ISSN 2085-4579

Ultima InfoSys

Jurnal Sistem Informasi

SI
Sistem Informasi
Universitas Multimedia Nusantara

VOL 16 No.1

Engineering | Vol.16 | No.1 | Page 01-39 | June 2025 | ISSN 2085-4579

EDITORIAL BOARD

Ultima InfoSys : Jurnal Ilmu Sistem Informasi is a Journal of Information Systems 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. ULTIMA InfoSys Journal is published regularly twice a year (June and December) by Faculty of Engineering and Informatics in cooperation with UMN Press.

EDITORIAL ADDRESS

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

Call for Papers

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. IJNMT is published twice a year by Faculty of Engineering and Informatics of Universitas Multimedia Nusantara in cooperation with UMN Press.

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.

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 2085-4579

FOREWORD

Greetings!

Ultima InfoSys : Jurnal Ilmu Sistem Informasi is a Journal of Information Systems 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. ULTIMA InfoSys Journal is published regularly twice a year (June and December) by Faculty of Engineering and Informatics in cooperation with UMN Press.

In this June 2025 edition, ULTIMA InfoSys enters the 1st Edition of Volume 16. In this edition there are five scientific papers from researchers, academics and practitioners in the fields covered by Ultima Infosys. Some of the topics raised in this journal are: Implementation of Convolutional Neural Network Model for Apple Leaf Disease Detection; Improving the Mie GACOAN Online Food Ordering Application Using the System Usability Scale (SUS); Effect of Data Quality and Task Technology Fit on SIHA Performance in West Papua Province; Examining the Role of Social Media in Shaping Undergraduate Students' Study Preferences; Use of the K-Medoids Algorithm for Food Clustering Using Nutritional Value and Evaluation of the Elbow Method and the Davies Bouldin Index Method; Project Management and Cost Optimization in QR Code-Based Parking Reservation System Development: A PERT, Monte Carlo and PMBOK Approach.

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: International Journal of New Media Technology (IJNMT), Ultimatics : Jurnal Teknik Informatics, 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 ultimainfosys@umn.ac.id and the web page of our Journal here.

Finally, we would like to thank all contributors to this December 2024 Edition of Ultima Infosys. 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,

Monica Pratiwi, S.ST., M.T.
Editor-in-Chief

ISSN 2085-4579

TABLE OF CONTENT

ISSN 2085-4579

Implementation of Convolutional Neural Network Model for Apple Leaf Disease Detection

Ageng Cahyo Widjaya[1], Kimi Axel Wijaya[1], Joaquin Noah Soegono[1], Primus Kartika Varrel[1], Monika Evelin Johan[1]*

[1] Information System Study Program, Universitas Multimedia Nusantara, Tangerang, Indonesia
monika.evelin@umn.ac.id

Accepted on October 25, 2024
Approved on February 25, 2025

***Abstract*— Apple growers may suffer large financial losses as a result of apple leaf diseases. To reduce crop losses, apple leaf diseases must be identified early and treated. Nevertheless, conventional techniques for identifying apple leaf diseases, like professional manual visual inspection, can be time-consuming and difficult. Thus, the goal of this research is to detect diseases in apple tree leaves using convolutional neural networks (CNNs). By using deep learning, the disease detection process becomes automated, saving time and resources. It is proven that after roughly 20 epochs, the accuracy rise slows down and begins to fluctuate, but it keeps rising until it surpasses 90%. The results of the CNN model's performance in predicting disease types have a high level of accuracy and can be used as a model for detecting disease types in apple plant leaves.**

***Index Terms*— Apple Leaves Disease; Convolutional Neural Network; Deep Learning; Detection.**

I. INTRODUCTION

The production of fruits and food crops around the world has been greatly affected by various diseases, especially tree-leaf diseases. One type of tree that is widely cultivated throughout the world is the apple tree, one of the most popular fruits consumed worldwide [1]. Despite its high consumption rate, apple trees is prone to a number of diseases brought on by insects and microbes like bacteria. There are various types of diseases that affect this including Altenaria Boltch, AppleScab, Blackrot, CedarAppleCrust, and Rust. Therefore, apple leaf disease can cause significant economic losses for apple farmers. An early detection method for apple leaf disease is needed to minimize crop losses. However, traditional methods for detecting apple leaf disease, such as manual visual inspection by experts, can be time-consuming and labor-intensive.

Technological advances have led to deep learning and soft computing methods being applied in this area [2] becoming increasingly helpful for automatically identifying and categorizing apple tree leaf diseases. By using deep learning, the disease detection process can be done automatically so that it can save time and resources. In addition, the true benefit of deep learning is that its methods may be used directly on unprocessed data in a variety of file types, including.csv,.jpg, and others [3], [4]. Advances in computer vision and deep learning have resulted in the adoption of convolutional neural networks (CNNs) [5], [6]. Currently, CNN models are widely used as the primary choice for tasks related to image classification. The main strength of CNN lies in its architecture and effective feature extraction and transmission to the next layer. CNN usually consists of two main modules: a feature extraction module and a classifier module [7]. The feature extraction module is responsible for extracting relevant features from the image through processes such as convolution and pooling. CNN is robust to data variance, can automatically capture the hierarchy of features from low to high levels, and can solve gradient problems and appropriately modify parameters to focus on key characteristics with designs like Improved ResNet-50. This makes it better at automatically classifying pests and diseases of apple leaves. Based on the results obtained, it is evident that the model outperforms some previously proposed models in terms of performance metrics such as accuracy. Merging has effectively minimized the prediction variance, and can improve the accuracy [8], especially in scenarios involving many diseased leaves [9]. In addition, the model proposed in this study is also implemented through a web application, thus creating convenient accessibility for apple farmers [10], [11].

II. LITERATURE STUDY

A. *Convolutional Neural Network*

CNN is a deep learning algorithm used to classify labeled data, such as images. CNN operates by learning patterns in the data to predict target variables. In this study, several CNN parameters were used, including:

1. Convolutional Layer: A layer that functions to extract features from an input image by

repeatedly applying a convolutional kernel to identify patterns in the image [12], [13].

2. Activation Function: This study uses the Rectified Linear Unit (ReLU). The kernel or filter size used for each convolutional layer in this study is 3×3, which aims to to increase identification accuracy and expedite the training process [14], [15].
3. Pooling: is a process that reduces the spatial size of feature maps, allowing CNNs to be trained faster. Pooling also improves CNN's translational invariance, allowing it to recognize objects even if their location in the image changes [12], [16].
4. Flattening: After the pooling stage is complete, the flattening process will be carried out, namely flattening the pooling results into a fully connected layer. Flattening is the process of changing a feature map from a 2D form to a 1D form [12], [17]. This is done because fully connected layers only accept input in 1D.
5. Full Connection: Every neuron in a fully connected layer is coupled to every other neuron in the preceding and subsequent layers, enabling it to understand the connections between the features that the convolutional layers have collected [5], [12].

B. Previous Research

Several previous studies are the basis for the development of this research. Research [18] discusses the rapid development of deep learning (DL) techniques that make it possible to detect and recognize objects from images. This highlights that DL approaches have recently entered various agricultural applications after being successfully applied in various fields. According to the study's findings, farmers can increase agricultural yields by managing their crops more skillfully with the use of autonomous plant disease detection. Study [19] implements CNN as a solution used for plant disease classification. CNN consistently performs well in leaf disease classification due to its ability to automatically extract complex features, with accuracy increasing as the model is scaled up or combined with techniques such as data augmentation and transfer learning. CNNs use its layered structure—which includes pooling layers that lower the dimensionality of the input while preserving crucial information and convolutional layers that identify local variables like color patterns, textures, and shapes—to effectively diagnose plant illnesses. CNNs are able to collect progressively more sophisticated visual characteristics at each layer, ranging from basic patterns to more abstract illness representations, thanks to this hierarchical approach. Furthermore, the backpropagation technique optimizes weights to increase accuracy, and the usage of activation functions like ReLU aids the model in learning non-linearities on the data. CNNs are a potent tool for diagnosing plant diseases because of these methods, which enable them to differentiate between even similar disease signs [20]. Another study [21] found that the Deep Convolutional Neural Network (CNN) model is very effective in detecting apple leaf diseases, with an accuracy of 97.62% in classifying various diseases. The study showed that the use of data augmentation techniques and hyperparameter optimization improved the performance of the model, making it a reliable tool for automatic plant disease diagnosis. CNNs are successful in diagnosing plant diseases because it able to automatically extract and efficiently learn visual features such as color, texture, and shape that are typical for various diseases. CNNs' capabilities also include its capacity to handle vast data, adapt to changes in lighting, orientation, and image background, and be easily incorporated with transfer learning and data augmentation methods to increase diagnostic precision. Research [22] discusses the latest deep-learning methods for plant disease detection and classification. This study also addresses various challenges in this task, such as the variability of plant diseases and the lack of publicly available datasets. This study suggests that future research should focus on developing methods that can generalize well in classifying plant diseases. Research [23] compares classic deep learning and machine learning approaches in detecting plant diseases using images. The study's findings indicate that the Convolutional Neural Network algorithm has the highest performance compared to the Fully Connected Neural Network algorithm and machine learning such as SVM (Support Vector Machine) and KNN (K Nearest Neighbor). Based on previous research, we propose a solution using the CNN model to classify the type of disease in apple leaves, due to CNN's ability to accurately determine if a leaf is ill or not. According to the study published in [24], their lightweight CNN model demonstrated great classification efficacy while preserving computational efficiency, achieving an accuracy of 97.95% on a benchmark dataset for apple leaf disease. This outcome demonstrates the model's promise for resource-constrained and real-time applications, like in-field plant disease monitoring. Using CNN as a foundation for the Single Shot Detector (SSD) algorithm is suggested as a way to achieve the highest level of accuracy in detecting and categorizing apple leaf diseases. The suggested model beats other models of a similar nature, with an accuracy of 96.62% when benchmarked against models like AlexNet and ResNet-50 [25].

III. METHODOLOGIES

A. *Cross-Industry Standard Process for Data Mining (CRISP-DM)*

The research methodology used is CRISP-DM. The research will be continued through the following stages based on the CRISP-DM framework:

1. Business Understanding: The purpose of this project is to use a convolutional neural network (CNN) to identify illnesses in apple tree leaves.
2. Data Understanding: For this study, the dataset used [26] was images of apple tree leaves in diseased conditions consisting of 5 classes, namely AlternariaBoltch, AppleScab, BlackRot, CedarAppleRust, and Rust sourced from the Mendeley Dataset. The total size of this dataset is 861 MB with total 1,414 images. An example of apple leaf disease image data used is shown in Figure 1 below.
3. Data Preparation: Preprocessing image data to make it suitable for use in a CNN model, such as resizing images, normalizing pixel values, and dividing the data into training and testing sets.

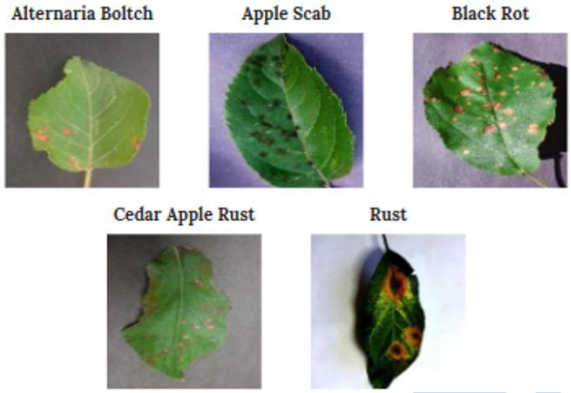


Fig 1. Apple Leaves with Diseases

4. Modeling: Creating a CNN model with CNN architecture, layer construction, and activation function selection.
5. Evaluation: To assess model performance, accuracy, and loss graphs, as well a confusion matrix will be used.
6. Deployment: Implementing a model that can classify diseases in apple leaves based on the input of apple tree leaf images on a website.

IV. RESULT AND DISCUSSIONS

A. *Data Preprocessing*

The data preparation stage is to perform image augmentation using the ImageDataGenerator library from Keras. Image augmentation is needed to create new training data from existing samples so that the trained data will be more. Parameters for the augmentation carried out include rescale to normalize image pixels, zoom_range for random zoom, rotation_range for random rotation, horizontal_flip and vertical_flip to flip the image randomly, shear_range for random shear transformation, fill_mode to fill newly created pixels after rotation or shift, width_shift_range and height_shift_range for random horizontal and vertical shifts, and validation_split to determine the fraction of data to be used as the validation set. Table 1 shows the parameters used for data (image) augmentation process.

TABLE I. PARAMETERS FOR DATA AUGMENTATION

Rescale	1.0/256.0
zoom_range	0.4
rotation_range	30
horizontal_flip	True
vertical_flip	True
shear_range	0.3
fill_mode	nearest
width_shift_range	0.2
height_shift_range	0.2
validation_split	0.2

B. *Convolutional Neural Network (CNN) Model*

The CNN model is made with a sequential model from Keras with one input layer and one output layer. In the first to third layers, the parameters used are a filter of 32, ReLU activation function, kernel size (3,3), and pool size (2,2). In the fourth to sixth layers, the parameters used are the same as the previous layers, but the filter used becomes 64. In the seventh to ninth layers, the filter parameters used are 128. In the tenth to twelfth layers, the filter used is 256. In the thirteenth to 17th layers, the parameters used are units of 256, 128, and 5, relu and softmax activation functions, and Dropout of (0.1). Then modeling is carried out with a total of 80 epochs. The model that is made is then loaded into h5 format for the deployment stage. Table 2 shows the details of the CNN model used.

