI score data shows positive skewness (0.7568865) and high kurtosis (3.347485), indicating a left-leaning distribution and a concentration of scores below 5. This suggests that Undergraduate students generally have good sleep quality. The overall score data for math001 subjects exhibits negative skewness (-0.3684753) and platykurtic kurtosis (2.524035), indicating a right-leaning distribution with scores tending to be above 50. This implies that Undergraduate Informatics students from the class of 2022 have a diverse range of scores.

D. Residual Analysis

The following are the results of the residual test in the form of a normality test, Generalized Least Model (GLM), and Partial Regression Plot Analysis.

1) Normality Test

```
> data_nilai <- as.numeric(dataset$'SKOR NILAI')
> shapiro.test(data_nilai)

        Shapiro-Wilk normality test

data:  data_nilai
W = 0.97866, p-value = 0.1274
```

Fig. 7. Normality test of Math001 Exam Scores

The Shapiro-Wilk Normality Test in Figure 7 was conducted on Undergraduate Informatics Students' Math001 exam grades Class of 2022 with a resulting W-statistic of 0.97866 and a p-value of 0.1274. Given a significance level (alpha) of 0.05, the p-value exceeded this threshold. Consequently, the null hypothesis indicating a normal distribution of exam grades was accepted. This suggests that there is insufficient evidence to claim a significant departure from normality, and researchers can reasonably assume that the data is approximately normally distributed.

```
> data_psqi <- as.numeric(dataset$'FINAL PSQI SCORE')
> shapiro.test(data_psqi)

        Shapiro-Wilk normality test

data:  data_psqi
W = 0.94478, p-value = 0.0005975
```

Fig. 8. Normality Test of PSQI Scores

The Shapiro-Wilk test on PSQI scores for Undergraduate Informatics students (Class of 2022) yielded a W statistic of 0.94478 and a p-value of 0.0005975, indicating a significant departure from normality. However, considering the sample size, if it exceeds 30, the study may proceed with parametric statistical methods. The larger sample size mitigates the impact of minor deviations from normality, aligning with the Central Limit Theorem, and allows for reasonable use of parametric analyses despite the departure indicated by the test.

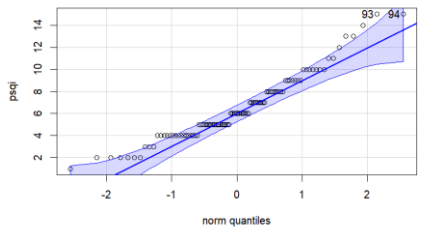


Fig. 9. Q-Q Plot of PSQI with Normal Area Highlighted in Blue

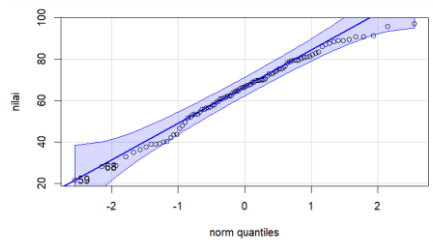


Fig. 10. Q-Q Plot of Math001 Exam Scores with Normal Area Highlighted in Blue

The plots above illustrate the distribution of PSQI and exam scores. The PSQI plot suggests some data points outside the normal zone, indicating a departure from normality. However, given a sample size exceeding 30, it is assumed the data approximates a normal distribution, allowing for subsequent tests. In contrast, the exam scores q-q plot shows all data within the normal zone, confirming a normal distribution.

2) Homoscedasticity Test

The Breusch-Pagan test is used to check for heteroscedasticity, which means the variability of the residuals is not constant across all levels of the independent variable(s).

Figure 11. shows the results of a Breusch-Pagan test done to test for homoscedasticity. The value of the BP is 9.611, with a degree of freedom of 1, and yielding a p-value of 0.001934.

```
> # Homoscedasticity Test (Breusch-Pagan)
> bptest(model)

        studentized Breusch-Pagan test

data:  model
BP = 9.611, df = 1, p-value = 0.001934
```

Fig. 11. Homoscedasticity Test for Constant-error variance

The small p-value suggests evidence against homoscedasticity in the residuals. Therefore, the null hypothesis of homoscedasticity is rejected, and the residuals of the data are indicated to be heteroscedastic

3) Autocorrelation Test

The Durbin-Watson test is used to check for autocorrelation in the residuals, which is a violation of the assumption of independence of residuals.

```
> # Autocorrelation Test (Durbin-Watson)
> dwtest(model)

        Durbin-Watson test

data:  model
DW = 1.2908, p-value = 0.0001202
alternative hypothesis: true autocorrelation is greater than 0
```

Fig. 12. Autocorrelation Test results

Figure 12 shows the results of a Durbin-Watson test done to test for autocorrelation in the residuals. The value of the DW is 1.2908, with a p-value of 0.0001202. The results indicate that the residuals sit against the null hypothesis of non-autocorrelation. Therefore, the residuals of the variables are autocorrelated.

E. T-test

A t-test is a statistical hypothesis test used to determine if there is a significant difference between the means of two groups. It's particularly useful when dealing with small sample sizes or when the population variances are unknown. A one-sample t-test is a statistical hypothesis test used to determine if the mean of a single sample is significantly different from a known or hypothesized population mean [22].

```
> #T-Test
> t_test_result <- t.test(psqi, mu = 5, alternative = "greater")
> print(t_test_result)

        One Sample t-test

data:  psqi
t = 4.7776, df = 93, p-value = 3.302e-06
alternative hypothesis: true mean is greater than 5
95 percent confidence interval:
 5.964501      Inf
sample estimates:
mean of x
 6.478723

> t_test_result <- t.test(nilai, mu = 55, alternative = "greater")
> print(t_test_result)

        One Sample t-test

data:  nilai
t = 5.6976, df = 93, p-value = 7.085e-08
alternative hypothesis: true mean is greater than 55
95 percent confidence interval:
 62.10056      Inf
sample estimates:
mean of x
 65.02333
```

Fig. 13. T-test Results

The t-value indicated in the t-table for alpha = 0.05 and df = 93 is 1.6614. This means that t-values greater than 1.6614 or less than -1.6614 suggest strong evidence against the null hypothesis, thus allowing for the null hypothesis to be rejected. In the t-test above, both PSQI and exam score t-tests result in t-values greater than 1.6614, with the t-value for the PSQI t-test sitting at 4.7776 and the t-value for the exam score t-test sitting at 5.6976.

Based on the t-test above, it can be concluded that the mean PSQI score data is 6.48, showing that UMN Informatics students enrolled in math001 courses on average have poor sleep quality. The mean of the math001 overall score is 65, showing that Undergraduate Informatics students enrolled in math001 courses on average have passed the math001 subject.

The t-test results for PSQI indicate that the null hypothesis, which states students' average PSQI score is less than 5, can be rejected, meaning that students' average PSQI is above 5, indicating poor sleep quality.

The t-test results for exam scores indicate that the null hypothesis, which states students' average aggregate math001 exam score is less than 55, can be rejected, meaning that on average, students' average aggregate math001 exam scores are greater than 55, thus, on average, students pass the math001 course.

F. Regression Model

A regression model is a statistical tool used to establish and analyze the relationship between a dependent variable (also known as the outcome, response, or target variable) and one or more independent variables (also called predictors or explanatory variables). Regression analysis is the statistical method used to determine the structure of a relationship between two variables (single linear regression) or three or more variables (multiple regression). The simple regression model for the data is as follows.

```
lm(formula = nilai ~ psqi, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-42.327 -10.874   0.268  10.673  30.310

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  76.9038     4.0010  19.221  < 2e-16 ***
psqi         -1.8338     0.5609  -3.269  0.00152 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16.23 on 92 degrees of freedom
Multiple R-squared:  0.1041,    Adjusted R-squared:  0.09435
F-statistic: 10.69 on 1 and 92 DF,  p-value: 0.001517
```

Fig. 14. Simple Regression Model Summary

Figure 14 shows the coefficient of determination for the data, which is at 10.41%, with the adjusted coefficient of determination being at 9,435%. This value indicates that 10% of the variance in the dependent variable is explained by the independent variable in the model. The low percentage also indicates that this regression model may not have a good fit for the data; therefore, other regression models may be used for the data.

G. The Generalized Least Squares (GLS) model

The Generalized Least Squares (GLS) model is an advanced regression technique used when the assumptions of Ordinary Least Squares (OLS) regarding the error terms are violated, specifically when errors exhibit heteroscedasticity (unequal variances across observations) or autocorrelation (errors are correlated with each other over time or space) [23].

```
> gls_model <- gls(nilai ~ psqi, data = data, correlation = corAR1())
>
> # Print model summary
> summary(gls_model)
Generalized least squares fit by REML
  Model: nilai ~ psqi
  Data: data
       AIC      BIC    logLik
  779.7439 789.831 -385.8719

Correlation Structure: AR(1)
 Formula: ~1
 Parameter estimate(s):
      Phi
0.3773654

Coefficients:
            Value Std.Error   t-value p-value
(Intercept) 78.08829  5.824803 13.406169  0.0000
psqi        -2.00378  0.809180 -2.476311  0.0151

 Correlation:
      (Intr)
psqi -0.903

Standardized residuals:
        Min          Q1         Med          Q3         Max
-2.57413862 -0.67133286 -0.01234902  0.62754174  1.86457389

Residual standard error: 16.44116
Degrees of freedom: 94 total; 92 residual
```

Fig. 15. GLS (Generalized Least Squares) Summary

The GLS model, considering heteroskedasticity and autocorrelation, investigated the relationship between math001 exam scores ("nilai") and PSQI scores. The AIC and BIC values (779.74 and 789.83) indicate a reasonable fit, with a specified AR (1) correlation structure. The autocorrelation parameter

(Phi) was approximately 0.377. The intercept (78.09) is significantly different from zero (t-value = 13.41, p-value = 0.0000). The coefficient for PSQI scores (-2.00) is statistically significant (t-value = -2.48, p-value = 0.0151), suggesting a negative relationship: as PSQI scores increase, exam scores tend to decrease. The linear regression function is y =78.08829 - 2.00378*psqi. The correlation between the intercept and PSQI score is -0.903. Standardized residuals range from -2.57 to 1.86, indicating how well the model explains variability. The residual standard error is 16.44, suggesting the typical magnitude of residuals around fitted values. With 92 residual degrees of freedom from 94 observations, the model implies a significant negative relationship between exam scores and PSQI, associating higher PSQI values with lower exam scores.

The following figure shows a partial regression scatterplot for a regression model, showing the trendline, individual data points, and a normal line with a normal zone.

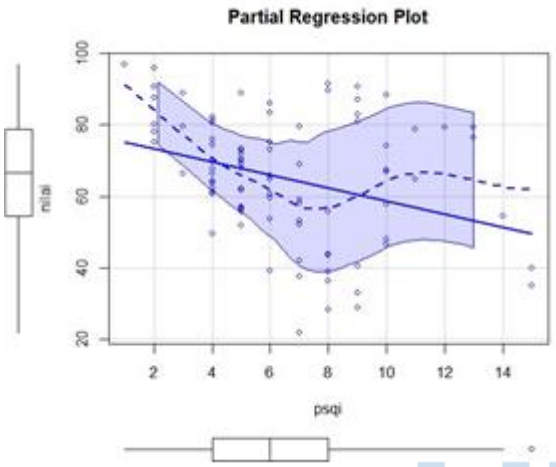


Fig. 16. Partial Regression Plot

Outliers, represented by points outside the normal zone, are visible. The trendline shows a negative linear relationship between "value" (test score) and "psqi" (PSQI score). The data points are evenly distributed, with concentrations near the trendline and within the normal zone, indicating variability in the values of "value" for different values of "psqi." This variability could be due to a variety of influencing factors. The concentration of data points near the trendline implies a systematic change in "value" with changes in "psqi," which is consistent with expectations in a normal distribution. The plot reflects a linear relationship between "psqi" and "value," indicating that changes in "psqi" correspond to changes in "value."

IV. CONCLUSIONS

The math001 exam scores exhibit a slightly negatively skewed distribution with a mean of 65.02, indicating average performance. PSQI scores have a mean of 6.48, reflecting average sleep quality. Covariance and correlation matrices reveal an inverse relationship between PSQI and exam scores, suggesting that better sleep quality correlates with higher academic performance. However, the weak correlation coefficient (-0.323) implies a modest linear relationship. Normality tests indicate PSQI scores deviate significantly, but given the sample size, parametric analyses may be reasonable. Regression analyses reveal a significant negative relationship between PSQI and exam scores, with outliers indicating variability. Sleep quality has a moderate effect on exam scores, rejecting the null hypothesis. However, it is not the sole determinant of academic performance, with other external factors playing a role.

The data analysis results show that the first null hypothesis of "Students' sleep quality does not affect their math001 exam scores" can be rejected. The second null hypothesis of "Students' average PSQI score is less than 5" can be rejected, indicating that the average sleep quality of students enrolled in Math001 is poor but not significantly. The third null hypothesis of "Students' average aggregate math001 exam score is less than 55" can be rejected, indicating that, on average, students pass the math001 course.

Moving forward, it is recommended to broaden the scope of research by incorporating additional quantitative variables beyond sleep quality. Factors such as study habits, stress levels, and prior academic achievement could provide a more comprehensive understanding of the determinants of math001 exam scores. Longitudinal studies tracking changes in sleep quality over time and exploring their correlation with academic performance could offer valuable insights. Moreover, considering external influences such as lifestyle and health in future quantitative analyses may contribute to a more holistic perspective on the complex interplay of factors influencing student success.

ACKNOWLEDGMENT

The author would like to thank all lecturers and students who have helped and participated in this research. The author also expresses his deepest gratitude to Universitas Multimedia Nusantara (Multimedia Nusantara University) for providing the opportunity and resources needed to conduct this research. The support and data received played a very important role in completing this research.

REFERENCES

[1] J. F. Dewald, A. M. Meijer, F. J. Oort, G. A. Kerkhof, and S. M. Bögels, "The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review," *Sleep Med Rev*, vol. 14, no. 3, pp. 179–189, Jun. 2010, doi: 10.1016/j.smrv.2009.10.004.

[2] M. Hirshkowitz *et al.*, "National Sleep Foundation's sleep time duration recommendations: methodology and results

summary," *Sleep Health*, vol. 1, no. 1, pp. 40–43, Mar. 2015, doi: 10.1016/j.sleh.2014.12.010.
[3] N. M. Punjabi, "The Epidemiology of Adult Obstructive Sleep Apnea," *Proc Am Thorac Soc*, vol. 5, no. 2, pp. 136–143, Feb. 2008, doi: 10.1513/pats 200709-155MG.
[4] J. S. Young, J. A. Bourgeois, D. M. Hilty, and K. A. Hardin, "Sleep in hospitalized medical patients, Part 1: Factors affecting sleep," *J Hosp Med*, vol. 3, no. 6, pp. 473–482, Nov. 2008, doi: 10.1002/jhm 372.
[5] K. L. Nelson, J. E. Davis, and C. F. Corbett, "Sleep quality: An evolutionary concept analysis," *Nurs Forum (Auckl)*, vol. 57, no. 1, pp. 144–151, Jan. 2022, doi: 10.1111/nuf 12659.
[6] R. Jalali, H. Khazaie, B. Khaledi Paveh, Z. Hayrani, and L. Menati, "The Effect of Sleep Quality on Students' Academic Achievement," *Adv Med Educ Pract*, vol. Volume 11, pp. 497–502, Jul. 2020, doi: 10.2147/AMEP.S261525.
[7] M. Sejbuk, I. Mirończuk-Chodakowska, and A. M. Witkowska, "Sleep Quality: A Narrative Review on Nutrition, Stimulants, and Physical Activity as Important Factors," *Nutrients*, vol. 14, no. 9, p. 1912, May 2022, doi: 10.3390/nu14091912.
[8] G. CURCIO, M. FERRARA, and L. DEGENNARO, "Sleep loss, learning capacity and academic performance," *Sleep Med Rev*, vol. 10, no. 5, pp. 323–337, Oct. 2006, doi: 10.1016/j.smrv.2005.11.001.
[9] D. Perez-Chada and C. Drake, "121 Sleep disordered breathing (SDB), sleep habits, and sleepiness are associated with poor academic performance in teenagers. A study using the pediatric daytime sleepiness scale (PDSS)," *Sleep Med*, vol. 7, pp. S8–S9, Sep. 2006, doi: 10.1016/j.sleep.2006.07.025.
[10] J. Duarte, P. Nelas, C. Chaves, M. Ferreira, E. Coutinho, and M. Cunha, "Sleep-wake patterns and their influence on school performance in Portuguese adolescents," *Aten Primaria*, vol. 46, pp. 160–164, Nov. 2014, doi: 10.1016/S0212-6567(14)70085-X.
[11] P. P. Lai and Y. H. Say, "Associated Factors of Sleep Quality and Behavior among Students of Two Tertiary Institutions in Northern Malaysia," *Med J Malaysia*, vol. 68, no. 3, pp. 195–203, Jun. 2013.
[12] M. C. Deak and R. Stickgold, "Sleep and cognition," *WIREs Cognitive Science*, vol. 1, no. 4, pp. 491–500, Jul. 2010, doi: 10.1002/wcs 52.
[13] J. M. Ellenbogen, "Cognitive benefits of sleep and their loss due to sleep deprivation," *Neurology*, vol. 64, no. 7, Apr. 2005, doi: 10.1212/01.wnl.0000164850.68115.81.
[14] K. Schwab, B. Moseley, and D. Dustin, "Grading Grades as a Measure of Student Learning," *SCHOLE: A Journal of Leisure Studies and Recreation Education*, vol. 33, no. 2, pp. 87–95, Jul. 2018, doi: 10.1080/1937156X.2018.1513276.
[15] R. Maulana, A. Kington, and J. Ko, "Editorial: Effective teaching: Measurements, antecedents, correlates, characteristics, and links with outcomes," *Front Educ (Lausanne)*, vol. 8, Mar. 2023, doi: 10.3389/feduc 2023.1170854.
[16] K. Okano, J. R. Kaczmarzyk, N. Dave, J. D. E. Gabrieli, and J. C. Grossman, "Sleep quality, duration, and consistency are associated with better academic performance in college students," *NPJ Sci Learn*, vol. 4, no. 1, p. 16, Oct. 2019, doi: 10.1038/s41539-019-0055-z.
[17] S. P. Gilbert and C. C. Weaver, "Sleep Quality and Academic Performance in University Students: A Wake-Up Call for College Psychologists," *J College Stud Psychother*, vol. 24, no. 4, pp. 295–306, Sep. 2010, doi: 10.1080/87568225.2010.509245.
[18] M. A. Armand, F. Biassoni, and A. Corrias, "Sleep, Well-Being and Academic Performance: A Study in a Singapore Residential College," *Front Psychol*, vol. 12, May 2021, doi: 10.3389/fpsyg 2021.672238.
[19] A. Monter-Pozos and E. González-Estrada, "On testing the skew normal distribution by using the Shapiro–Wilk test," *J Comput Appl Math*, vol. 440, p. 115649, Apr. 2024, doi: 10.1016/j.cam.2023.115649.
[20] T. S. Breusch and A. R. Pagan, "A Simple Test for Heteroscedasticity and Random Coefficient Variation," *Econometrica*, vol. 47, no. 5, p. 1287, Sep. 1979, doi: 10.2307/1911963.
[21] W. K. S. Lee, "The Deception on the Critical Value Table of the Durbin-Watson Test," Oct. 2024. doi: doi.org/10.2139/ssrn.5005826.
[22] A. Brown, "The strange origins of the Student's t-test," *Physiology News*, no. Summer 2008, pp. 13–16, Jul. 2008, doi: 10.36866/pn 71.13.
[23] E. Antonian, G. W. Peters, and M. Chantler, "Kernel generalized least squares regression for network-structured data," *PLoS One*, vol. 20, no. 5, p. e0324087, May 2025, doi: 10.1371/journal.pone.0324087.