TABLE II. CNN MODELS ARCHITECTURE

Layer (type)	Output Shape	Param
conv2d (Conv2D)	(None, 256, 256, 32)	896
conv2d_1 (Conv2D)	(None, 256, 256, 32)	9248
max_pooling2d (MaxPooling2D)	(None, 128, 128, 32)	0
conv2d_2 (Conv2D)	(None, 128, 128, 64)	18496
conv2d_3 (Conv2D)	(None, 128, 128, 64)	36928
max_pooling2d_1 (MaxPooling2D)	(None, 64, 64, 64)	0
conv2d_4 (Conv2D)	(None, 64, 64, 128)	73856
conv2d_5 (Conv2D)	(None, 64, 64, 128)	147584
max_pooling2d_2 (MaxPooling2D)	(None, 32, 32, 128)	0
conv2d_6 (Conv2D)	(None, 32, 32, 256)	295168
conv2d_7 (Conv2D)	(None, 32, 32, 256)	590080
max_pooling2d_3 (MaxPooling2D)	(None, 16, 16, 256)	0
flatten (Flatten)	(None, 655536)	0
dense (Dense)	(None, 256)	16777472
dropout (Dropout)	(None, 256)	0

dense_1 (Dense)	(None, 128)	32896
dense_2 (Dense)	(None, 5)	654

C. *Evaluation Model CNN*

The results of the modeling that has been created are depicted in the accuracy and loss graphs, as well as the confusion matrix which is explained in Figure 2 below.

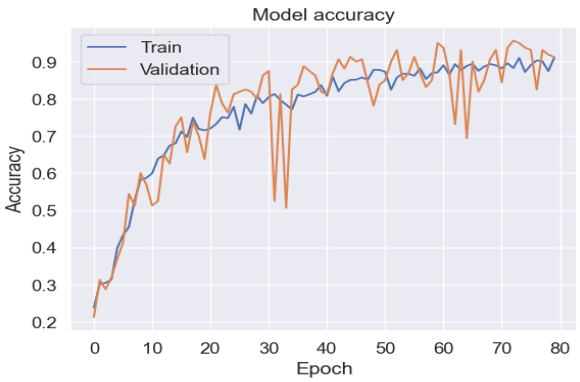


Fig 2. Model Accuracy with 80 Epochs

Based on Figure 5, training and validation accuracy increase with the number of epochs, indicating that the model learns from the data over time. After about 20 epochs, the increase in accuracy slows down and starts to fluctuate, but the accuracy continues to increase until it reaches above 90%. Validation accuracy is more volatile or has many spikes than training accuracy which is more consistent because the model generally performs better with previously seen data (training data) than with freshly acquired data (validation data).

Fig 3. Model Loss with 80 Epochs

Based on Figure 3, the model loss throughout the epoch continues to decrease significantly, indicating that the model learns from the data over time and becomes better at predicting the target.

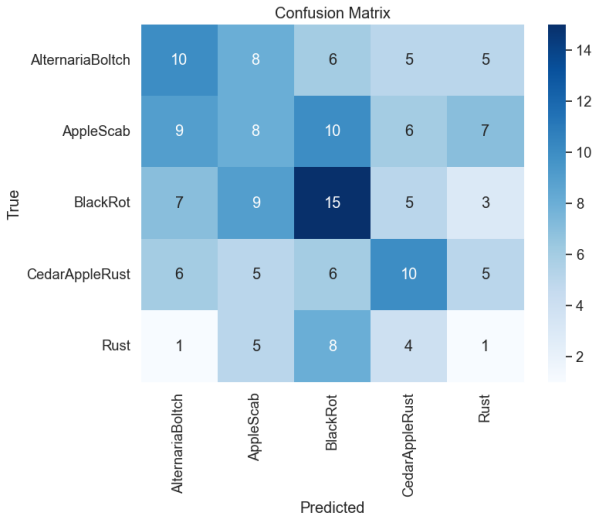


Fig 4. Confusion Matrix Result

Figure 4 shows the confusion matrix of correct predictions for each class. Based on the confusion matrix results, Table 3 below describes the Precision, Recall, and F1 Score values of each class. These results are taken from testing 164 test images.

TABLE III. SUMMARY OF THE RESULTS

Class	Precision (%)	Recall (%)	F1 Score (%)
AlternariaBoltch	30.30	29.41	29.85
AppleScab	22.86	20.00	21.33
BlackRot	33.33	38.46	35.71
CedarAppleRust	33.33	31.25	32.36
Rust	4.76	5.26	5.00

From the test results, it can be concluded that the CNN model created is the best at predicting BlackRot disease which is marked by the darkest area on the true and predicted diagonals. The model can also predict other disease classes well such as AlternariaBoltch, AppleScab, and CedarAppleRust. However, the model is not good or is still wrong in classifying the Rust class disease. This is likely because the image of the disease leaf in the Rust class is somewhat similar to the AppleScab class, predicting the type of disease slightly wrong.

D. *Deployment*

After the modeling and evaluation are complete, the model is deployed to the website. This deployment process is carried out using the Flask framework. On the website that is created, users can upload photos of apple leaves affected by the disease, then the model will process it and will tell the type of apple leaf disease from the image uploaded by the user. Figure 5 shows the interface of the web-based apple leaf disease classification, starting from the initial display, image upload results, and the classification results that come out.

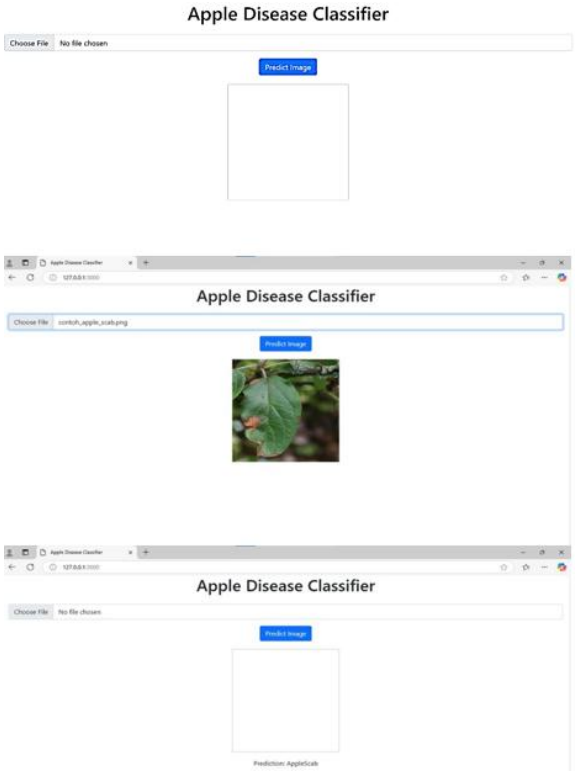


Fig 5. Deployment of Apple Leaf Disease Classification

V. CONCLUSION

Apple leaf disease can cause significant economic losses to apple farmers. Therefore, early detection and treatment of this disease are very important to minimize crop losses. The application of Convolutional Neural Networks (CNN) in apple leaf disease detection has been proven to be efficient and accurate in several previous studies. CNN's strength is primarily found in its capacity to extract useful features from images. As the number of epochs increased, there was a noticeable increase in both the training and validation accuracy graphs, indicating that the model has learned well from the data. The evaluation results with the confusion matrix also provide an overview of the model's performance in predicting disease types. The model successfully predicted several disease classes with a high level of accuracy, however, there were weaknesses in the classification of the "Rust" type of disease which was similar to the "AppleScab" type. In addition, the developed CNN model is not only efficient in analysis but can also be integrated into web applications. This increases accessibility and convenience of use for apple farmers. Suggestions for further research are to improve the accuracy of the model in predicting certain disease classes, especially the "Rust" class which is the weakness of the model in this study. One way that can be done is by increasing the number of datasets for Rust classes, or implementing more powerful image processing techniques such as augmentation.

ACKNOWLEDGMENT

We extend our sincere gratitude to Universitas Multimedia Nusantara for the invaluable support and resources provided throughout this project. The university's commitment to academic excellence and innovation has greatly contributed to the success of this work. We are especially grateful for the collaborative environment and the opportunities to enhance our research skills that UMN has facilitated.

REFERENCES

[1] A. Bonkra et al., "Apple Leave Disease Detection Using Collaborative ML/DL and Artificial Intelligence Methods: Scientometric Analysis," Int. J. Environ. Res. Public Health, vol. 20, no. 4, p. 3222, 2023, doi: 10.3390/ijerph20043222.

[2] L. Mohammed and Y. Yusoff, "DETECTION AND CLASSIFICATION OF PLANT LEAF DISEASES USING DIGTAL IMAGE PROCESSING METHODS: A REVIEW," ASEAN Eng. J., vol. 13, no. 1, pp. 1–9, 2023, doi: 10.11113/aej.v13.17460.

[3] P. Borkar, V. A. Wankhede, D. T. Mane, S. Limkar, J. V. N. Ramesh, and S. N. Ajani, "Deep learning and image processing-based early detection of Alzheimer disease in cognitively normal individuals," Soft Comput., 2023, doi: 10.1007/s00500-023-08615-w.

[4] B. V. D. S. Sekhar and A. K. Jagadev, "Efficient Alzheimer's disease detection using deep learning technique," Soft Comput., vol. 27, no. 13, pp. 9143–9150, 2023, doi: 10.1007/s00500-023-08434-z.

[5] M. Xin and Y. Wang, "Research on image classification model based on deep convolution neural network," EURASIP J. Image Video Process., vol. 2019, no. 1, p. 40, 2019, doi: 10.1186/s13640-019-0417-8.

[6] F. Sultana, A. Sufian, and P. Dutta, "Advancements in Image Classification using Convolutional Neural Network," in 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), IEEE, 2018, pp. 122–129. doi: 10.1109/ICRCICN.2018.8718718.

[7] X. Zhang et al., "Classification and Identification of Apple Leaf Diseases and Insect Pests Based on Improved ResNet-50 Model," Horticulturae, vol. 9, no. 9, p. 1046, 2023, doi: 10.3390/horticulturae9091046.

[8] D. A. Shah, E. D. De Wolf, P. A. Paul, and L. V Madden, "Accuracy in the prediction of disease epidemics when ensembling simple but highly correlated models," PLOS Comput. Biol., vol. 17, no. 3, p. e1008831, 2021, doi: 10.1371/journal.pcbi.1008831.

[9] N. Sharma, J. Dev, M. Mangla, V. M. Wadhwa, S. N. Mohanty, and D. Kakkar, "A Heterogeneous Ensemble Forecasting Model for Disease Prediction," New Gener. Comput., vol. 39, no. 3–4, pp. 701–715, 2021, doi: 10.1007/s00354-020-00119-7.

[10] R. Fabregas, T. Harigaya, M. Kremer, and R. Ramrattan, "Digital Agricultural Extension for Development," in Introduction to Development Engineering, Springer International Publishing, 2023, pp. 187–219. doi: 10.1007/978-3-030-86065-3_8.

[11] K. Phasinam, T. Kassanuk, and M. Shabaz, "Applicability of Internet of Things in Smart Farming," J. Food Qual., vol. 2022, pp. 1–7, 2022, doi: 10.1155/2022/7692922.

[12] L. Alzubaidi et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," J. Big Data, vol. 8, no. 1, p. 53, 2021, doi: 10.1186/s40537-021-00444-8.

[13] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," Artif. Intell. Rev., vol. 53, no. 8, pp. 5455–5516, 2020, doi: 10.1007/s10462-020-09825-6.

[14] R. H. K. Emanuel, P. D. Docherty, H. Lunt, and K. Möller, "The effect of activation functions on accuracy, convergence speed, and misclassification confidence in CNN text classification: a comprehensive exploration," J. Supercomput., 2023, doi: 10.1007/s11227-023-05441-7.

[15] Y. Jiang, J. Xie, and D. Zhang, "An Adaptive Offset Activation Function for CNN Image Classification Tasks," Electronics, vol. 11, no. 22, p. 3799, 2022, doi: 10.3390/electronics11223799.

[16] R. Nirthika, S. Manivannan, A. Ramanan, and R. Wang, "Pooling in convolutional neural networks for medical image analysis: a survey and an empirical study," Neural Comput. Appl., vol. 34, no. 7, pp. 5321–5347, 2022, doi: 10.1007/s00521-022-06953-8.

[17] E. Jeczmionek and P. A. Kowalski, "Flattening Layer Pruning in Convolutional Neural Networks," Symmetry (Basel)., vol. 13, no. 7, p. 1147, 2021, doi: 10.3390/sym13071147.

[18] B. Tugrul, E. Elfatimi, and R. Eryigit, "Convolutional Neural Networks in Detection of Plant Leaf Diseases: A Review," Agriculture, vol. 12, no. 8, p. 1192, 2022, doi: 10.3390/agriculture12081192.

[19] J. Lu, L. Tan, and H. Jiang, "Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification," Agriculture, vol. 11, no. 8, p. 707, 2021, doi: 10.3390/agriculture11080707.

[20] D. S. Joseph, P. M. Pawar, and R. Pramanik, "Intelligent plant disease diagnosis using convolutional neural network: a review," Multimed. Tools Appl., vol. 82, no. 14, pp. 21415–21481, 2023, doi: 10.1007/s11042-022-14004-6.

[21] P. Bansal, R. Kumar, and S. Kumar, "Disease Detection in Apple Leaves Using Deep Convolutional Neural Network," Agriculture, vol. 11, no. 7, p. 617, 2021, doi: 10.3390/agriculture11070617.