ISSN 2085-4552

Developing an Android-based Tour Guide Application

Imanuel Vicky Sanjaya[1], Restyandito[2], Maria Nila Anggia Rini[3], Andhika Galuh Prabawati[4]
[1,2,3]Department of Informatics, Universitas Kristen Duta Wacana, Yogyakarta, Indonesia
[4]Department of Information System, Universitas Kristen Duta Wacana, Yogyakarta, Indonesia
[1]imanuel.vicky@ti.ukdw.ac.id, [2]dito@ti.ukdw.ac.id, [3]nila@ti.ukdw.ac.id, [4]andhika.galuh@staff.ukdw.ac.id

Accepted 25 June 2025
Approved 01 July 2025

***Abstract*— Yogyakarta, renowned as a student city, is also one of Indonesia's top tourist destinations. Its distinctive local culture attracts many visitors, particularly for traditional shopping experiences. An application named Sipasar now offers information about the traditional markets in Yogyakarta. Additionally, the tourism information currently available from the government agency TIC (Tourist Information Center) is only provided in the form of brochures. This information is not very comprehensive as it is limited to the brochure and is also not very up to date because the brochures are printed only in certain years. There is a need to develop an application that can fulfill the tourism information needs of Yogyakarta from the Sipasar application. Besides containing market information, this application also serve as a tour guide capable of providing route and navigation guidance, as well as helping tourists obtain information about local tourist destinations. To address these requirements, an Android-based application called Sipasar Tour Guide was developed. The application underwent blackbox testing to verify the functionality of all features. Moreover, the application was evaluated by respondents using the MARS (Mobile Application Rating Scale) method. This method involved 25 respondents using the application first, then answering questions by giving ratings on a scale of 1 which is the worst to 5 which is the best score. The data analysis results showed that the designed Sipasar Tour Guide application successfully met the needs of providing tourism information to users which is proven by the score of the average from all the answers at 4,5 of 5 scales or 90%, as well as providing routes and navigation for trips. This was proven not only by the results of the blackbox testing but also by the respondents' ratings during the application testing using the MARS method. The highest ratings from the four dimensions (usability, functionality, aesthetics, and information quality) were in functionality and information quality, with percentages of 92.4% and 91.6%, respectively.**

Index Terms*— *Android Application; Local Tourist Destinations; Tour Guide; Route; Navigation.

I. INTRODUCTION

Yogyakarta is recognized as a major tourist hub in Indonesia. In 2009, there are about 1,2 million visitors. This number rose substantially in 2017 to approximately 4,7 million tourists, and by 2018, Yogyakarta saw around 5 million visitors [1]. One of the attractions that tourists often choose is traditional shopping. In traditional markets, tourists can find traditional foods, iconic items, and so on. Although the locations of these markets are already available on the SiPasar application, which can be accessed through mobile or web applications, Sipasar can give information about traditional market in Yogyakarta, start from traditional things, market's map, and many more. Furthermore, data from the Yogyakarta City Tourism Office indicates that numerous tourists struggle to access information about tourist attractions [2]. A significant number of these visitors seek help by going to government offices, like the Tourism Information Center (TIC) at the Yogyakarta City Tourism Office. However, the TIC can only help by providing brochures of available tourist spots. Thus, the recommendations for tourist spots are not very up to date, and the information presented is incomplete, as it is limited to the brochures [2].

Currently, there are applications that can be used to guide tourists during their travels. Applications like Google Maps are now frequently chosen to guide people to their destinations. However, Google Maps does not provide cultural information or unique tourist icons at those locations, so users must find that information on their own. To get information about the culture of a specific location, tourists can also use the TripAdvisor application, but that application cannot provide direct guidance to tourists to reach their intended destinations.

Based on these issues, this research will develop an application that can serve as a tour guide for tourists by adding new tourist data to the SiPasar application. This application will use a text-to-speech feature, so tourists can enjoy their journey while listening to information about the location. This is because Text to Speech is a system that converts text into audio signals [3]. Additionally, to reach specific locations, users can select multiple points at once and use real-time turn-by-turn navigation to reach their destinations.

To develop the application, design planning is required first. According to Rahmasari, design planning is a stage that must be undertaken to create several alternative solutions before the development and implementation of the application [4]. Therefore, a data collection stage must be conducted first to understand what kind of application design will be carried out.

In this application, after conducting benchmarking, it is envisioned that there will be a map that can navigate and thus serve as a tour guide. This service can be provided using Mapbox, a location platform company that offers tools for displaying and analyzing spatial data [5]. A tour guide functions as an interpreter, educator, community ambassador, and a mediator for interactions between local culture, residents, and tourists from both abroad and within the country [6]. Therefore, it is crucial for tour guides to convey information effectively. To meet this need, a text-to-speech feature has been integrated into the travel navigation system to provide audio information about tourist sites. This application is being developed using Kotlin, a programming language that runs on the Java Virtual Machine [7].

The objective of this study is to develop and evaluate the *Sipasar Tour Guide* application, which aims to provide travel routes for tourists and offer information about tourist attractions in Yogyakarta. This application is expected to help tourists familiarize themselves with Yogyakarta's culture and tourist destinations through the provided information. Additionally, it facilitates tourists in determining their travel destinations by offering recommendations based on their proximity to related traditional markets. With this application, tourists can have a more guided and informative experience while exploring Yogyakarta.

This research is conducted by considering the existing studies. From these studies, the beneficial aspects are adopted, and for the less beneficial aspects, alternative solutions are sought to find more relevant options. This way, the research can be carried out optimally to achieve the desired results.

Collecting data through interviews with official government agencies to obtain concrete information, as done by Mersing et al. which also developing an android application for tourism objects in Batam city, seems appropriate for this research [8]. This approach can be applied in this study because of the similarity in the need to build a system.

For the system design model, in the research conducted by Mersing et al. which also developing android application for tourism object in Batam and Adil et al. which also using Mapbox API to map tourism, they used the waterfall model and succeeded [8][9]. However, considering the nature of that model, which is not flexible, it may not be suitable for implementation in this study. This is because this research requires a lot of improvements and changes based on the data collection conducted.

In Adil et al.'s research to map tourism, they use Mapbox, they use its API with JavaScript programming language [9]. Mapbox would be suitable for implementation in this study because it is free to use, and it provides comprehensive features. However, using JavaScript based on the web platform might not be appropriate for this research because the features to be used are compatible with mobile programming.

In the research conducted by Ahmadi et al. that concerning in pickup-and-delivery problems using electric vehicles with Mapbox multistops and also Choudhary et al. that using Augmented Reality for the Turn-by-Turn Navigation System for mobile devices, they used Mapbox features such as multi-stop and turn-by-turn [10][11]. Multi-stop is used to determine multiple destination points, while turn-by-turn can be used to navigate users in real-time to reach their destinations. These features would certainly be very helpful for tourists if implemented in this study.

Modeling with the SDLC prototype has more flexibility, compared to waterfall, making it easier to make corrections if there are errors [12][13]. Considering the practicality of coding and the continuity of application development, Kotlin would be more suitable to implement than Java

II. METHODOLOGY

A. Research Object

The object used in this research is a prototype of a tourist guide application designed to help tourists decide on travel destinations and obtain information about those destinations. It was tested on local tourists to determine the application's success in meeting their needs.

B. Research Subject

This research focuses on usage by tourists, with 10 respondents for prototype testing and 25 local tourists for system application testing. All respondents are aged 17-30 years and come from outside the city of Yogyakarta, as the application is focused on walking tours typically undertaken by this age group. The 10 respondents in the prototype testing assessed the system's suitability to user needs.

The study focuses on respondents aged 17-30 years, as this demographic represents the most active users of mobile applications and digital tourism services. This age group is also considered more adaptive to technology and frequently uses mobile applications for travel planning.

Based on data from the Central Statistics Agency of the Special Region of Yogyakarta Province, the number of international tourists in Yogyakarta Province is 6,258 [14]. Compared to the 469,227 international tourists visiting Bali, as reported by the Central Statistics Agency of Bali Province, the number of tourists in Yogyakarta can certainly be considered small [15].