[22] Saleem, Potgieter, and M. Arif, "Plant Disease Detection and Classification by Deep Learning," Plants, vol. 8, no. 11, p. 468, 2019, doi: 10.3390/plants8110468.

[23] D. Radovanovic and S. Dukanovic, "Image-Based Plant Disease Detection: A Comparison of Deep Learning and Classical Machine Learning Algorithms," in 2020 24th International Conference on Information Technology (IT), IEEE, 2020, pp. 1–4. doi: 10.1109/IT48810.2020.9070664.

[24] L. Fu, S. Li, Y. Sun, Y. Mu, T. Hu, and H. Gong, "Lightweight-convolutional neural network for apple leaf disease identification," Front. Plant Sci., vol. 13, p. 831219, 2022.

[25] S. Subbotin, A. Oliinyk, T. Kolpakova, V. Perehuda, and D. Borovyk, "Apple leaf diseases detection using convolutional neural networks," 2024.

[26] L. Zhou, Q. Xiao, M. Taha, C. Xu, and C. Zhang, "Plant Disease Dataset," Mendeley Dataset, 2022. https://data.mendeley.com/datasets/tsfxgsp3z6/2.

ISSN 2085-4579

Improving the Mie GACOAN Online Food Ordering Application Using the System UsabilityScale (SUS)

Raihannissa Hatrinidinar Rasya[1], Denta Devina Tiara Maharani[1], Royan Alfianto[1], Cornelius Mellino Sarungu[1]

[1] Information System Department, Binus Online Learning, Bina Nusantara University, Jakarta, Indonesia
raihannissa.rasya@binus.ac.id, denta.maharani@binus.ac.id, royan.alfianto@binus.ac.id, cornelius.sarungu@binus.ac.id

Accepted on April 02, 2025
Approved on June 16, 2025

***Abstract*—Improvement of the online food ordering application through the System Usability Scale (SUS) method to enhance user satisfaction and streamline the ordering process. Driven bythe rapid expansion of online food ordering in Indonesia, this study aims to address long queues and improve customer convenience through optimized digital solutions. Current challenges include delayed ordering processes and limitations in existing application features, such as navigation issues, unclear language, and limited ordering options for dine-in, takeaway, and delivery. The research employs a mixed-methods approach,combining survey and interview data with competitive analysis, to identify critical areas for enhancement. Design improvements focus on user interface (UI) and user experience (UX) to simplify navigation, optimize payment processes, and make language more user-friendly. Initial usability tests resulted in a SUS scoreof 48, categorizing the application as "poor." After iterative improvements based on Agile Development, the score rose to 70, indicating an "acceptable" level of usability. This research highlights the importance of continuous UI/UX refinement to meet consumer expectations, with recommendations for ongoing updates and professional design involvement to maintain and enhance app performance. The findings indicate that improvements in the usability and functionality aspects of the application have a significant impact on customer satisfaction levels. This directly reinforces the position of online food ordering services as a competitive alternative in the digital food service industry.**

***Index Terms*—agile development; user interface design; usability scale; digital food service**

I. INTRODUCTION

In the era of Industry 4.0, technological advancements are experiencing significant acceleration, impacting various aspects of daily life. According to a survey conducted by the Indonesian Internet Service Providers Association (APJII) in 2019, it was recorded that 73.7% of the population—equivalent to 196.71 million individuals—were connected to the internet. This figure increased to 78.19% in 2023, reflecting a 4.49% growth in internet usage over a four-year period [1]. This increase in digital connectivity has a real impact on various sectors, including the food and beverage industry, where online food ordering services have shown significant growth. In Indonesia, the online food ordering market is estimated to grow at an annual rate of 11.5% from 2020 to 2024, indicating the increasingly important role of these services in modern consumption patterns [2].

Online food ordering services offer various advantages, including the creation of job opportunities and increased customer convenience through an easier and faster ordering process. In Indonesia, many restaurants have integrated this service into their operations, one of which is Mie Gacoan managed by PT Pesta Pora Abadi. Mie Gacoan is widely known for its distinctive noodle dishes at affordable prices, making it a favorite choice among students and young consumers. Despite achieving high popularity, Mie Gacoan still faces operational challenges, particularly in terms of long queues which are often the main complaint from customers [3].

The model proposed in this study focuses on optimizing the usability and functionality aspects of the online food ordering application Mie Gacoan. This research applies the System Usability Scale (SUS) method combined with an Agile Development approach to address various limitations in the user interface (UI) and user experience (UX). The SUS method is used as a reliable evaluation instrument to measure the system's usability level based on direct feedback from users. Meanwhile, the Agile approach allows for an iterative and adaptive development process, so that each system update can be responsively adjusted to the evolving needs of users.

The System Usability Scale (SUS) method was chosen in this research because it has proven effective in identifying usability issues with relatively low resource requirements. Its wide use across various fields has demonstrated a high level of reliability in evaluating user satisfaction and system effectiveness. Meanwhile, the Agile Development approach was

selected because of its adaptive and iterative characteristics, allowing for the direct integration of feedback to support continuous improvements in user interface (UI) and user experience (UX). The combination of these two methods results in a user-centered approach, making the developed solutions more aligned with customer needs and expectations.

This research makes a significant contribution to this field by demonstrating that systematic improvements in user interface and user experience (UI/UX) can directly enhance customer satisfaction levels in the food delivery service sector. By identifying and addressing specific usability challenges—such as navigation difficulties, unclear language usage, and inefficient payment processes—this study produces practical insights that can be applied not only by Mie Gacoan but also by similar businesses. The findings of this research, indicated by the increase in the System Usability Scale (SUS) score from 48 (poor category) to 70 (acceptable category), affirm the potential for ongoing development in UI/UX design to enhance customer experience and support overall operational efficiency.

User Experience (UX) refers to the perceptions, feelings, and level of satisfaction experienced by users when interacting with a product or service. A positive UX is created when software effectively meets the needs and goals of users, such as helping to complete tasks easily, thereby enhancing overall comfort and user satisfaction [4]. By considering user preferences, perceptions, emotions, and cognitive and physical reactions before, during, and after use, UX plays a crucial role in ensuring that products can provide optimal and sustainable services for their users [5].

The User Interface (UI) plays a strategic role in supporting the interaction between users and systems by providing visually appealing designs, easily accessible features, and well-organized navigation structures. An effective UI design contributes to enhanced usability by aligning visual elements—such as colors, animations, and responsive feedback—with the needs and expectations of users, thereby strengthening the overall user experience and satisfaction. Additionally, the UI serves as a communication bridge between users and the system, enabling efficient and intuitive information exchange across various platforms, including web applications, mobile, and software. To achieve optimal results, UI design must integrate aesthetic and performance aspects in a balanced manner to ensure that the appearance does not hinder the core functionality of the application [6].

Usability is a product evaluation method conducted by directly testing with users to assess the ease of use of a website's interface. This evaluation includes several key aspects, such as learnability, which measures how quickly users can complete basic tasks when first using the design; efficiency, which refers to the speed of task completion after users become familiar with the design; memorability, which evaluates how easily users can navigate back to the site after a period of not using it; error management, which encompasses the frequency, severity of errors, and ease of recovery from errors; and satisfaction, which describes the level of comfort and enjoyment users experience while interacting with the design [7].

The System Usability Scale (SUS) is a usability evaluation method developed by John Brooke in 1986 to measure user satisfaction with a system or application. This method uses a questionnaire consisting of 10 statements with a 5-point Likert scale, ranging from strongly agree to strongly disagree. SUS is designed to provide a quick and efficient assessment of ease of use and effectiveness of the user interface. The strength of this method lies in its ability to be applied in various contexts with a relatively small number of respondents, making it very suitable for use in the early stages of product development to effectively and reliably identify usability issues [8].

Usability testing measures how user-friendly an application is, involving testers to identify defects before launch. It applies to both web and mobile apps, focusing on ease of use, flexibility, and control. Reports indicate that 50% of developers' efforts go toward resolving usability issues, highlighting the importance of early- stage usability testing to meet user expectations. Additionally, experts found that 97% of mobile app users prioritize ease of use as their top requirement [9].

A prototype is an early version or model of a product created to test concepts, designs, or processes before mass production. Prototypes serve as tools to explore ideas, identify issues, and refine designs before the final product is developed. They allow designers and developers to see how a product will function in real life and make necessary adjustments to enhance functionality, aesthetics, and usability. While crucial for creating innovative solutions, prototype development can incur high costs and involve significant uncertainty, potentially leading to resource waste if not managed effectively [10].

Hybrid mobile applications combine native app features with web technologies, relying on HTML, CSS, and JavaScript for cross-platform development. The Ionic Framework facilitates this by using HTML, CSS, and JavaScript to create hybrid mobile apps, integrating Angular and Cordova for communication with mobile devices. In the context of m-commerce applications, particularly food delivery services from restaurants, the presence of an intuitive and user-friendly interface becomes very crucial. Additionally, the system must provide a secure payment processing mechanism, real-time order tracking features, as well as an easy-to-operate menu management system. This allows restaurants to update items and adjust prices effectively and efficiently. [11].

The Agile Development Method is an approach to software development that is widely known for its ability to quickly adapt to changes in user needs during the development cycle. This approach emphasizes short iterations that allow for continuous

evaluation and adjustment, as well as strengthening direct collaboration between development teams and users to ensure that the resulting software can meet dynamic business needs. Previous studies have indicated that Agile methods are very effective in the development of web and mobile applications, especially in fast-changing environments such as mobile commerce (m-commerce). The application of Agile in hybrid m-commerce applications enables efficient development processes and easier integration with external systems, thus enhancing application performance while improving the user experience [12].

Research on the development of hybrid applications for restaurant delivery services is motivated by issues related to a less intuitive user experience, which affects customers' difficulties in completing the purchasing process.The Agile Development method was used to develop a hybrid application that combines the benefits of both native and web-based applications. The resulting solution allows users to access m-commerce features, such as product browsing, ordering, and payments, through a single integrated platform [13].

Another study focused on the development of a web-based self- service system for food ordering in restaurants. This topic was chosen as a solution to minimize physical contact in restaurants during the COVID-19 pandemic. In this study, the System Usability Scale (SUS) method was used to test the usability of the application. The proposed solution involves designing a self-service application that allows customers to independently order food using QR code access, complemented by an admin management feature to support restaurant operations [14].

Study on the GrabFood mobile food ordering app revealed a lengthy loading screen, an invisible map feature when selecting addresses, missing food photos for some restaurants, an excessive number of restaurants in the “Delivering to You” section, and the absence of search filters by price or distance. The researchers then applied the Design Thinking method to evaluate and enhance the user experience for GrabFood. Additionally, they used Heuristic Evaluation to analyze usability issues and prioritize improvements based on severity ratings. The implemented changes significantly improved application efficiency, providing a better user experience, with users reporting increased satisfaction after these modifications [15].

The study on hybrid-based emergency call application that was developed to assist the Tangerang Regency community in handling emergency situations, such as accidents or fires has revealed the main issues that are the public’s limited awareness of emergency numbers and the difficulty in providing accurate location information [11]. The solution is an application that utilizes Location-Based Services (LBS) and the Ionic framework to provide quick access to emergency numbers and display critical locations, such as police stations, hospitals, and fire departments, through Google Maps. With these features, the app is expected to help people receive faster assistance and enhance safety during emergencies.

Therefore, a thorough investigation of the Gacoan application is needed to assess the level of user satisfaction and identify any challenges they may face. Thus, the scope of this paper will be limited to:

- The research focus is directed towards Gacoan noodle customers who are experiencing issues with long queues and delays when ordering food.
- This research is limited to the discussion of user interface and user experience development in the Gacoan noodle ordering application.
- User experience analysis is used to assess how satisfied Gacoan noodle customers are with the features of the developed Gacoan noodle application.

II. METHODOLOGY

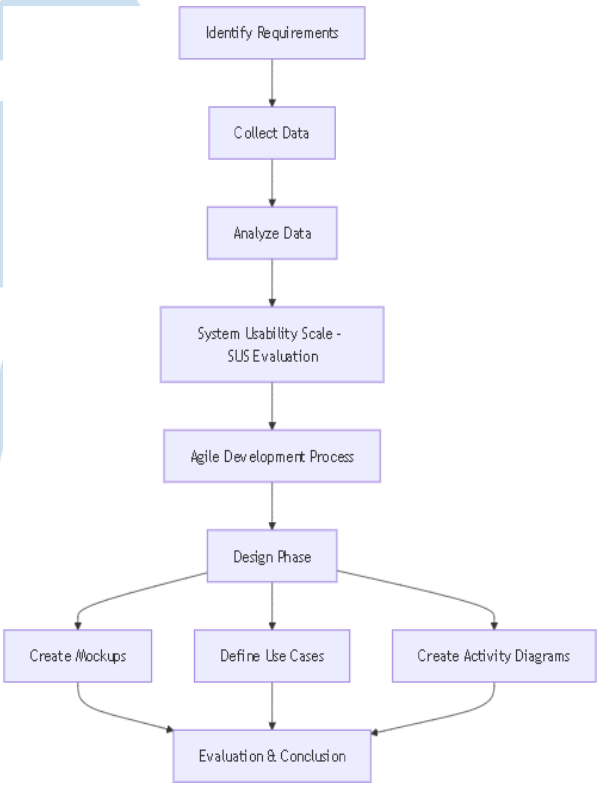


Fig. 1. Research Methodology.