Thus, considering this comparison in numbers, the margin of error used is 20%. To achieve a confidence level of 80% or a margin of error of 20%, the number of respondents needed according to the Slovin's formula is as follows:

$$N = 6.258(\text{number of population or visitors})$$

$$e^2 = ((20\%)/100)^2 = 0.2^2$$

$$n = \frac{6.258}{1 + (6.258 \times 0{,}2^2)}$$

$$n = 24{,}9 = 25 \text{ people/respondents}$$

Thus, there are a total of 35 respondents: 10 people for prototype testing and 25 people for system application testing, as well as research using the MARS (Mobile Application Rating Scale) model.

The limited number of respondents in prototype testing (10 people) is due to the qualitative nature of this phase, which aims to gather detailed feedback before the full system application testing. Meanwhile, 25 respondents were used for system application testing as a representative sample within the targeted user demographic in Yogyakarta.

The number of 25 respondents for application testing was determined using Slovin's formula, which is commonly used to calculate an appropriate sample size from a given population while considering a margin of error.

C. Testing Procedure

For the 25 respondent, the testing procedure can be outlined as follows:

1. Respondents are informed about the purpose of the research. Respondents must also be prepared to use the application for testing using the MARS method. This aims to ensure that respondents feel involved in this research.
2. Respondents are required to fill out the distributed form after using the application. This is necessary for conducting calculations in research using the MARS method.

D. Research Block Diagram

This research uses the SDLC Prototype method. In this research the first step is literature review to know people's needs for tourism. In SDLC Prototype, the research develops the prototype using Figma and develops the application based on the prototype design using Android Studio. To make sure the application is running well, this research uses blackbox testing. Then after the application passes the blackbox testing, MARS testing can be done to know the user's evaluation. The stages of this research can be seen in Figure 1.

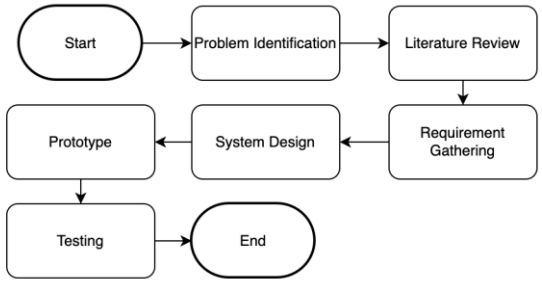


Fig. 1. Research Diagram

D.1. Problem Identification

The research begins by identifying the need for an application that provides travel routes and tourism information in Yogyakarta, particularly integrating traditional markets as part of the tourism experience.

D.2. Literature Review

Relevant studies on tourism applications, route recommendation systems, and usability testing methods are reviewed to establish a strong theoretical foundation.

D.3. Requirement Gathering

To understand user needs and expectations, data is collected through surveys. This stage helps define essential features such as market information, tourism categories, and route recommendations.

D.3.1. Collection of Tourist Place Data

During this stage, observations are made on tourist locations around the market to gather information about these tourist spots. This is useful to ensure that the tourist data collected is actual. Therefore, surveys and browsing related to this matter are necessary.

D.3.2. Use Case Diagram

The Sipasar Tour Guide application features a main tour guide for tourist spots around traditional markets, with tourism categories, tourist spot choices, and detailed information about each place. Users can view travel routes to selected destinations and use real-time navigation. These features are depicted in the Use Case Diagram in Figure 2.

Fig. 2. Use Case Diagram

D.4. System Design

Based on the findings from the literature review and requirement gathering, the application prototype is designed.

D.4.1. Application Flowchart

The details of this application flow can be seen in Figure 3.

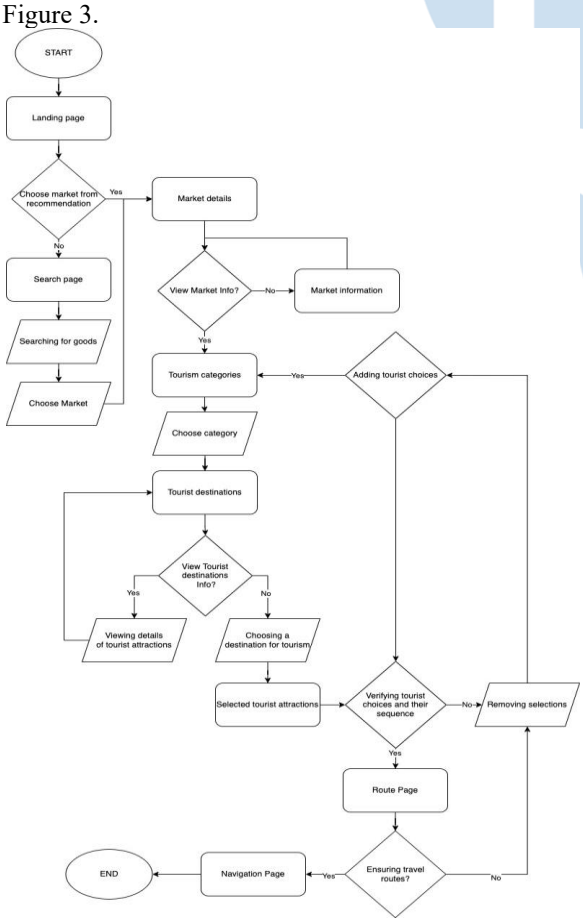


Fig. 3. Application Flowchart

D.4.2. System Requirements Analysis

An Android application named "Sipasar Tour Guide" was built to help users determine their travel destinations. The application aims to offer the following features:

a. Displaying a list of goods and markets that sell those goods

b. Displaying detailed market information

c. Displaying selectable categories of tourist destinations

d. Displaying options of various tourist destinations according to the selected category

e. Displaying travel routes in the form of a map for the selected tourist destinations

f. Displaying real-time travel navigation that provides travel instructions

D.4.3. Functional Requirements Analysis

The functional requirements analysis focuses on identifying and defining the key features that the application must provide to meet user needs effectively. This stage is based on data gathered during the requirement gathering phase, including user expectations and tourism-related use cases. The functional features identified in the analysis are as follows:

1. **Search** - On the search page, users can view and search for goods as well as market options selling those goods. Once a user selects a market, they will be directed to the detailed page of that market.
2. **Market Details** - The market detail page displays photos and a market layout, with an information tab containing descriptions, history, notes, item tags, and its stalls. The tourism tab showcases three relevant tourism categories associated with the market.
3. **Tourist Destination Options** - The tourist destination options page allows users to view 3 different tourist destinations within each selected category. They can see the distance from each destination to the targeted market, as well as detailed information about each place, including the number of reviews, photos, address, operating hours, and description.
4. **List of Selected Tourist Destinations** - On this page, users can view the list of tourist destinations they have selected. Users can also sort these tourist spots again.
5. **Travel Routes** - The travel routes page allows users to view routes to their selected tourist destinations, with estimated travel times between points. They can also remove tourist spots if the route does not align with their intended destination.

6. **Travel Navigation** - The navigation page provides real-time travel instructions at every road turn and displays information about tourist destinations upon arrival. Users can read or listen to this information using the text-to-speech feature.

D.4.4. Non-Functional Requirements Analysis

In addition to functional requirements, non-functional requirements play a crucial role in determining the overall performance, usability, and reliability of the "SiPasar Tour Guide" application. The following sections outline the necessary software and hardware requirements for the development and use of the application.

1. **Software Requirements Analysis**
   The applications and website that support the development of the "Sipasar Tour Guide" application in this research are Figma and Android Studio (Flamingo 2022.2.1).
2. **Hardware Requirements Analysis**
   The minimum hardware specifications required to use the "Sipasar Tour Guide" application are Android 8.0 (API level 26) and it must be an Android device.

D.5. Prototype

The prototyping phase is essential in the development of the "SiPasar Tour Guide" application, as it allows for iterative improvements based on user feedback. This stage involves designing and refining the application's interface and features before moving on to full-scale development.

1. **Concept *Design*** - This step begins with creating a design for the desired prototype concept in an abstract manner.
2. **Low Fidelity Prototype** - Once the concept and type of prototype are determined, in the initial stage of prototype development, the creation starts from a low-fidelity prototype of the application system. The prototype consists of landing page, search page, market detail page, tourism categories, tourist destination options page, places detail information, travel route, and travel navigation.
3. **High Fidelity Prototype** - In addition to creating designs for each feature and page, in the prototype stage, the existing designs are also converted into a High-Fidelity Prototype. This is necessary because the prototype will be tested by 10 respondents. Therefore, the prototype tested must accurately represent the functionality of the application system to be developed.

D.5.2. Database Design

Database Design in this research is carried out to determine the appropriate database structure for managing existing data. In the development of the Sipasar Tour Guide application, a database is needed to manage the selected tourist destinations. Since this function is relatively simple, the suitable database to use is Room Database.

Room Database is suitable for managing relatively simple and lightweight data, such as the selected tourist data in this application.

Within the system flowchart in Figure 3, this database is specifically utilized for storing tourism destination data. The primary structure used in this implementation is the Choices table, which holds information about selected tourist destinations. Figure 4 is the data structure employed in this research.