The evaluation of the Mie Gacoan application was conducted to identify deficiencies in the user interface (UI/UX) and provide relevant development solutions. In this study, surveys and interviews targeted application users and loyal customers of Mie Gacoan restaurants [16]. These subjects were selected because they were considered capable of providing objective feedback regarding their experience using the application.

This survey involved a total of 73 respondents collected in one day at one of the Mie Gacoan branches. The selection of this sample size was based on the consideration of data representativeness while maintaining practicality in the analysis process according to the time constraints of the research. The respondents participated in usability testing of the application and were asked to fill out a System Usability Scale (SUS) questionnaire to evaluate their experience using the application.

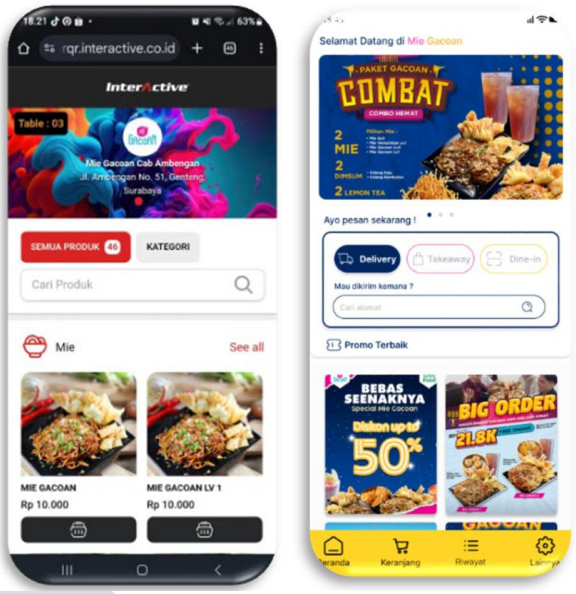
This research begins with a system development approach that starts with a needs identification process, aimed at understanding user expectations and formulating the overall system objectives. The next stage involves data collection through surveys and/or interviews, which are then analyzed to gain insights that support the design of an appropriate system solution. To assess the level of system usability, the System Usability Scale (SUS) instrument is used. In the development process, the Agile method is implemented to allow for flexibility and rapid adaptation through short and continuous iterations. [17].

Subsequently, the team moves to design, incorporating several key elements such as mockups as early visual representations, use cases to depict interaction scenarios between users and the system, and activity diagrams to illustrate process flows. Once all stages are completed, the results and evaluation of the development process lead to conclusions and recommendations for future improvements [18].

III. Results and Discussions

Testing of the Gacoan application was carried out by utilizing the smartphone used by the respondents, recording and observing. The following are the results of solving the user interface design problems along with their explanations:

1. The first problem in this research is related to the *usability* of food ordering applications, where the ordering process often takes a long time and reduces user convenience. Customers frequently face challenges in selecting the appropriate ordering service, whether for Delivery, Take Away, or Dine- In. Therefore, this research aims to design an application that can simplify the process for customers to manage queues and choose services according to their needs. With this application, it is expected that the ordering process will be faster and more efficient, ultimately enhancing the overall userexperience [20].



Before After

Fig. 2 Dashboard page.

2. The second problem in this research is to simplify navigation in the Gacoan app by adding features that were previously unavailable. These include optionsto increase the quantity of food and drinks ordered, as well as a timer for the payment process. If payment is not completed within the specified time, the order will automatically be canceled and returned to the main menu. This feature is particularly useful for Dine-In services, where customers may occasionally overlook payment, causing tables to remain marked as "occupied" and unavailable for other customers. It is hoped that this feature will improve the efficiency of table use.

Before After

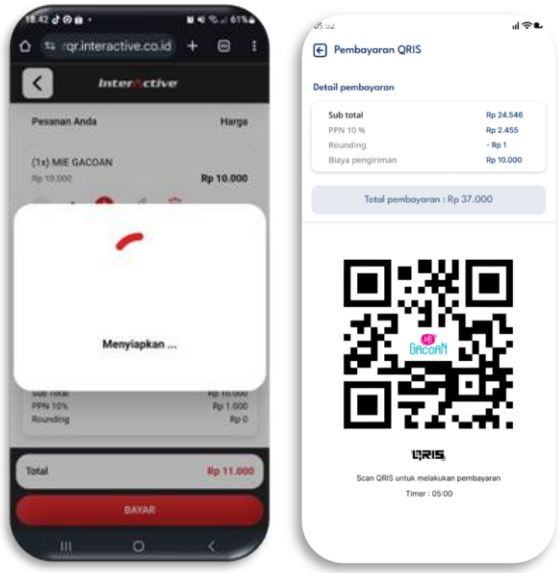


Fig. 3. Payment page.

3. The third problem in this research concerns the language used in the Gacoan app, which still poses difficulties for customers during the food ordering process. Many customers experience confusion due to unclear and unintuitive wording or terminology, which slows down the ordering process and reducesthe user experience. Therefore, improvements in the language used in the app are needed to make it more easily understood by a diverse range of users. With simpler and more familiar language, customers are expected to place orders more quickly and comfortably, without encountering confusion whileusing the app [21].

Before After

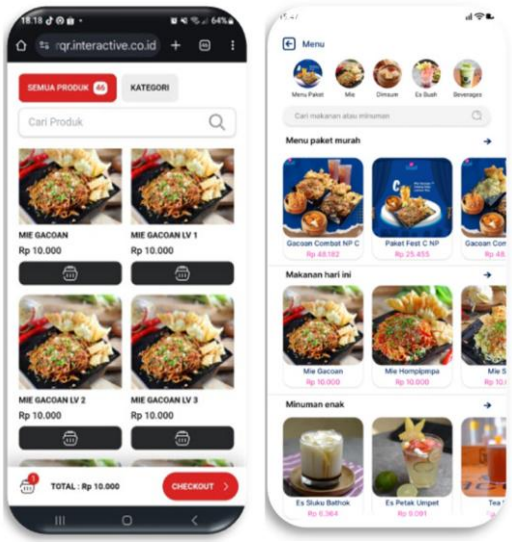


Fig. 4. Receipt page.

In this study, we distributed a questionnaireto Mie Gacoan customers to evaluate the application. The questionnaire aimed to gather user feedback both before and after testing. The data collected helped assess customer satisfaction, user experience, and areas for improvement. These insights were used to analyze the app's strengths and weaknesses and to identify necessary development steps. This approach ensures the app better meets user needs, enhances customer satisfaction, and provides an optimal user experience. The results of the questionnaire serve as a foundation for further application development. Below are the findings from the distributed questionnaire. The following are the final calculated scores for "after"assessments:

TABLE I. Questionnaire 1 score.

Question: I feel that the feature in the Mie Gacoan application are functioning well.

Response	Frequency	Percent	Valid Percent	Cumulative Percent
Very Disagree	25	34.2%	34.2%	34.2%
Disagree	18	24.7%	24.7%	58.9%
Neutral	17	23.3%	23.3%	82.2%
Agree	5	6.8%	6.8%	89.0%
Very Agree	8	11.0%	11.0%	100.0%
Total	**73**	**100.0%**	**100.0%**	

TABLE II. Questionnaire 2 score.

Question: I feel that the design of Mie Gacoan application is quite appealing.

Response	Frequency	Percent	Valid Percent	Cumulative Percent
Very Disagree	1	1.4%	1.4%	1.4%
Disagree	3	4.1%	4.1%	5.5%
Neutral	7	9.6%	9.6%	15.1%
Agree	24	32.9%	32.9%	47.9%
Very Agree	38	52.1%	52.1%	100.0%
Total	**73**	**100.0%**	**100.0%**	

TABLE III. Questionnaire 3 score.

Question: I feel the need for the addition of takeaway and delivery feature.

Response	Frequency	Percent	Valid Percent	Cumulative Percent
Sangat Tidak Setuju	1	1.4%	1.4%	1.4%
Netral	6	8.2%	8.2%	9.6%
Setuju	18	24.7%	24.7%	34.2%
Sangat Setuju	48	65.8%	65.8%	100.0%
Total	**73**	**100.0%**	**100.0%**	

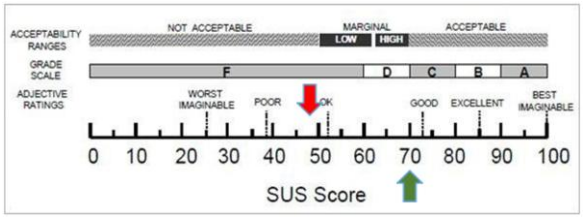


Fig. 5. Likert score after testing.

4. The fourth problem Gacoan application is that its usability level is still low. Based on the calculations that have been performed, it is known that the SUS score obtained before the development of the Gacoan noodle ordering application was 48. This aligns with the established SUS score criteria, wherea score of 48 falls into the adjective ratings category of "poor," marginal acceptability ranges of "low," and a grade scale of "F." After development, another usability test was conducted. Based on the results ofthe test, the SUS score obtained was 70, which indicates that this score falls within the adjective ratings of "Ok," marginal acceptability ranges of "high," and a grade scale of "D" [19].

IV. CONCLUSION

The research findings indicate that customer satisfaction with the Mie Gacoan ordering application has increased after development. The post-development System Usability Scale (SUS) score reached 70, which falls into the high category, while the pre-development score was only 48, classified as low. Based on the results of interviews and questionnaires, the appearance and features of the application have undergone significant improvements and now meet the needs and comfort of customers. Customers have stated that the changes in appearance and the addition of features provide a better user experience. The development of the application's features and appearance has proven to enhance the user experience. The newly added delivery and takeaway features greatly simplify the process for customers to order food online, eliminating the need for long queues.

PT. Pesta Pora Abadi should conduct regular maintenance of the Mie Gacoan application by updating existing features and enhancing the application's appearance to ensure it meets customer expectations. PT. Pesta Pora Abadi is encouraged to hire a professional UI/UX designer to improve the design and user experience of the Mie Gacoan ordering application, ensuring it is user-friendly and visually appealing.

ACKNOWLEDGMENT

Raihannissa Hatrinidinar Rasya contributes in conceptualization, methodology, formal analysis, data curation, and writing original draft. **Denta Devina Tiara Maharani** contributes in conceptualization, methodology, formal analysis, data curation, and writing original draft. **Royan Alfianto** contributes in conceptualization, methodology, formal analysis, data curation, and writing original draft. **Cornelius Mellino Sarungu** providing expert guidance, supervision, writing review either in conception or writing processes.

REFERENCES

[1] J. M. Dumalang, C. E. J. C. Montolalu, and D. Lapihu, "Perancangan UI/UX Aplikasi Penjualan Makanan Berbasis Mobile Pada Umkm Di Kota Manado Menggunakan Metode Design Thinking," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 2, no. 2, pp. 41–52, Sep. 2023. doi:10.58602/jima-ilkom.v2i2.19

[2] I. Aprilianti and F. Amanta, "Memajukan Keamanan Pangan Pada Layanan Pesan Antar Makanan Daring Di Indonesia," *Center for Indonesian Policy Studies*, vol. 28, p. 7, 2020. doi:10.35497/324009

[3] L. Valentine, M. N. Aisyah, S. B. Riono, A. Yulianto, and M. B. Zaman, "Analisis Persepsi Harga, Cita Rasa, Kualitas Layanan Dan Fasilitas Terhadap Keputusan Pembelian Ulang Mie Gacoan Tegal," *GEMILANG: Jurnal Manajemen dan Akuntansi*, vol. 3, no. 4, pp. 409–428, Oct. 2023. doi:10.56910/gemilang.v3i4.1309

[4] S. R. Henim and R. P. Sari, "Evaluasi User Experience Sistem Informasi Akademik Mahasiswa Pada Perguruan Tinggi Menggunakan User Experience Questionnaire," *Jurnal Komputer Terapan*, vol. 6, no. 1, pp. 69–78, May 2020. doi:10.35143/jkt.v6i1.3582

[5] R. Auliazmi, G. Rudiyanto, and R. D. Utomo, "Kajian Estetika Visual Interface Dan User Experience Pada Aplikasi RUANGGURU Aesthetic Studies of Visual Interface and User Experience of The Ruangguru Application," *Jurnal Seni dan Reka Rancang: Jurnal Ilmiah Magister Desain*, vol. 4, no. 1, pp. 21–36, Aug. 2021. doi:10.25105/jsrr.v4i1.9968

[6] S. Milah, M. A. Sunandar, and M. andayani Komara, "Perancangan User Interface Dan User Experience APLIKASI Penyewaan Peralatan Bayi Berbasis Mobile Menggunakan Metode Human Centered Design (HCD) Di Babystuffrent Purwakarta," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 3, pp. 1714–1721, Nov. 2023. doi:10.36040/jati.v7i3.7066

[7] I. Setiawan, Z. U. Hani, S. Serli, S. Rahayu, and W. Audiana, "Evaluasi Kegunaan Website Mahasiswa Universitas Amikom Purwokerto Menggunakan Metode Usability Testing," *Merkurius : Jurnal Riset Sistem Informasi dan Teknik Informatika*, vol. 2, no. 6, pp. 119–130, Oct. 2024. doi:10.61132/merkurius.v2i6.426

[8] N. Hibban, R. Albaihaqi, D. B. Rifai, and H. Aiman, "Evaluasi Usability Aplikasi Zenly Menggunakan Metode Usability Testing," *Jurnal SAINTEKOM*, vol. 13, no. 2, pp. 136–148, Sep. 2023. doi:10.33020/saintekom.v13i2.399

[9] S. Jain and P. Purandare, "Study of the usability testing of e-commerce applications," *Journal of Physics: Conference Series*, vol. 1964, no. 4, p. 042059, Jul. 2021. doi:10.1088/1742-6596/1964/4/042059

[10] K. M. Edwards, B. Man, and F. Ahmed, "Sketch2Prototype: Rapid Conceptual Design Exploration and Prototyping With Generative AI," *Proceedings of the Design Society*, vol. 4, pp. 1989–1998, May 2024. doi:10.1017/pds.2024.201

[11] R. M. Repanovici, Șerban Nedelcu, L. A. Tarbă, and S. Busuioceanu, "Improvement of Emergency Situation Management Through An Integrated System Using Mobile Alerts," *Sustainability*, vol. 14, no. 24, p. 16424, Dec. 2022. doi:10.3390/su142416424

[12] S. Al-Saqqa, S. Sawalha, and H. Abdelnabi, "Agile Software Development: Methodologies and Trends," *International*

Journal of Interactive Mobile Technologies, vol. 14, no. 11, pp. 246–270, 2020. doi:10.3991/ijim.v14i11.13269

[13] U. Rio, D. Haryono, W. Agustin, and S. Imardi, "Hybrid Application Mobile Commerce Resto Delivery Menggunakan Metode Agile Developmet," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 3, no. 3, pp. 518–525, Jan. 2023. doi:10.37859/coscitech.v3i3.4498

[14] A. A. Susila and D. M. Sri Arsa, "Analisis System Usability Scale (Sus) Dan Perancangan Sistem Self Service Pemesanan Menu Di Restoran Berbasis Web," *Majalah Ilmiah UNIKOM*, vol. 21, no. 1, pp. 3–8, Apr. 2023. doi:10.34010/miu.v21i1.10683

[15] T. A. Adinugraha and E. Z. Astuti, "Rancangan Bangun model desain tampilan user experience (UX) Sistem Aplikasi layanan Hiburan Pemesanan Tiket bioskop berbasis mobile Dengan Menggunakan Konsep Metode design thinking," *Techno Creative*, vol. 1, no. 1, p. 76, Aug. 2023. doi:10.62411/tcv.v1i1.1383

[16] I. Dewi *et al.*, "Marketing strategy and consumer satisfaction in gen Z's purchasing decision (study case at mie Gacoan bogor)," *E3S Web of Conferences*, vol. 454, p. 03021, 2023. doi:10.1051/e3sconf/202345403021

[17] M. K. Anam, S. Susanti, N. Nurjayadi, F. Zoromi, and A. K. Sari, "Evaluation of an existing system using the system usability scale (SUS) as a guideline for system improvement," *Jurnal Teknik Informatika*, vol. 18, no. 1, pp. 122–132, Apr. 2025. doi:10.15408/jti.v18i1.40766

[18] B. Hnatkowska and M. Cebinka, "Activity diagram generation based on use-case textual specification," *Computing and Informatics*, vol. 40, no. 4, pp. 772–795, 2021. doi:10.31577/cai_2021_4_772

[19] R. Novariany and R. D. Kusuma, "Evaluasi Pengalaman Pengguna aplikasi HIIP Indonesia Menggunakan METODE use questionnaire Dan usability testing," *Jurnal Informatika Kesatuan*, vol. 3, no. 2, Sep. 2023. doi:10.37641/jikes.v3i2.2352

[20] Y. Xu, X. Liu, Z. (Eddie) Mao, and J. Zhou, "Mobile food ordering apps, restaurant performance, and customer satisfaction," *Cornell Hospitality Quarterly*, vol. 65, no. 4, pp. 497–508, Jan. 2024. doi:10.1177/19389655231223376

[21] H. M. Az-zahra, A. Pinandito, and H. Tolle, "Usability evaluation of mobile application in culinary recommendation system," *2015 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, pp. 89–94, Aug. 2015. doi:10.1109/apwimob.2015.7374938

ISSN 2085-4579

Effect of Data Quality and Task Technology Fit on SIHA Performance in West Papua Province

Tresia Novita Susana Wake[1], Dedi I. Inan[1], Ratna Juita[1], Muhamad Indra[1]
[1] Department of Informatics Engineering, Universitas Papua, Papua Barat, Indonesia
202165021@student.unipa.ac.id
d.inan@unipa.ac.id
r.juita@unipa.ac.id
m.indra@unipa.ac.id

Accepted on June 02, 2025
Approved on June 16, 2025

***Abstract*— HIV/AIDS remains a major public health issue in Indonesia, especially in geographically challenging regions like West Papua Province. To support monitoring efforts, the HIV/AIDS and IMS Information System (SIHA) has been in use since 2012. However, poor data quality often limits its effectiveness. This study analyzes how data quality in SIHA affects the effectiveness of HIV/AIDS monitoring using the Task Technology Fit (TTF) framework. A quantitative method was applied through the distribution of questionnaires to 103 SIHA users, and the data were analyzed using SEM-PLS. The results indicate that data quality significantly influences task-technology fit (TTF) and system performance, while task, technology, and individual characteristics did not show a significant influence. The R-Square values for TTF and performance were 0.234 and 0.212 respectively, meaning that this model only explains approximately 21–23% of the variance in these variables. This suggests that there are still other influential factors that have not yet been investigated in this study. These findings provide fundamental insights into the importance of data quality for health information systems in challenging areas, but further research is needed to explore other variables that could enhance the effectiveness of SIHA in HIV/AIDS monitoring and control in West Papua.**

Keywords: Data Quality, Disease Spread Monitoring, HIV/AIDS, SIHA, Task Technology Fit, West Papua

I. INTRODUCTION

Infectious diseases remain a major challenge in many countries, including Indonesia. One disease that receives special attention is HIV, because until now no country has been completely free from this problem [1]. HIV itself is a virus that attacks the immune system, making those infected more vulnerable to various other infections [2]. HIV continues to be a significant global public health issue because individuals who do not receive treatment will experience chronic infection that can progress to Acquired Immunodeficiency Syndrome (AIDS), a condition resulting from a weakened immune system that carries a risk of mortality [3]. In Indonesia, HIV case detection within the national population from 2021 to 2023 shows an increase in tests conducted, correlating with a rise in identified positive individuals. The majority of people with HIV (PLHIV) are within the productive age group of 25–49 years (69.5%), although the positivity rate has remained relatively stable at around 0.9% to 1.1%. As of 2024, the total number of PLHIV recorded was 503,261, with 351,378 individuals aware of their status. Of this number, 217,482 people (62%) are receiving antiretroviral (ARV) treatment, 99,463 individuals (46%) have undergone viral load testing, and 91,662 individuals (42%) have successfully achieved viral suppression in their bodies [4]. A similar situation is observed in developing regions, such as West Papua. According to reports from the Provincial Health Office Chief, from 2013 to 2025, over 214,000 people have undergone HIV tests, and nearly 6,000 of them tested positive. However, only about 1,400 individuals have consistently received ARV therapy to suppress viral development in their bodies [5].

Amid various challenges, the use of information technology in the digital era has increasingly supported efforts to address HIV/AIDS. As digital transformation progresses, information technology has become an essential component in facilitating these initiatives. Based on Government Regulation No. 46 of 2014 concerning the Health Information System, health data, information, and indicators must

be classified and regulated within the Health Information System to support the implementation of health improvement [6]. The Health Office has launched several applications to assist in managing health data and information through electronic platforms, particularly in the field of Communicable Disease Prevention and Control (P2PM). One such application is the HIV-AIDS and STIs Information System (SIHA). SIHA serves as the official system for documenting and reporting cases of HIV/AIDS and Sexually Transmitted Infections (STIs) across national, provincial, and district or city levels. SIHA has been used by the Health Office since 2012 [7]. Therefore, it is important to analyze SIHA to ensure the system can provide accurate data that meets user needs. This analysis helps identify system weaknesses, assess technological suitability, and improve data quality to support more effective HIV/AIDS monitoring and control.

A study conducted by [8] aimed to identify challenges in HIV surveillance data quality in the United States, including issues of representativeness, completeness, accuracy, and the level of detail of information affecting the country's ability to track and respond to HIV epidemiological trends. Therefore, this study emphasizes the importance of high-quality data in information systems for effective monitoring of HIV/AIDS transmission. The study by [9] explains that Task Technology Fit (TTF) involves the relationship between existing tasks, individual capabilities, and the role of technology. This means that the ability to complete a task depends on the support of technological functions. Similarly, a study by [10] also developed the Task Technology Fit (TTF) theory in the context of mobile information systems supporting administrative tasks. This research highlights the importance of the fit between tasks, technology, and the context of mobile information system usage.

In this study, the authors apply the Task Technology Fit (TTF) theory, which centers on the compatibility between task demands and the capabilities of the supporting technology. This includes the extent to which the system assists users, regardless of their skill level [9]. According to [11], the Task Technology Fit (TTF) model is a formal framework, also known as Task-Technology Fit, which refers to the degree to which technology features align with task needs within a work context, particularly the potential of information technology to support task execution.

The studies conducted by [9], [10], and [11] have several limitations in examining Task Technology Fit (TTF). Goodhue [9] did not fully consider contextual factors, the role of users, and technological developments, making the model less flexible. Gabauer [10] also did not thoroughly discuss organizational factors and the long-term impact of mobile information systems. Meanwhile, [11] only used a partial set of TTF variables, employed a cross-sectional research design that did not capture changes over time, and produced findings that are difficult to generalize to other sectors. Overall, these three studies emphasize the importance of alignment between tasks and technology, but still have limitations in considering other variables impacting the success of information system.

This study aims to fill the shortcomings of earlier research by examining not just the alignment between tasks and technology as described by the Task Technology Fit (TTF) theory, but also by incorporating an additional external factor, namely Data Quality (DQ). The inclusion of the DQ variable can close previous gaps because DQ ensures that data reported through SIHA is more accurate, complete, consistent, and timely. This, in turn, enhances the reliability of information for decision-making and program evaluation. Therefore, it's crucial to understand how data quality within SIHA contributes to HIV/AIDS monitoring performance in this region. Given the significant role of data quality in supporting reliable health information systems, especially in HIV/AIDS reporting, a more in-depth study of the system in use is necessary. This research focuses on data quality within SIHA and its impact on the performance of HIV/AIDS transmission monitoring in West Papua. Using the Task Technology Fit (TTF) framework, this research seeks to examine how data quality and technology alignment relate to and influence the success of monitoring and controlling the spread of HIV/AIDS. However, unlike previous studies, this research is conducted in a developing region, West Papua, an area with unique geographical characteristics and accessibility challenges that add complexity to health data management.

Through the findings of this research, we hope to gain in-depth insights that are not only relevant for West Papua but can also serve as a reference for other regions facing similar conditions in improving the effectiveness of health information systems for epidemic control. The Task Technology Fit (TTF) framework is applied to evaluate how well the SIHA system meets the practical requirements of field users, including activities like documenting, reporting, and overseeing HIV/AIDS case management. By understanding this fit, the research can identify whether the technology used supports or hinders task execution. Furthermore, the addition of the Data Quality (DQ) aspect is crucial to ensure that the collected data is accurate, complete, consistent, and timely. Good data quality is vital for effective monitoring and decision-making, serving as a critical foundation for developing targeted interventions. Thus, the integration of the Task Technology Fit and Data Quality approaches is expected to make a tangible contribution to strengthening health information systems, not only in West Papua but also nationally, in enhancing the effectiveness of health information systems for epidemic control.

This research centers on examining the influence of Data Quality within the HIV/AIDS Information System (SIHA) on the effectiveness of monitoring HIV/AIDS transmission in West Papua from the Task Technology Fit (TTF) perspective. Digital information systems, such as SIHA, play a crucial role in managing public health data, particularly in monitoring and controlling infectious diseases like HIV/AIDS. However, the effectiveness of SIHA is not solely dependent on its existence but also on the extent to which users perceive the system as having high-quality data and how this influences its performance in supporting healthcare tasks, such as recording, reporting, and decision-making related to HIV/AIDS transmission monitoring.

II. METHODOLOGY

A. *Task Technology Fit in SIHA*

The Task-Technology Fit (TTF) approach developed by Goodhue [12] serves as the basis for designing the user evaluation instrument. This instrument aims to assess the extent to which the information system supports the execution of managerial tasks, particularly in utilizing organizational data for decision-making. Additionally, these instruments are used to test various propositions related to factors influencing user evaluations and their resulting impacts.

According to [12], Task-Technology Fit (TTF) consists of several key elements, including task characteristics, which encompass complexity, information dependence, and the nature of the task; technology characteristics, which include ease of use, system capabilities, and supporting features; and the degree of fit between technology and tasks, which determines the effectiveness of an information system. TTF also looks at how it affects how well people or organizations perform. Basically, the better the match between the task and the tech, the more it helps improve work results.

The TTF approach developed by [9] introduces a distinct difference by incorporating the utilization of technology by users in performing their tasks. In this model, Goodhue and Thompson (1998) emphasize that the fit between tasks and technology not only has a direct impact on performance but also influences it through the extent to which the technology is actually used by its users.

B. *Data Quality in SIHA*

According to [13], Data Quality is assessed based on several key aspects, including data validity, which relates to the relevance and accuracy of information in supporting decision-making.

In the use of SIHA, Data Quality plays a crucial role in supporting the effectiveness of monitoring and decision-making related to HIV/AIDS mitigation. Relevant and accurate data enable healthcare professionals to obtain the right information for planning and implementing HIV/AIDS prevention and treatment measures. Additionally, maintaining accurate data is also a key factor, which includes efforts to ensure that data remains up-to-date, consistent, and error-free throughout the processes of collection, storage, and utilization.