Choices	
PK	id (int)
	nama (string)
	foto (string)
	alamat (string)
	bintang (string)
	ulasan (int)
	jamOperasional (string)
	latitude (double)
	longitude (double)
	deskripsi (string)
	sumber (string)
	sumberGambar (string)

Fig. 4. Choices Table Structure in Room Database

D.5.3. User Interface Design

1. **Initial Design**
   In the initial design stage, after testing the prototype with 10 respondents, the implementation of these designs into the application system development begins. This involves building all desired features, starting from the landing page, search page, market details, tourism category options, tourist destination choices, travel routes, and travel navigation.

   However, upon reviewing the feedback and suggestions received, both from guidance sessions and from respondents, it appears that there are still shortcomings in the implemented application system. The deficiency lies in the search page, where all displays are presented in a list format, including item names that number in the tens. This layout is considered less effective and efficient for users. The design of the search page in the first iteration can be seen in Figure 5.

Fig. 5. Search Page in Initial Design

2. **Second Design**

After improving the application system, it turns out there is still an error in the search page layout. Although the displayed item names are different, the market options for each item remain the same. This is a fairly significant mistake. This error can be seen in Figure 6.

Because of this error, tourists can only see these markets, which may not necessarily sell the displayed items. Additionally, in this improvement, there is a distance between the markets and the users, so that users can more easily choose where to shop. However, because the list of markets is not correct, this is still not relevant.

Fig. 6. Search Page in Second Design Iteration

D.6. Testing

D.6.1. Testing Method

This application is tested using the blackbox testing method. The definition of blackbox testing itself is a method used to test a system without knowing its code structure [16]. After the application successfully passes blackbox testing, it is also assessed using the Mobile Application Rating Scale (MARS).

The MARS assessment is a reliable type of assessment that is simple, multidimensional in scale, and requires little training to implement. There are 4 aspects evaluated in MARS: engagement, functionality, aesthetics, and information quality. Each feature is rated on a scale from "inadequate" to "excellent" [17]. In the MARS (Mobile Application Rating Scale) method, calculations using the Cronbach's alpha formula are also conducted to determine whether the questions posed are the appropriate parameters for testing the application.

In the study by Roberts, et al. [17], there are six levels that serve as boundaries for determining the degree of internal consistency. These levels are very good, good, acceptable, questionable, poor, and the last is unacceptable. These boundaries are also used to determine whether the average rating of the application is very good or not. The calculation formula for MARS is as follows:

$$Average\ Score = \frac{Total\ Score}{Total\ of\ Respondents}$$

$$Total\ Score = 1st + 2nd + \cdots + N\ dimension\ average\ score$$

Thus, the application is evaluated not only for its functionality in being usable but also validated by users that it is informative and useful enough for them, addressing the issues intended.

D.6.2. Testing Phase

In the blackbox testing phase, independent testing is conducted on the developed application. This is used to test for bugs, errors, or incorrect information in the application. The stages involved in blackbox testing are as follows:

1. **Interface Testing**

Interface testing is conducted to determine whether the Sipasar Tour Guide application can display results that correspond to the input provided by the user. Describing the interface testing of the final implementation of the "Sipasar Tour Guide" application. On the landing page, several test cases were conducted, including market selection, start button, search input, item name button, and market selection. The test results showed that the system displayed the market details page and search results according to the given input. On the search page, the search input and item name button were also tested, and the results showed that the system could display the searched item name along with related markets. The market detail page was tested with several features such as market images, layout images, information tab, tourism tab, and back button, all of which showed results that met expectations.

On the tourism selection page, the system was tested to display the details of the selected tourist spots, save the selected spots, and remove

unchecked spots. The test results showed that the system could save and remove tourism selections correctly and display the number of selected spots on the number button. The selected tourism page was also tested with results showing that the system could remove data of unchecked spots.

Testing on the route page showed that the system could display the route map, remove tourism spots from the list, orient the map view to the North, and start route navigation. On the route navigation page, features such as re-center button, overview, volume, route instructions banner, and close button were tested, and the results showed that the system could follow and display the route map properly during the journey.

2. **Basic System Functionality Testing**

Basic system functionality testing is conducted to verify if the features available in the application can function correctly according to user needs, and to address any discrepancies in the displayed output. Describing the basic function testing of the final implementation, where the system was tested for item search, selecting tourist spots, deleting tourist spots, viewing travel routes, and using travel navigation. The test results showed that the system could display the searched items, save and delete selected tourist spots, display travel routes, and effectively guide users to the selected tourist spots.

a. Validation Testing

Validation testing is conducted to determine whether the designed validation system works successfully without any errors. Focusing on the validation of the number of selected tourist spots. The system was tested to provide a message when the user has selected 5 tourist spots and to disable checkboxes for further selections. The test results showed that the system successfully provided the message and disabled the checkboxes according to the specified limit.

b. Real Testing

Real testing is conducted to assess whether the application functions well for users. In this research, real testing is performed using the MARS method after ensuring that the application system runs smoothly without any errors. The real testing section explains that the final testing in the application research was conducted by having respondents use the application to fill out the MARS (Mobile Application Rating Scale) form. Some tasks given in this testing included searching for batik and selecting Beringharjo Market, choosing 5 tourist attractions including Empal Bu Warno, deleting 4 attractions and leaving Empal Bu Warno, and starting navigation and waiting for a message to appear.

III. Result and Discussion

A. Implementation

In this implementation phase, we start by developing the Sipasar Tour Guide application. Based on data collected from observations and prototype testing of the application, we have developed an application that includes not only market information system features but also tourist information around those markets. In this application, there are several pages that contain information about markets and tourist spots.

A.1. Database Implementation

Room Database is implemented in the application to manage the tourist data selected by users, through database initialization, table initialization, and DAO (Data Access Object) implementation. The integration of Room Database in the application can be seen in the code below.

Managing Android application development requires DAOs that contain methods to access the database such as query, insert, update, and delete. DAOs ensure that interactions with the database are performed through safe and consistent methods that can be called from other classes.

A.2. *Final Implementation*

In the final implementation, corrections and improvements were made based on the shortcomings of the previous implementation. The list of markets and the distance between markets and users are displayed correctly.

A.2.1. Landing Page

On the landing page, users are presented with instructions about the application's features and recommendations for items such as batik and snacks, based on the habits of local tourists who enjoy culinary tourism and shopping for traditional clothing. This allows them to bring souvenirs for their family and friends.

A.2.2. Search Page

On the search page, users can view up to 7 items and 7 related markets. They can choose from the options displayed or search for desired items themselves.

The markets are displayed based on the distance from the user to facilitate the walking tour. Recommendations for the nearest markets help users determine their destinations.

A.2.3. Market Detail Page

This market detail page contains all the information about the market, such as photos, history, address, stalls, and items sold. Users can zoom in on images, including market photos and maps. This page has 2 accessible fragments.

The first fragment contains detailed information about the market, while the second fragment displays categories of tourist attractions around the market. Each market in Yogyakarta has its own uniqueness based on history, stereotypes, and culture, resulting in different categories of tourist attractions. The market detail page is shown in Figure 7 and Figure 8.

A.2.4. All Selected Tourist Destinations Page

Users can view, delete, and see the details of the selected tourist attraction on this page. The distance between the market and the tourist attraction, as well as the total distance from the first to the last tourist attraction, is also displayed. This can be seen in Figure 9.

A.2.5. Travel Route Page

Users can view the travel route that includes the selected tourist attractions. These attractions can be removed if deemed unsuitable. The user's target market will be the final destination on the map, as shown in Figure 10.

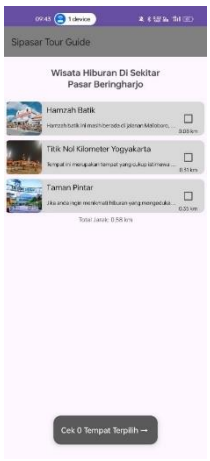


Fig. 7. Tourist Options

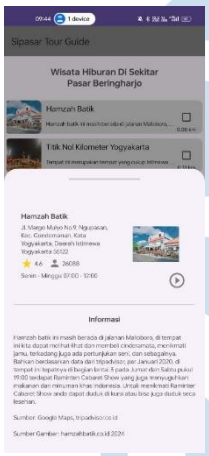


Fig. 8. Tour Details

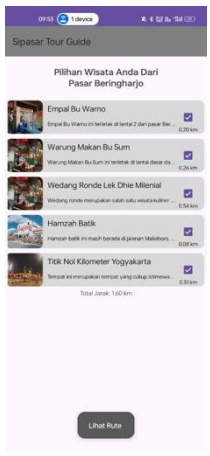


Fig. 9. Selected Tourist Options

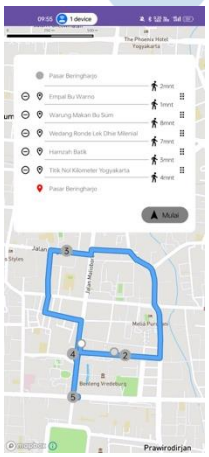


Fig. 10. Travel Route

A.2.6. Travel Navigation Page

After selecting a place, users can press the start button for real-time navigation with Mapbox's text-to-speech feature, providing turn-by-turn directions (Figure 11).

In Figure 12, when users reach their destination, information about the destination will be displayed and read aloud using a different text-to-speech feature from Mapbox. Users can stop this feature and read the information themselves in the bottom sheet dialog that appears at the destination.

B. Analysis

After finding an application that can meet the needs of local tourists, 25 respondents from outside Yogyakarta, aged between 17 and 30, tested the application. During the application testing, users were asked to fill out a form to evaluate the application using the Mobile Application Rating Scale (MARS) method.

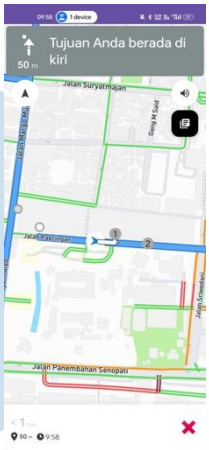


Fig. 11. Real-time navigation

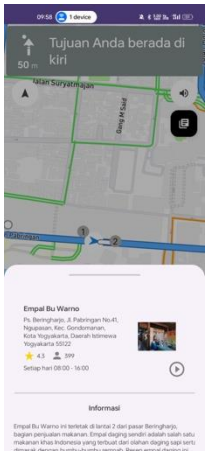


Fig. 12. Destination's Information

At this stage, research was conducted from four different dimensions. The first dimension is engagement, the second dimension is functionality, the third dimension is aesthetics, and the fourth dimension is information quality. This research was conducted by having respondents fill out a form containing questions for them to provide answers on a scale of 1 to 5.

The 1-5 rating scale was chosen based on the Likert scale methodology, which is widely used in user satisfaction and usability research. According to previous studies, a 5-point Likert scale effectively captures user feedback with balanced positive and negative response options.

From these four dimensions, the main focus of this research is on functionality and information quality dimensions. Because the aim of this research is to give travel tour which is the functionality and providing Yogyakarta tourist information which is related to the information quality. Here is the list of questions provided to the respondents:

1. Usability/Engagement: This dimension assesses the ease of users in recognizing the application system, to determine how familiar users are with the provided features. In the assessment of the Usability dimension, there are 3 questions.
   Q1. How intuitive is the navigation of this application for you?
   Q2. How easy is it for you to understand the messages or instructions given by the application?
   Q3. How responsive is the application to your inputs or interactions (e.g., deleting destinations, adding destinations)?
2. Functionality: This dimension assesses the utility of the application for users. Whether users feel that this application is useful for exploring the city of Yogyakarta or not. There are 6 questions provided in this dimension.
   Q1. How quickly does the application respond to your selected tourist destinations?
   Q2. How often do you experience lag or delays in accessing tourist information within the application?
   Q3. How stable is the application in performing its tourist functions without experiencing crashes or recurring errors?
   Q4. How smooth are the transitions between pages or menus within the application?
   Q5. How often would you use the main features of this application (especially the tourism features)?
   Q6. How easy is it for you to find and access the tourism features and information you need?
3. Aesthetics: The aesthetic dimension assesses the beauty of the application, focusing on the combination of colors, elements, and layout in the application's interface. There are 3 questions in the testing for this dimension.
   Q1. How appealing is the layout and visual design of this application to you?
   Q2. How consistent is the design style (colors, fonts, icons) throughout this application?
   Q3. How well does the application utilize design elements to enhance user experience (images, text-to-speech, infographics)?
4. Information Quality: The information quality dimension itself can be assessed based on whether the information provided is substantial, relevant, and helpful for users to understand the tourist attractions they intend to visit. There are 5 questions in the testing for this dimension.
   Q1. How clear and understandable is the information structure presented in the application?
   Q2. How useful are the information features and tourist options provided by this application?
   Q3. How clear is the information presented by this application?
   Q4. How well does the application provide solutions or answers to your tourism problems or needs?
   Q5. How well does the application handle your needs to provide tourist recommendations and suggest travel routes for those recommendations?

B.1. Dimension of Usability

Based on TABLE I, the Usability dimension obtained an average score of 4.53. This score, compared to the highest possible score of 5, is quite good. In percentage terms, this indicates the application's engagement level is at 90.6% , which can be considered "excellent" [17].

After testing the respondents and reviewing their responses, it seems they were pleasantly surprised and satisfied with the travel navigation instructions provided by the application. Despite the instructions being in Indonesian with an English accent, they were quite satisfied because the instructions were still easily understandable.

TABLE I. RESULT OF USABILITY DIMENSION

Questions	Respondents Rated (Likert 1-5)					Average Score (max 5)
	5	*4*	*3*	*2*	*1*	
Q1	14	11	0	0	0	**4.56**
Q2	14	10	1	0	0	**4.52**
Q3	15	8	2	0	0	**4.52**
					Average	**4.53 of 5**
					Average in persent	**90.6%**

B.2. Dimension of Functionality

Based on TABLE II, the average of all six questions in functionality dimension is 4.62. This score can be considered "excellent," as it translates to 92.4% in percentage terms [17]. This result is very satisfactory because it meets the research objective of providing travel routes for tourists and providing tourist information in Yogyakarta effectively through the application.

From the review of the application, questions 3 and 6 received the highest score of 5 mainly because the application has never experienced errors or crashes. Furthermore, tourist destinations can be easily found because their positions are clearly marked like the market's detailed information. Users can simply tap on a tourist destination to view its details.

TABLE II. RESULT OF FUNCTIONALITY DIMENSION

Questions	Respondents Rated (Likert 1-5)					Average Score (max 5)
	5	*4*	*3*	*2*	*1*	
Q1	15	9	1	0	0	**4.56**

Questions	Respondents Rated (Likert 1-5)					Average Score (max 5)
	5	*4*	*3*	*2*	*1*	
Q2	15	10	0	0	0	**4.60**
Q3	19	6	0	0	0	**4.76**
Q4	14	11	0	0	0	**4.56**
Q5	15	10	0	0	0	**4.60**
Q6	16	9	0	0	0	**4.64**
					Average	**4.62 of 5**
					Average in persent	**92.4%**

B.3. *Dimension of Aesthetics*

The results, as shown in TABLE III, are somewhat unsatisfactory in terms of aesthetics because when averaged, the first question only received 4.16, the second question received 4, and the last question received 4.6. Overall, in terms of average ratings across all questions, the aesthetic dimension only scored 4.25 out of the highest scale rating of 5.

It's undeniable that the development of this application has been more focused on functionality and the quality of information provided. Therefore, at this point, the visual design of the application still lacks satisfaction in terms of color combination and elements. However, with a percentage of 85% from the score of 4.25, it falls within the "good" rating category [17].

From the redesign review of this application, the first and second questions received low scores because the colors used are still too monotonous. Additionally, the layout of the elements presented in this application is also not very attractive. Although it has improved compared to the prototype, it is still not sufficient to be considered excellent.

TABLE III. RESULT OF AESTHETICS DIMENSION

Questions	Respondents Rated (Likert 1-5)					Average Score (max 5)
	5	*4*	*3*	*2*	*1*	
Q1	10	11	2	2	0	**4.16**
Q2	7	13	3	2	0	**4.00**
Q3	17	6	2	0	0	**4.60**
					Average	**4.25 of 5**
					Average in persent	**85%**

B.4. *Dimension Information Quality*

Based on TABLE IV, the overall average for the information quality dimension is 4.58. These results are quite satisfactory as they exceed 90%, specifically 91.6%. This percentage in the information quality dimension can be considered "excellent" [17].

From the user evaluations and application reviews regarding information quality, respondents indeed feel that the provided information is sufficient for them to determine their tourist destinations. Details such as addresses, operating hours, and other relevant information about the attractions have been very helpful. Additionally, the fact that this information is sourced from reliable sources adds to its credibility.

TABLE IV. RESULT OF AESTHETICS DIMENSION

Questions	Respondents Rated (Likert 1-5)					Average Score (max 5)
	5	*4*	*3*	*2*	*1*	
Q1	17	8	0	0	0	**4.64**
Q2	11	10	0	0	0	**4.52**
Q3	16	9	0	0	0	**4.60**
Q4	14	11	0	0	0	**4.56**
Q5	15	10	0	0	0	**4.60**
					Average	**4.58 of 5**
					Average in persent	**91.6%**

C. Discussion

In this section, we discuss the evaluation results of the developed application. The discussion includes two key aspects: Cronbach's Alpha Score, which assesses the reliability of the questionnaire used in the study, and a comparison with similar applications, highlighting the distinguishing features and advantages of our system. These aspects provide a deeper understanding of the application's effectiveness and usability.

C.1. *Cronbach's Alpha Score*

In addition to evaluating its usability, this study also conducted an analysis using Cronbach's alpha. This was done to determine whether all the questions posed were reliable tools for measuring the application under study. This analysis helps validate statements about whether the application is good or bad in its respective dimensions.

The formula for Cronbach's alpha is as follows:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma_{total}^2}\right)$$

α = Cronbach's alpha value

k = The number of items in the scale

σ_i^2 = The variance of each item

σ_{total}^2 = The variance of total score

The calculation using Cronbach's alpha for this application study resulted in an alpha value of 0.81 [17]. This result falls into the category of good, as seen within the established thresholds.

C.2. *Benchmarking with Similar Applications*

Aplikasi Sipasar Tour Guide juga dibandingkan dengan fitur-fitur pada aplikasi sejenis lainnya. The applications studied were A = "Google Maps," B = "Tripadvisor," and C = "Roadtrippers." The results of benchmarking these similar applications can be seen in Table V.

TABLE V. BENCHMARKING RESULTS

Feature	A	B	C	Sipasar Tour Guide
Login	Yes	Yes	Yes	No
Search	No	Yes	Yes	Yes
Tourist Recommendations	Yes	No	Yes	Yes
Travel Route	Yes	No	Yes	Yes
Selecting Multiple Destinations	Yes	No	No	Yes
Travel Navigation	Yes	No	Yes	Yes

Overall, across the four dimensions tested, this application received a rating of 4.5 out of 5 from 25 respondents. Based on the established scale, this data falls into the "very good" category, as it is at 90% [17]. The comparison graph for each dimension is shown in Figure 13.