It can be concluded that good Data Quality not only enhances the efficiency of system users but also strengthens trust in the information generated, thereby supporting more effective and targeted decision-making. With high-quality data, monitoring performance can be improved through early case detection, more accurate tracking of transmission trends, and better-targeted prevention and treatment programs. Conversely, poor data can lead to delays, analytical errors, and misguided decisions, ultimately hindering the effectiveness of disease mitigation efforts.

This study employs a quantitative method, which focuses on collecting and analyzing numerical data to examine the relationship between variables [14]. In this study, a quantitative method is applied to analyze the impact of data quality in the HIV/AIDS Information System on the performance of disease transmission monitoring in West Papua.

This study's research framework is based on the Task Technology Fit (TTF) model, which has been extended by adding several independent variables: Task Characteristics, Technology Characteristics, Individual, and Data Quality. The dependent variable is Performance Impact. Meanwhile, SIHA's Task Technology Fit works as a mediator that helps show how those independent variables affect system performance.

The research model indicates that the Task Technology Fit (TTF) within the HIV/AIDS Information System (SIHA) directly affects how well HIV/AIDS monitoring is carried out. Data quality, along with task, technology, and individual characteristics, affects TTF, while data quality also has a direct impact on performance. In conclusion, high-quality data and a system that aligns with user needs are key factors in enhancing the effectiveness of HIV/AIDS monitoring in West Papua.

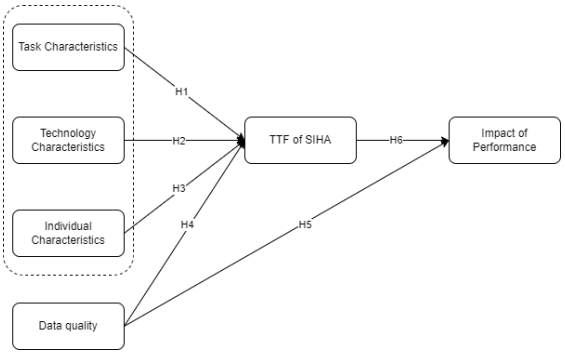


Fig 1. Research Model

From the conceptual framework shown above, all the variables are linked to each other. A hypothesis is a provisional statement that requires testing. In this context, hypotheses can be categorized into research hypotheses and working hypotheses (statistical hypotheses). A researcher needs hypotheses to guide their research plan and steps. A hypothesis is stated as a temporary truth and serves as both a working foundation and a guideline for data analysis [15].

According to [16] Task characteristics are defined as actions performed by individuals to transform inputs into outputs in order to meet the information needs of task fit. Therefore, this study proposes the following hypothesis (H1):

H1: Task Characteristics have a significant influence on the Task Technology Fit in the use of SIHA.

According to [17] Technology characteristics indicate that information technology consists of various tools that not only assist individuals in managing information and completing data processing tasks but also include communication technologies that enable information transmission to support organizational needs. Therefore, this study proposes the following hypothesis (H2):

H2: Technology Characteristics have a significant influence on the Task Technology Fit in the use of SIHA.

According to [16] The role of individuals in the use of information technology is crucial. This is because technological skills, experience, awareness, and perceived benefits of technology can influence the level of acceptance and effectiveness of its use. Therefore, this study proposes the following hypothesis (H3):

H3: Individual Characteristics have a significant influence on the Task Technology Fit in the use of SIHA.

According to [13] Data quality is assessed based on several key aspects, including data validity, which relates to the relevance and accuracy of information in supporting decision-making. Additionally, maintaining accurate data is also a crucial factor. Therefore, this study proposes the following hypotheses (H4 and H5):

H4: Data Quality has a significant influence on the Task Technology Fit in the use of SIHA.

H5: Data Quality has a significant influence on the Impact of Performance in the use of SIHA.

According to User evaluation of Task Technology Fit will have explanatory power in predicting the perceived impact on performance. Therefore, this study proposes the following hypothesis (H6):

H6: Task Technology Fit has a significant influence on the Impact of Performance in the use of SIHA.

C. Sampel

The population in this study consists of SIHA users in West Papua Province. To assess the validity of the analysis, the author also used the G*Power tool. By setting the effect size at 0.15, the alpha significance level at 5%, and the statistical power at 95%, with six predictor variables, the minimum required sample size was 75 respondents. However, to enhance the reliability of the results, the author collected a larger dataset, totaling 103 respondents [18]. Data collection was conducted by distributing questionnaires via Google Forms, with assistance from the Head of the Disease Prevention and Control Division of the West Papua Provincial Health Office.

D. Analysis Method

The research and data collection process was conducted over three months, from October to December 2024, in West Papua, with a total of 103 valid respondents. The demographic data collected included gender, age, highest level of education, and work unit.

TABLE I. RESPONDENTS DEMOGRAPHIC DESCRIPTION

Category	Item	Quantity	Percentage
Gender	Male	33	32%
	Female	70	68%
Age	18 - 35 years	49	47,60%
	36 - 50 years	54	52,40%
Highest Level of Education	(D1/D2/D3	54	52,40%
	(S1)	47	45,60%
	(S2)	1	1%
	Other	1	1%

III. RESULT AND DISCUSSION

A. Measurement Model Evaluation (Outer Model)

The outer model is the first of two steps in data analysis. In this study, three criteria are used to evaluate the outer model: convergent validity, composite reliability, and discriminant validity. The first criterion, convergent validity, is assessed by examining the Loading Factor (LF), where an LF value is generally considered acceptable if it is greater than 0.7 [16]. Next, the measurement of Average Variance Extracted (AVE) is considered, with an AVE value of 0.5. An AVE value of 0.5 or higher is deemed adequate, indicating that the latent construct can explain more than half of the variance in its indicator variables [19]. In the reliability test, each variable is evaluated using two methods: Cronbach's Alpha (CA) and Composite Reliability (CR). Each variable indicator is considered reliable if the CA and CR values are greater than 0.70 [20].

In this research instrument, some statements were removed. To ensure that the indicator removal process was conducted scientifically, planned, and justifiable, the researchers established instrument evaluation criteria based on relevant references, where, Loading Factor (LF) An indicator is declared valid if its loading factor value is ≥ 0.7. Average

Variance Extracted (AVE) A construct is declared valid if its AVE is ≥ 0.5. And Cronbach's Alpha (CA) and Composite Reliability (CR) A construct is considered reliable if its CA and CR values are ≥ 0.7. Therefore, indicators that did not meet these criteria in this study were considered for removal and were not included in the Research Instrument table. After the indicator removal, the AVE, CA, and CR values of each construct were re-evaluated to ensure they continued to meet the established criteria. This process was carried out gradually and iteratively until all remaining indicators were valid and reliable.

TABLE II. RESEARCH INSTRUMENT

Variable	Statement Item	Code	LF	Reference
Task Characteristics (TSC) CA, CR, AVE = 0.856, 0.933, 0.874	I often encounter problems with data input and online patient data reporting through SIHA.	TSC1	0.962	[19]
	I often feel that the SIHA usage problems I deal with are : difficulties in updating data, data recording, and reporting.	TSC2	0.938	
Technology Characteristics (TCH) CA, CR, AVE = 0.725, 0.862, 0.760	I find that using SIHA online presents no difficulties and makes it easy for healthcare workers to input data that needs to be updated.	TCH 2	0.962	[19]
	I feel that using SIHA online improves time efficiency in managing data without having to go through complicated manual processes.	TCH 3	0.771	
Individual Characteristics (IC) CA, CR, AVE = 0.758, 0.891, 0.803	I feel the work experience in the healthcare field with SIHA has a positive impact on the continuous monitoring of patients.	IC2	0.919	[14]
	I feel that the personality of healthcare workers influences their understanding of the system and their attitude towards SIHA use.	IC3	0.873	
Data quality (DQ) CA, CR, AVE = 0.739, 0.884, 0.793	I feel the data used in SIHA is indeed accurate and provides information that has been rigorously validated.	DQ1	0.902	[11]
	I feel that using SIHA provides me with administrative data in carrying out my HIV/AIDS reporting tasks.	DQ2	0.879	
Task Technology Fit (TTF) CA, CR, AVE = 0.909, 0.943, 0.846	I feel that SIHA supports my tasks as a healthcare worker in monitoring treatment, report generation, and patient data recording.	TTF1	0.92	[14]
	I feel that SIHA provides me with the suitability for the demands of the tasks I have to complete as a healthcare worker.	TTF2	0.888	
	I feel that SIHA assists healthcare workers in making better decisions regarding the care needed by patients.	TTF3	0.929	
Impact of Performance (IOP) CA, CR, AVE = 0.901, 0.938, 0.874	I feel that SIHA can assist me in making decisions such as: prioritizing areas, allocating resources, evaluating programs, identifying trends, and developing faster and more accurate education strategies.	IOP1	0.899	

Variable	Statement Item	Code	LF	Reference
	I feel that SIHA can improve efficiency in all aspects such as: continuous monitoring, automated analysis, precise allocation, efficient strategy evaluation, and data integration.	IOP2	0.903	[11]
	I feel that SIHA can improve data accessibility such as: data can be accessed from anywhere, easy-to-understand data visualization, access rights according to user roles, and ,mobile integration.	IOP3	0.938	

Discriminant validity can be evaluated using an alternative method, namely HTMT [21]. This method uses the multitrait-multimethod matrix as the basis for measurement. The HTMT value must be less than 0.90 to ensure discriminant validity between two reflective constructs [22].

Based on Table III, all values are below 0.90, which means that discriminant validity has been met. This indicates that each construct in the reflective model has a clear distinction from one another and does not experience conceptual overlap issues.

TABLE III. RESULTS OF DISCRIMINANT VALIDITY TEST HTMT

	DQ	IC	IOP	TCH	TSC	TTF
DQ						
IC	0.304					
IOP	0.484	0.339				
TCH	0.249	0.103	0.247			
TSC	0.380	0.220	0.144	0.310		
TTF	0.542	0.276	0.422	0.213	0.237	

B. Structural Model Evaluation (Inner Model)

Structural model analysis is crucial for understanding how variables interact and influence each other within a model. In this study, the evaluation of the magnitude and significance of these relationships will be conducted to test the proposed hypotheses. Several tests are employed in the structural model analysis, including the Variance Inflation Factor (VIF) test and the Determinant Coefficient (R-Square) test, to assess the validity of the hypotheses [23].

The VIF test is used to assess collinearity among constructs in the research model. This test ensures that the constructs do not exhibit excessive interdependence, which could affect the model's validity. The VIF value for a construct should be greater than or equal to 5 and less than or equal to 0.2, indicating that the construct has collinearity issues [21].

In SEM-PLS analysis, a VIF value above 5 indicates a multicollinearity issue, while a value below 5 suggests no multicollinearity problem [23] [24]. In this study, based on the test results in Table 4 , all constructs have VIF values ranging from 1.152 to 1.163. These values are well below the threshold of 5, indicating that there is no multicollinearity issue in this research model.

The hypothesis testing in this study aims to determine whether the proposed hypotheses can be accepted or rejected . Hypothesis testing is conducted by comparing the obtained t-statistic and p-value. According to [25], the t-statistic value must be >1.96 with a significance level of 5%, and the p-value must be <0.05 for the hypothesis to be accepted.

TABLE IV. MULTICOLLINEARITY TEST RESULTS VARIANCE INFLATION FACTOR (VIF)

	DQ	IC	IOP	TCH	TSC	TTF
DQ			1.249			1.152
IC						1.071
IOP						
TCH						1.068
TSC						1.163
TTF			1.249			

TABLE V. HYPOTHESIS TESTING RESULTS

Hypothesis	Variable	T-Statistic	P-Values	Description
H1	TSC → TTF	0.542	0.588	Rejected
H2	TCH → TTF	1.244	0.214	Rejected
H3	IC → TTF	1.613	0.107	Rejected
H4	DQ → TTF	4.008	0.000	Accepted
H5	DQ → IOP	2.494	0.013	Accepted
H6	TTF → IOP	2.878	0.004	Accepted

Furthermore, the R-square value is used to assess the impact of independent latent variables on the dependent variable. The R-square value ranges from

0 to 1, with values closer to 1 indicating a stronger model fit. An R-square value of 0.75 is considered strong, 0.50 is considered moderate, and 0.25 is considered weak [23].

TABLE VI. R-SQUARE TEST RESULTS

Variable	R-square	Category
IOP	0.212	Weak
TTF	0.234	Weak

Based on the analysis results, data quality is a significant factor influencing both Task Technology Fit and Impact of Performance in the SIHA system. However, the R-Square values obtained for these two variables are still relatively weak (0.234 and 0.212, respectively). This indicates that the current research model can only explain approximately 21–23% of the variation in Task Technology Fit and Impact of Performance , while the remaining variability is likely influenced by other factors not yet included in the model. Therefore, although the factors identified in this study are important, further research is needed to explore other factors that could contribute to improving the effectiveness of the SIHA system.

Based on Table VI, six hypotheses were proposed and tested. Three hypotheses have a T-statistic > 1.96 and a p-value < 0.05, indicating that these hypotheses are accepted. Meanwhile, three hypotheses have a T-statistic < 1.96 and a p-value > 0.05, meaning that these hypotheses are rejected as they do not meet the hypothesis criteria.

H1 is rejected because it has a T-Statistic value < 1.96 (0.542) and a P-Value > 0.05 (0.588). This indicates that Task Characteristics does not have a significant impact on the Task Technology Fit of SIHA. In other words, the complexity or simplicity of a task does not directly influence whether the technology used is suitable for supporting that task.

Furthermore, H2 is rejected because it has a T-Statistic value < 1.96 (1.244) and a P-Value > 0.05 (0.214). Technology Characteristics, such as ease of use or feature sophistication, do not have a significant impact on the Task Technology Fit of SIHA. This indicates that the compatibility of SIHA with tasks is not solely dependent on its features.

Furthermore, H3 is rejected because it has a T-Statistic value < 1.96 (1.613) and a P-Value > 0.05 (0.107). Individual Characteristics, such as user experience or skills, do not have a significant impact on the Task Technology Fit of SIHA. This indicates that individual factors are not the primary determinants of technology-task compatibility in the use of SIHA.

Furthermore, H4 is accepted because it has a T-Statistic value > 1.96 (4.008) and a P-Value < 0.05 (0.000). Data Quality has a significant impact on the Task Technology Fit of SIHA. This indicates that accurate, complete, and reliable data enhance the compatibility of technology with the tasks performed. It highlights the importance of data quality in supporting the effective use of SIHA.

Furthermore, H5 is accepted because it has a T-Statistic value > 1.96 (2.494) and a P-Value < 0.05 (0.013). Data Quality has a significant impact on the Impact of Performance. This indicates that high-quality data enhance the effectiveness and efficiency of the system in supporting decision-making and monitoring the spread of HIV/AIDS.

And finally, H6 is accepted because it has a T-Statistic value > 1.96 (2.878) and a P-Value < 0.05 (0.004). Task Technology Fit of SIHA has a significant impact on the Impact of Performance. This indicates that the better the technology aligns with task requirements, the better the performance of SIHA. This finding is consistent with previous research by [13]. These findings confirm that Task Technology Fit serves as a mediating variable that explains how data quality contributes to performance impact through better alignment between technology and user tasks in the use of SIHA.

Based on the hypothesis testing results, it was found that Data Quality has a significant impact on Task Technology Fit (TTF) and Impact of Performance (IOP). Meanwhile, Task Characteristics, Technology Characteristics, and Individual Characteristics do not have a significant impact on TTF. This indicates that the effectiveness of SIHA usage is more determined by the quality of available data rather than other factors such as task, technology, or user characteristics.

This study aims to analyze the relationship between data quality, technology fit, and their impact on the effectiveness of monitoring and controlling the spread of HIV/AIDS in West Papua Province. The research employs a quantitative method using the Task Technology Fit (TTF) approach.

The findings of this study prove that Data Quality in SIHA in West Papua significantly influences Task Technology Fit and its Impact on Performance. The test results show that the R-Square value for Impact on Performance (IOP) is 0.212 and for Task Technology Fit (TTF) is 0.234. This means that the model explains approximately 21.2% of the variability in IOP and 23.4% of the variability in TTF. This indicates that other factors outside the model also contribute to the effectiveness of SIHA in West Papua. Based on hypothesis testing results, it was found that Data Quality has a significant influence on Task Technology Fit (TTF) and Impact on Performance (IOP). Additionally, Task Technology Fit also significantly influences Impact on Performance, meaning that the better the alignment between technology and user task requirements, the better the performance in monitoring the spread of HIV/AIDS in West Papua.

However, this study also found that three factors do not have a significant influence on Task Technology Fit in West Papua. First, Task Characteristics do not have a significant effect

because the tasks performed by respondents in this study are routine or standardized, making technology use not the primary factor in determining the fit between tasks and technology. Second, Technology Characteristics do not have a significant effect, possibly because the technology used in this study is already considered standard and no longer serves as a differentiating factor in supporting task execution. If the available technology has common features or is similar to previous systems, users may not perceive a significant improvement in fit. Third, Individual Characteristics do not have a significant effect due to uniformity in the level of skills, experience, or user perceptions of technology within the study environment in West Papua. If most individuals have relatively similar abilities in using technology, differences in individual characteristics do not significantly impact the fit between tasks and technology.

These findings indicate that the success of SIHA in supporting HIV/AIDS monitoring in West Papua is primarily determined by data quality rather than technical factors or individual user characteristics. Overall, this study confirms that improving data quality is the key factor in the effectiveness of SIHA in West Papua, compared to technical or user-related factors. Therefore, efforts to develop SIHA in West Papua should focus on data validation, healthcare worker training, and strengthening technological infrastructure to ensure accessibility and accuracy of information in HIV/AIDS monitoring.

C. Theoretical Implications

This study provides a significant contribution to the development of the Task Technology Fit (TTF) theory, particularly within the context of health information systems in regions with geographical challenges, such as West Papua. The findings demonstrate that among the four independent variables tested (task characteristics, technology characteristics, individual characteristics, and data quality), only data quality has a significant effect on both task-technology fit and performance impact. This highlights the dominant role of data quality in enhancing system effectiveness, extending the theoretical understanding of the TTF model.

The first theoretical implication is the importance of integrating data quality dimensions as a central component within the TTF model. While prior research has primarily focused on the alignment between task, technology, and user characteristics, this study reveals that, in the context of health information systems, data accuracy, completeness, and validity play a more critical role in determining system performance. Therefore, the TTF model may need to be revised or extended to account for external factors such as data quality that significantly influence system effectiveness.

The second implication is the unexpected insignificance of task characteristics, technology characteristics, and individual characteristics in shaping TTF. This contradicts core assumptions of the original TTF theory. It suggests that in certain settings such as standardized systems with relatively homogeneous users these variables may lose their predictive power. Thus, the application of TTF should be context-sensitive, with awareness of contextual moderators that may reduce the relevance of traditional constructs.

The third implication is the need for a more holistic approach in evaluating the effectiveness of health information systems. This study shows that even when technology and user characteristics are neutral or non-significant, high-quality data can still drive performance improvements through its influence on TTF. This strengthens the argument that system evaluations should go beyond the fit between task and technology and focus on the quality of the information generated by these systems.

Overall, the findings of this research suggest that the TTF model should be modified or expanded to more accurately reflect the realities of health information system implementation especially in under-resourced or geographically challenging settings like West Papua. Incorporating external variables such as data quality into the model enhances its explanatory power and relevance for real-world applications in healthcare monitoring and decision-making.

D. Practical Implications

Based on the findings of this study, several practical implications can be applied to enhance the effectiveness of monitoring the spread of HIV/AIDS in West Papua. This research indicates that Data Quality (DQ) and Task Technology Fit (TTF) have a significant impact on the Impact of Performance (IOP) in the use of SIHA. Therefore, more accurate, complete, and reliable data enable healthcare professionals to conduct more comprehensive analyses, make faster decisions, and improve efficiency in reporting processes and public health interventions. The timeliness of data provision also supports a more proactive response to disease transmission trends. Thus, data quality serves as a crucial foundation for effective HIV/AIDS prevention and control efforts.

The findings of this study align with the three accepted hypotheses: Data Quality, Task Technology Fit (TTF), and Impact of Performance. The alignment between technology and the tasks performed by healthcare workers has been proven to strengthen the relationship between data quality and monitoring performance. With a system that meets user needs, the processes of data recording, processing, and reporting can be carried out more efficiently and accurately. Therefore, improving data quality should

be a top priority in the future development of SIHA. This includes implementing stricter data validation, enhancing system integration among relevant institutions, and adopting user-friendly technology to facilitate use by healthcare workers in the field.

Additionally, these findings provide valuable insights for policymakers and stakeholders in managing health information systems. The government, healthcare institutions, and non-governmental organizations can utilize this research to develop sustainable system improvement strategies. These strategies include enhancing capacity and training for healthcare workers in using SIHA, strengthening technological infrastructure to ensure optimal data availability and accessibility, and refining standard operating procedures (SOPs) for HIV/AIDS data recording and reporting. Furthermore, cross-sector collaboration must be reinforced to facilitate smooth data exchange and coordination in HIV/AIDS prevention and control efforts.

In the long term, the development of SIHA, supported by high-quality data, will not only enhance monitoring effectiveness but also strengthen the overall healthcare system in West Papua. These efforts are expected to accelerate HIV/AIDS case management, reduce new infection rates, and improve the quality of life for affected communities. Therefore, the synergy between technology, reliable data, and appropriate policy support is the key to establishing an effective and sustainable health information system.

E. Limitation

This study has several limitations. One of the main limitations is the low R-Square values for the for the Impact of Performance and 0.234 for Task Technology Fit. This indicates that there are still other factors outside the model that contribute to the performance of SIHA, which have not been included in this study. To improve the R-Square value, it is necessary to add external variables, use a combinative research method, and expand the scope of the study. With these improvements, it is expected that the research model will be stronger in explaining the factors influencing the effectiveness of SIHA, thereby providing more accurate recommendations for the development of health information systems in the future.

Second, this study was conducted only in the West Papua region, so the results may not be generalizable to other areas with different characteristics. Third, this study employs a quantitative method with a survey approach, which does not explore in depth the qualitative factors that may influence the effectiveness of SIHA. Fourth, this study focuses solely on the aspects of Data Quality, Task Technology Fit, and Impact of Performance, without considering other external variables that may contribute to the successful implementation of the health information system.

F. Future Research Directions

For future research, several aspects can be explored further. First, a qualitative approach should be employed to gain a deeper understanding of the challenges and obstacles in the implementation of SIHA. Second, a broader study with a larger geographical scope will provide a more comprehensive understanding of SIHA's effectiveness across different regions. Third, future research can integrate other factors such as health policies, management support, and technological infrastructure readiness to holistically understand the success factors of SIHA in HIV/AIDS monitoring. Thus, future research findings can provide more accurate recommendations for improving health information systems in Indonesia.

IV. CONCLUSION

This study indicates that data quality is a significant factor influencing the fit between tasks and technology, as well as the performance impact in the implementation of SIHA for HIV/AIDS monitoring in West Papua. Other factors such as task characteristics, technological characteristics, and individual characteristics were not found to be significant in this model.

However, it should be noted that the R-squared values for Task Technology Fit (0.234) and Performance Impact (0.212) suggest that this model only explains approximately 21–23% of the variance in the dependent variables. This implies that a large portion of the variation—more than 77%—is likely influenced by other factors not yet included in the current model.

Therefore, although data quality is proven to have a significant influence, these findings need to be interpreted with caution. Further research is needed to identify additional factors that may affect the effectiveness of SIHA in HIV/AIDS monitoring. Efforts to improve data accuracy, completeness, and reliability remain crucial, but future development also needs to consider other factors such as organizational aspects, environmental factors, or policies to maximize the impact of health information systems in challenging regions.

REFERENCES

[1] S. Mukarromah and M. Azinar, “Penghambat Kepatuhan Terapi Antiretroviral pada Orang dengan HIV/AIDS (Studi Kasus pada Odha Loss To Follow Up Therapy),” Indones. J. Public Heal. Nutr., vol. 1, no. 1, pp. 101–113, 2021. [Online]. Available: http://journal.unnes.ac.id/sju/inde

x.php/IJPHN

[2] V. Silalahi, "Pendidikan Kesehatan Tentang Perilaku Seksual Berisiko HIV / AIDS Pada Remaja Health Education On HIV / AIDS Sexual Risk Behavior Among Teenagers," J. Ilm. Tatengkorang, vol. 5, no. 2, pp. 50–55, 2021.

[3] R. D. Govender, M. J. Hashim, M. A. Khan, H. Mustafa, and G. Khan, "Global Epidemiology of HIV/AIDS: A Resurgence in North America and Europe," *J. Epidemiol. Glob. Health*, vol. 11, no. 3, pp. 296–301, 2021, doi: 10.2991/jegh.k.210621.001.

[4] Kementrian Kesehatan Republik Indonesia, "Laporan Penilaian Risiko Cepat MPOX MPOX di Indonesia," 2023.

[5] orideknews.com, "6.000 Orang di Papua Barat Positif HIV, Mayoritas Usia Produktif," Accessed: May 07, 2025. [Online]. Available: https://orideknews.com/2025/05/07/6-000-orang-di-papua-barat-positif-hiv-mayoritas-usia-produktif/

[6] Permenkes No. 24, "Peraturan Menteri Kesehatan RI No 24 tahun 2022 tentang Rekam Medis," *Peratur. Menteri Kesehat. Republik Indones. Nomor 24 Tahun 2022*, vol. 151, no. 2, pp. 1–19, 2022.

[7] D. K. Artati *et al.*, "Tingkat Penerimaan Aplikasi SIHA 2 . 1 pada Petugas Kesehatan di Provinsi Jawa Tengah," vol. 5, no. 3, pp. 186–193, 2024, doi: 10.25047/j-remi.v5i3.4620.

[8] B. K. Ngugi, B. Harrington, E. N. Porcher, and R. G. Wamai, "Data quality shortcomings with the US HIV/AIDS surveillance system," *Health Informatics J.*, vol. 25, no. 2, pp. 304–314, 2019, doi: 10.1177/1460458217706183.

[9] D. L. Goodhue, "Development and measurement validity of a task-technology fit instrument for user evaluations of information systems," *Decis. Sci.*, vol. 29, no. 1, pp. 105–138, 1998, doi: 10.1111/j.1540-5915.1998.tb01346.x.

[10] J. Gebauer, M. J. Shaw, and M. L. Gribbins, "Task-Technology Fit for Mobile Information Systems," 2020. [Online]. Available: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=43a388e74e1de114d6ac249a5ffd19dffc2423ff

[11] K. C. Maulina and E. S. Astuti, "Pengaruh Karakteristik Tugas, Teknologi Informasi Dan Individu Terhadap Task-Technology Fit (TTF), Utilisasi Dan Kinerja," JISIP J. Ilmu Sos. Ilmu Polit., vol. 4, no. 1, pp. 108–119, 2015.