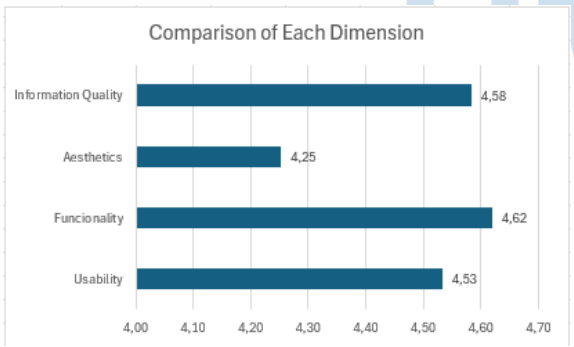


Fig. 13. Comparison of Each MARS Dimension

To determine whether the average rating of 4.5 is considered satisfactory compared to other applications, a comparison was made with the ratings of similar applications obtained from user reviews on the Play Store. From the research results of the average reviews of 3 similar applications, the following data was obtained, as shown in Figure 14:

- Google Maps
  - Average: 4,3
  - Number of Reviews: 17.228.001
- Tripadvisor (Mobile)
  - Average: 4,6
  - Number of Reviews: 1.378.167
- Roadtrippers – Trip Planner
  - Average: 4,3
  - Number of Reviews: 7.528

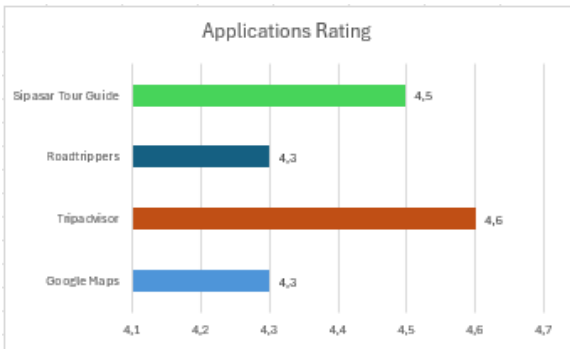


Fig. 14. Comparison of the 'Sipasar Tour Guide' Application with Similar Applications

Although the number of reviews for these compared applications is much higher, the 25 participants tested in this study are sufficient as research samples. This is because the study focuses more on the scope of tourism in the city of Yogyakarta alone. Meanwhile, these applications receive numerous reviews because they provide information not only for one city but for the entire world.

However, from the results obtained, it turns out that the application from this research is quite satisfactory compared to those applications. Although currently, this application receives a much lower average rating compared to Tripadvisor. This application is still very new and can be further developed to achieve much higher ratings. Development can be done across various dimensions.

IV. CONCLUSIONS

A. *Conclusion*

From the research and analysis carried out it can be concluded that the "Sipasar Tour Guide" application has succeeded in meeting the needs of local tourists. With the Sipasar Tour Guide, local tourists can effectively search for and obtain detailed information about tourist destinations in Yogyakarta.

This is proven by the average results of research analysis using the MARS method, where overall the four dimensions received a "very good" rating. What stands out most is the functionality dimension, which has the highest average rating in this application study, with a percentage of 92.2%. This shows that the application is very functional and usable for the users. Moreover, the dimension with the second highest ranking is the information quality dimension, with a score of 91.6%. This shows that the detailed tourist

information provided is very useful for users in determining their travel destination.

B. Recommendations

After drawing these conclusions, here are several recommendations that can help improve the "Sipasar Tourist Guide" application:

- It is recommended to support various languages other than Indonesian. This allows the application to be enjoyed not only by Indonesian citizens, but also by foreigners.
- Colors and elements can be made more attractive to increase the aesthetic appeal of the app.

It is hoped that the suggestions above can become valuable recommendations for the development of the "Sipasar Tour Guide" application.

ACKNOWLEDGMENT

We extend our gratitude to the Institute for Research and Community Service at Duta Wacana Christian University for funding our research. Special thanks are also extended to the respondents who agreed to participate in our team's pilot study.

REFERENCES

[1] S. Santoso, "Examining Relationships between Destination Image, Tourist Motivation, Satisfaction, and Visit Intention in Yogyakarta," *Expert Journal of Business and Management*, vol. 7, no. 1, pp. 82–90, 2019, [Online]. Available: http://Business.ExpertJournals.com.

[2] P. T. Prasetyaningrum, "Penerapan Analytical Hierarchy Process (Ahp) Untuk Mendukung Keputusan Pemilihan Desrinasi Tempat Wisata Daerah Istimewa Yogyakarta Untuk Para Wisatawan Mancanegara Non Asia," *Jurnal SIMETRIS*, vol. 10, no. 2, 2019.

[3] A. F. Octaviansyah, D. Darwis, and A. Surahman, "Sistem Pencarian Lokasi Bengkel Mobil Resmi Menggunakan Teknik Pengolahan Suara Dan Pemrosesan Bahasa Alami," 2019. [Online]. Available: http://maps.google.com.

[4] T. Rahmasari, P. Studi, K. Akuntansi, K. Kunci, and : Abstrak, "Perancangan Sistem Informasi Akuntansi Persediaan Barang Dagang Pada Toserba Selamat Menggunakan Php Dan Mysql," 2019.

[5] M. Rzeszewski, "Mapbox," in *Evaluating Participatory Mapping Software*, C. M. Burnett, Ed., Cham: Springer International Publishing, 2023, pp. 21–40. doi: 10.1007/978-3-031-19594-5_2.

[6] I. P. Hardani, "Sinergitas Antara Pemandu Wisata Dan Operator Tur (Study Kasus di CV Gondes Karya Mandiri)," 2021, [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/jp.

[7] S. Bose, "A Comparative Study: Java Vs Kotlin Programming In Android Application Development," *International Journal of Advanced Research in Computer Science*, vol. 9, no. 3, pp. 41–45, Jun. 2018, doi: 10.26483/ijarcs.v9i3.5978.

[8] P. Mersing, O. Veza, and N. H. Adi, "RANCANG BANGUN APLIKASI OBJEK WISATA KOTA BATAM BERBASIS ANDROID," 2018. [Online]. Available: https://www.researchgate.net/publication/343537519.

[9] A. Adil, B. Krismono Triwijoyo, I. Made, and Y. Dharma, "Implementasi Spasial Algoritma Harvesine pada Mapbox API untuk Pemetaan Pariwisata Spatial Implementation of the Harvesine Algorithm in The Mapbox API for Tourism Mapping," *Jurnal Bumigora Information Technology (BITe)*, vol. 5, no. 1, pp. 53–64, 2023, doi: 10.30812/bite/v5i1.2874.

[10] S. Ahmadi, G. Tack, D. Harabor, and P. Kilby, "Vehicle dynamics in pickup-and-delivery problems using electric vehicles," in *Leibniz International Proceedings in Informatics, LIPIcs*, Schloss Dagstuhl- Leibniz-Zentrum fur Informatik GmbH, Dagstuhl Publishing, Oct. 2021. doi: 10.4230/LIPIcs.CP.2021.11.

[11] A. Choudhary, I. Gogna, M. Deolekar, P. More, and R. Bhise, "Augmented Reality Based Turn-by-Turn Navigation System for Mobile Devices," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, no. 8, pp. 2278–3075, 2019, doi: 10.1007/978-3-319-47952.

[12] Y. Nugraha, "Information System Development With Comparison of Waterfall and Prototyping Models," 2020.

[13] A. Basith, Z. Arifin, M. Anshori, and A. Widya, "Aplikasi Pemandu Wisata Religi Di Jawa Timur Berbasis Android," 2021.

[14] "Badan Pusat Statistik Provinsi Daerah Istimewa Yogyakarta," 2024.

[15] "Badan Pusat Statistik Provinsi Bali," 2024.

[16] J. Teknika, D. Yuliawati, and A. Andriyadi, "Teknika 16 (02): 303-310 Pengujian Sistem Informasi E-Monitoring Pengelolaan Pembangunan Desa Dengan Menggunakan Metode Blackbox Testing," *IJCCS*, vol. x, No.x, no. 93, 2022.

[17] A. E. Roberts, T. A. Davenport, T. Wong, H. W. Moon, I. B. Hickie, and H. M. LaMonica, "Evaluating the quality and safety of health-related apps and e-tools: Adapting the Mobile App Rating Scale and developing a quality assurance protocol," *Internet Interv*, vol. 24, Apr. 2021, doi: 10.1016/j.invent.2021.100379.

AUTHOR GUIDELINES

1. Manuscript criteria

- The article has never been published or in the submission process on other publications.
- Submitted articles could be original research articles or technical notes.
- The similarity score from plagiarism checker software such as Turnitin is 20% maximum.
- For December 2021 publication onwards, Ultimatics : Jurnal Teknik Informatika will be receiving and publishing manuscripts written in English only.

2. Manuscript format

- Article been type in Microsoft Word version 2007 or later.
- Article been typed with 1 line spacing on an A4 paper size (21 cm x 29,7 cm), top-left margin are 3 cm and bottom-right margin are 2 cm, and Times New Roman's font type.
- Article should be prepared according to the following author guidelines in this template. Article contain of minimum 3500 words.
- References contain of minimum 15 references (primary references) from reputable journals/conferences

3. Organization of submitted article

The organization of the submitted article consists of Title, Abstract, Index Terms, Introduction, Method, Result and Discussion, Conclusion, Appendix (if any), Acknowledgment (if any), and References.

- Title
  The maximum words count on the title is 12 words (including the subtitle if available)
- Abstract
  Abstract consists of 150-250 words. The abstract should contain logical argumentation of the research taken, problem-solving methodology, research results, and a brief conclusion.
- Index terms
  A list in alphabetical order in between 4 to 6 words or short phrases separated by a semicolon (;), excluding words used in the title and chosen carefully to reflect the precise content of the paper.
- Introduction
  Introduction commonly contains the background, purpose of the research, problem identification, research methodology, and state of the art conducted by the authors which describe implicitly.
- Method
  Include sufficient details for the work to be repeated. Where specific equipment and materials are named, the manufacturer's details (name, city and country) should be given so that readers can trace specifications by contacting the manufacturer. Where commercially available software has been used, details of the supplier should be given in brackets or the reference given in full in the reference list.
- Results and Discussion
  State the results of experimental or modeling work, drawing attention to important details in tables and figures, and discuss them intensively by comparing and/or citing other references.
- Conclusion
  Explicitly describes the research's results been taken. Future works or suggestion could be explained after it
- Appendix and acknowledgment, if available, could be placed after Conclusion.
- All citations in the article should be written on References consecutively based on its' appearance order in the article using Mendeley (recommendation). The typing format will be in the same format as the IEEE journals and transaction format.

4. Reviewing of Manuscripts

Every submitted paper is independently and blindly reviewed by at least two peer-reviewers. The decision for publication, amendment, or rejection is based upon their reports/recommendations. If two or more reviewers consider a manuscript unsuitable for publication in this journal, a statement explaining the basis for the decision will be sent to the authors within six months of the submission date.

5. Revision of Manuscripts

Manuscripts sent back to the authors for revision should be returned to the editor without delay (maximum of two weeks). Revised manuscripts can be sent to the editorial office through the same online system. Revised manuscripts returned later than one month will be considered as new submissions.

6. Editing References

- **Periodicals**
  J.K. Author, “Name of paper,” Abbrev. Title of Periodical, vol. x, no. x, pp. xxx-xxx, Sept. 2013.

- **Book**
  J.K. Author, “Title of chapter in the book,” in Title of His Published Book, xth ed. City of Publisher, Country or Nation: Abbrev. Of Publisher, year, ch. x, sec. x, pp xxx-xxx.

- **Report**
  J.K. Author, “Title of report,” Abbrev. Name of Co., City of Co., Abbrev. State, Rep. xxx, year.

- **Handbook**
  Name of Manual/ Handbook, x ed., Abbrev. Name of Co., City of Co., Abbrev. State, year, pp. xxx-xxx.

- **Published Conference Proceedings**
  J.K. Author, “Title of paper,” in Unabbreviated Name of Conf., City of Conf., Abbrev. State (if given), year, pp. xxx-xxx.

- **Papers Presented at Conferences**
  J.K. Author, “Title of paper,” presented at the Unabbrev. Name of Conf., City of Conf., Abbrev. State, year.

- **Patents**
  J.K. Author, “Title of patent,” US. Patent xxxxxxx, Abbrev. 01 January 2014.

- **Theses and Dissertations**
  J.K. Author, “Title of thesis,” M.Sc. thesis, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year. J.K. Author, “Title of dissertation,” Ph.D. dissertation, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.

- **Unpublished**
  J.K. Author, “Title of paper,” unpublished. J.K. Author, “Title of paper,” Abbrev. Title of Journal, in press.

- **On-line Sources**
  J.K. Author. (year, month day). Title (edition) [Type of medium]. Available: http:// www.(URL) J.K. Author. (year, month). Title. Journal [Type of medium]. volume(issue), pp. if given. Available: http://www.(URL) Note: type of medium could be online media, CD-ROM, USB, etc.

7. Editorial Adress

Universitas Multimedia Nusantara
Jl. Scientia Boulevard, Gading Serpong
Tangerang, Banten, 15811
Email: **ultimatics@umn.ac.id**

Paper Title

Subtitle (if needed)

Author 1 Name[1], Author 2 Name[2], Author 3 Name[2]
[1] Line 1 (of affiliation): dept. name of organization, organization name, City, Country
Line 2: e-mail address if desired
[2] Line 1 (of affiliation): dept. name of organization, organization name, City, Country
Line 2: e-mail address if desired

Accepted on mmmmm dd, yyyy
Approved on mmmmm dd, yyyy

***Abstract*—This electronic document is a "live" template which you can use on preparing your ULTIMATICS paper. Use this document as a template if you are using Microsoft Word 2007 or later. Otherwise, use this document as an instruction set. Do not use symbol, special characters, or Math in Paper Title and Abstract. Do not cite references in the abstract.**

***Index Terms*—enter key words or phrases in alphabetical order, separated by semicolon (;)**

I. INTRODUCTION

This template, modified in MS Word 2007 and saved as a Word 97-2003 document, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. Margins, column widths, line spacing, and type styles are built-in here. The authors must make sure that their paper has fulfilled all the formatting stated here.

Introduction commonly contains the background, purpose of the research, problem identification, and research methodology conducted by the authors which been describe implicitly. Except for Introduction and Conclusion, other chapter's title must be explicitly represent the content of the chapter.

II. EASE OF USE

A. Selecting a Template

First, confirm that you have the correct template for your paper size. This template is for ULTIMATICS. It has been tailored for output on the A4 paper size.

B. Maintaining the Integrity of the Specifications

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them.

III. PREPARE YOUR PAPER BEFORE STYLING

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not add any kind of pagination anywhere in the paper. Please take note of the following items when proofreading spelling and grammar.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Abbreviations that incorporate periods should not have spaces: write "C.N.R.S.," not "C. N. R. S." Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units (SI units are encouraged).
- Do not mix complete spellings and abbreviations of units: "Wb/m2" or "webers per square meter," not "webers/m2." Spell units when they appear in text: "...a few henries," not "...a few H."
- Use a zero before decimal points: "0.25," not ".25."

C. Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number the equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop.

$$\int_0^{r_2} F(r,\phi)\, dr\, d\phi = [\sigma r_2 / (2\mu_0)] \qquad (1)$$

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the

equation. Use "(1)," not "Eq. (1)" or "equation (1)," except at the beginning of a sentence: "Equation (1) is ..."

D. Some Common Mistakes

- The word "data" is plural, not singular.
- The subscript for the permeability of vacuum μ_0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter "o."
- In American English, commas, semi-/colons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an "inset," not an "insert." The word alternatively is preferred to the word "alternately" (unless you really mean something that alternates).
- Do not use the word "essentially" to mean "approximately" or "effectively."
- In your paper title, if the words "that uses" can accurately replace the word using, capitalize the "u"; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones "affect" and "effect," "complement" and "compliment," "discreet" and "discrete," "principal" and "principle."
- Do not confuse "imply" and "infer."
- The prefix "non" is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the "et" in the Latin abbreviation "et al."
- The abbreviation "i.e." means "that is," and the abbreviation "e.g." means "for example."

IV. Using the Template

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention as below

ULTIMATICS_firstAuthorName_paperTitle.

In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper. Please take note on the following items.

A. Authors and Affiliations

The template is designed so that author affiliations are not repeated each time for multiple authors of the same affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

B. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include ACKNOWLEDGMENTS and REFERENCES, and for these, the correct style to use is "Heading 5."

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles, named "Heading 1," "Heading 2," "Heading 3," and "Heading 4", are prescribed.

C. Figures and Tables

Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation "Fig. 1," even at the beginning of a sentence.

TABLE I. TABLE STYLES

Table Head	**Table Column Head**		
	Table column subhead	***Subhead***	***Subhead***
copy	More table copy		

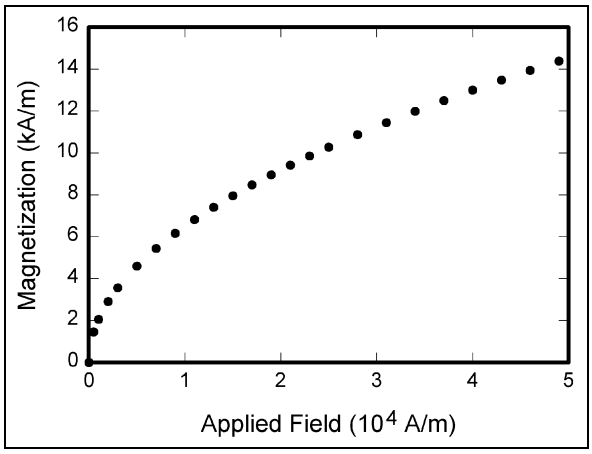


Fig. 1. Example of a figure caption

V. Conclusion

A conclusion section is not required. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

Appendix

Appendixes, if needed, appear before the acknowledgment.

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Avoid expressions such as "One of us (S.B.A.) would like to thank ... ." Instead, write "F. A. Author thanks ... ." You could also state the sponsor and financial support acknowledgments here.

References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

[1] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (*references*)

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[3] I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

ISSN 2085-4552

Universitas Multimedia Nusantara
Scientia Garden Jl. Boulevard Gading Serpong, Tangerang
Telp. (021) 5422 0808 | Fax. (021) 5422 0800