[12] D. L. Goodhue and R. L. Thompson, "Task-Technology Fit and Individual Performance," Inf. Syst. Res., vol. 19, no. 2, pp. 213–236, 1995. [Online]. Available: http://www.jstor.org/stable/24969689

[13] M. J. Wills, O. F. El-Gayar, and A. V. Deokar, "Evaluating task-technology fit and user performance for an electronic health record system," *15th Am. Conf. Inf. Syst. 2009, AMCIS 2009*, vol. 4, pp. 2708–2717, 2009.

[14] P. E. Setiawati, D. I. Inan, R. N. Wurahrah, R. Juita, and M. Sanglise, "How the perceived enjoyment effect m-payment adoption in west papua province: delone and mclean information systems success model," vol. 9, no. 3, pp. 1494–1505, 2024.

[15] M. Arifudin, M. H. Arfah, M. Riskayanti, S. F. Husaimah, and L. Rahim, "Makalah Mata Kuliah Statistika," no. June, 2022.

[16] J. D'Ambra, C. S. Wilson, and S. Akter, "Application of the task-technology fit model to structure and evaluate the adoption of E-books by academics," *J. Am. Soc. Inf. Sci. Technol.*, vol. 64, no. 1, pp. 48–64, 2013, doi: 10.1002/asi.22757.

[17] R. A. Putri, R. A. Putra, and M. L. Dalafranka, "Analisis Penerimaan Pengguna Sistem Informasi Akademik STIQ Al-Lathifiyyah Menggunakan Task Technology Fit," *J. Comput. Inf. Syst. Ampera*, vol. 3, no. 2, pp. 111–132, 2022, doi: 10.51519/journalcisa.v3i2.177.

[18] N. Kock and P. Hadaya, "Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods," *Inf. Syst. J.*, vol. 28, no. 1, pp. 227–261, 2018, doi: 10.1111/isj.12131.

[19] Y. A. Takke, L. F. Marini, and L. Y. Baisa, "Analisis Kepuasan Pengguna Aplikasi PELNI Mobile Menggunakan Metode Technology Acceptance Model (TAM) dan End User Computing Satisfaction (EUCS)," *G-Tech J. Teknol. Terap.*, vol. 8, no. 3, pp. 1921–1930, 2024, doi: 10.33379/gtech.v8i3.4700.

[20] S. Sayyida, "Structural Equation Modeling (Sem) Dengan Smartpls Dalam Menyelesaiakan Permasalahan Di Bidang Ekonomi," *J. MISSY (Management Bus. Strateg.*, vol. 4, no. 1, pp. 6–13, 2023, doi: 10.24929/missy.v4i1.2610.

[21] D. I. Inan *et al.*, "Service quality and self-determination theory towards continuance usage intention of mobile banking," *J. Sci. Technol. Policy Manag.*, vol. 14, no. 2, pp. 303–328, 2023, doi: 10.1108/JSTPM-01-2021-0005.

[22] E. Roemer, F. Schuberth, and J. Henseler, "HTMT2–an improved criterion for assessing discriminant validity in structural equation modeling," *Ind. Manag. Data Syst.*,

vol. 121, no. 12, pp. 2637–2650, 2021, doi: 10.1108/IMDS-02-2021-0082.

[23] A. Saputra, D. I. Inan, R. Juita, M. Sanglise, and M. Indra, "Determinants of User Satisfaction with a Mobile Application Using the Extended Webqual 4 . 0 Method and Structural Equation Modelling : A Case of the PLN Mobile Application," vol. 10, no. 1, pp. 39–48, 2024.

[24] C. F. Risdiyanto, D. I. Inan, R. N. Wurarah, and O. A. Fenetiruma, "Analisis Faktor-faktor Pendukung dan Penghambat Beralih Mengadopsi Mobile Banking di Papua Barat Memanfaatkan PLS-SEM dan Perspektif Status Quo Bias," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 2, pp. 637–646, 2024, doi: 10.57152/malcom.v4i2.1289.

[25] G. Di Leo and F. Sardanelli, "Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach," *Eur. Radiol. Exp.*, vol. 4, no. 1, 2020, doi: 10.1186/s41747-020-0145-y.

ISSN 2085-4579

Examining the Role of Social Media in Shaping Undergraduate Students' Study Preferences

Fransiscus A Halim [1], Gusti Ketut Siska Sri Nandini[1], Nova Indah Lumbantoruan[1]
[1]Information Systems Study Program, The Faculty of Engineering and Informatics,
Multimedia Nusantara University, Tangerang, Indonesia. www.umn.ac.id
fransiscus.ati@lecturer.umn.ac.id, gusti.ketut@student.umn.ac.id, nova.indah@student.umn.ac.id

Accepted on April 07, 2025
Approved on June 23, 2025

***Abstract* – Social media has undeniably permeated nearly every facet of contemporary daily life, establishing itself as an indispensable tool, especially for students. These digital platforms have evolved beyond mere communication channels, becoming critical conduits through which students access a wealth of information pertinent to their academic pursuits, including educational assignments, supplementary course materials, and collaborative learning resources. The study specifically targeted a homogenous group: students enrolled in the Undergraduate Students Information Systems Study Program, all belonging to the Class of 2022. The research framework was structured around two principal variables. The independent variable was defined as "social media use," encompassing the frequency, duration, and nature of students' interactions with various social media platforms. Conversely, the dependent variable was "students' interest in learning," which was assessed through various indicators reflecting their engagement, motivation, and curiosity towards academic content. The residual analysis indicated the presence of heteroscedasticity within the residuals of the social media and interest variables. This suggests that the variance of the errors was not constant across all levels of the independent variable, a factor that would require careful consideration in subsequent statistical modeling. However, a significant positive observation was the absence of autocorrelation among the residuals of key variable pairings, specifically between social media and interest, interest and time (presumably time spent on social media or studying), and interest and productivity. This lack of autocorrelation suggests that the error terms in the model were independent, strengthening the reliability of the correlation findings.**

Keywords - Interest in Learning, social media, educational assignment, Normality Analysis, Study Preferences

I. INTRODUCTION

Platforms like Facebook, Instagram, Twitter, and TikTok serve as avenues for social interaction with friends, family, and the community. According to data from datareportal.com 2024, There were 185.3 million internet users in Indonesia at the start of 2024, when internet penetration stood at 66.5 percent. Indonesia was home to 139.0 million social media users in January 2024, equating to 49.9 percent of the total population.[1] A total of 353.3 million cellular mobile connections were active in Indonesia in early 2024, with this figure equivalent to 126.8 percent of the total population [1]. Meanwhile, data published in the ad planning tools of top social media platforms indicates that there were 126.8 million users aged 18 and above using social media in Indonesia at the start of 2024, which was equivalent to 64.8 percent of the total population aged 18 and above at that time.[1]This data indicates that teenagers are significant users of social media.[2] Students often use social media for assignments and lecture information.[3]. While social media is ubiquitous, its precise impact on educational performance remains a contentious issue in educational discourse [4]

However, the excessive and uncontrolled use of social media can negatively impact students' interest in learning. One evident manifestation is the phenomenon of "FOMO" (fear of missing out), which refers to the fear of being left out of trending activities on social media, affecting students' learning interests. The allure of social media trends may distract students from their educational work, leading them to spend excessive time browsing social media and experiencing mental fatigue. This inability to disconnect from social media is commonly known as addiction.

The following research questions:

1. How significant is the influence of learning interests among Undergraduate students in the Information Systems Study Program, Class of 2022, on their use of social media?
2. What are the positive and negative impacts of

social media use, and which effect dominates the learning interest of Undergraduate students Information Systems Study Program students

Finally, the scope of this investigation is precisely defined by its target demographic: students enrolled in the Undergraduate students Information Systems Study Program, specifically those from the 2022 academic batch. This focused approach ensures that the findings are relevant and applicable to the intended population within the established methodological framework

A. Previous Research on Social Media and Educational Performance

The study assessed the impact of social media on economic growth in a global perspective hence the use of 198 countries as sample for the period 2009 - 2017. The study utilize and adopted panel data methodologies such as panel corrected standard errors, two-stage least square and panel quantile regression methods for its regression analysis[5]

The impact of social media on educational performance has been the subject of extensive research, with results varying considerably. Several studies have reported a negative correlation, attributing it to the distractions and diminished educational performance associated with excessive social media use.[6] Conversely, other researchers have emphasized potential benefits, including enhanced learning opportunities facilitated by social media and the development of educational support networks.[7]

This study reveals a complex relationship between social media usage and academic performance among college students. While excessive use correlates with negative educational outcomes, strategic utilization of social media for educational purposes can yield positive results. The results demonstrate that while excessive engagement with social media can negatively influence academic performance, the strategic and purposeful integration of social media for educational endeavors can positively impact academic outcomes. This dual effect highlights the necessity of balanced and mindful social media engagement among students.[4]

B. Research purposes

The study investigates the relationship between student learning interest and social media usage. This research provides insights for lecturers, education administrators, and students, enabling them to understand and mitigate the negative impacts of social media on student learning interests while maximizing its positive potential [8]. Thus, more effective strategies can be developed to utilize social media positively, increase student interest in learning as a whole, and help students maximize the benefits of using social media and maintain high learning interest.

The benefits of the research study are as follows:

1. Provide a better understanding of the influence of student learning interest on the use of social media. This understanding can aid in the development of more effective strategies and policies to enhance student learning interests.
2. Assist teachers and lecturers in comprehending the optimal ways to integrate social media into learning, enabling them to maximize the benefits of using social media while maintaining high levels of learning interest.
3. This research enhances the existing body of scientific literature on social media's impact on student learning interests, providing a robust resource for subsequent studies seeking to explore this dynamic relationship.
4. Raise awareness about the significance of positive and effective use of social media in education. It also helps the general public understand the impact of social media use on student interest in learning."

II. RESEARCH METHODS

A. Social Media and Interest in Learning

Social media is a form of media that comprises three essential components: information infrastructure and tools used for producing and distributing media content. The media content itself can encompass personal messages, news, ideas, and digital cultural products. Individuals, organizations, and industries are the participants who possess and consume digital media content.[3]

Interest can be defined as a persistent inclination to focus on and remember specific activities. When someone is interested in an activity, they consistently pay attention to it with pleasure. From a psychological perspective, learning is a process of change and a method of adapting to new circumstances. It involves a transformation in behavior resulting from interactions with the environment to meet one's life needs. These changes manifest in various aspects of behavior [9]

B. Prior Research

In this study, quantitative data were collected using a convenience sampling method through the distribution of questionnaire. Random sampling was employed to ensure the research results could be used

to estimate the population [10]. The validity of each indicator was assessed through a comparison of the corrected item-total correlation [11].

The research investigates three variables, with the dependent variable being interest in learning, and the independent variables being attitude toward using the internet and subjective norms of internet use. The hypotheses formulated for this study are as follows:

H1: Attitudes toward internet use significantly correlate with student interest in learning.
H2: Subjective norms of internet use significantly impact student interest in learning.

Based on the test results, it can be analyzed that respondents' attitudes toward using the internet influence students' interest in learning. Consequently, it can be concluded that students' attitudes toward internet use can foster their interest in learning [12]."

C. *Hypothesis*

We employed an associative hypothesis for this study, which is a provisional statement or research proposition that examines the association or relationship between two variables.

H0: There is no significant influence of Student Learning Interest on the use of social media.
H1: Student Learning Interest has a significant effect on the use of social media.

D. *Secondary Hypothesis*

Minor Hypothesis 1:

H0: There is no significant influence of study time on the use of social media among Undergraduate students Information System Class of 2022.
H1: Study time significantly affects the use of social media among Undergraduate students Information System, Class of 2022.

Minor Hypothesis 2:

H0: There is no significant influence of productivity on the use of social media among Undergraduate students Information System, Class of 2022.
H1: Productivity significantly influences the use of social media among Undergraduate students Information System, Class of 2022.

To maintain focus and ensure the discussion remains pertinent to the primary research problem, this study establishes specific limitations. Firstly, the chosen research methodology is descriptive, aiming to characterize the relationships observed rather than establish causal links. Secondly, the core objective of this research is to quantify the influence of learning interest when mediated through social media platforms.

III. Research and Discussion

A. *Research Objective and Data Collection*

This study adopts a quantitative approach, utilizing a survey as the research method. It examines both independent and dependent variables [13]. Quantitative method is the collection and analysis of numerical data to answer scientific research questions. Quantitative method is used to summarize, average, find patterns, make predictions, and test causal associations as well as generalizing results to wider populations. It allows us to quantify effect sizes, determine the strength of associations, rank priorities, and weigh the strength of evidence of effectiveness.[14] The independent variable in this study is the use of social media, while the dependent variable is student interest in learning. The research design employed is a causal-comparative study known as a criterion group design. Causal-comparative research aims to investigate the presence of a causal relationship between two variables.

In this study, no treatment or control is administered to the existing variables as they have already occurred [15]. The population under investigation consists of Undergraduate students Information Systems Study Program students from the Batch of 2022. The research techniques employed include questionnaires and interviews conducted with a select group of students. The sample size represents 80% of the total Undergraduate students enrolled in the Information Systems Study Program.

B. *Normality test using Shapiro-Wilk's test*

The skew normal (SN) family of distributions includes the normal distribution as a particular case as well as a wide variety of skew densities [16]. To evaluate the normality of the residuals for each variable, this study utilized the Shapiro-Wilk test. A data is normally distributed when the significant value (p-value) > 0.05 [17].The null hypothesis (H0) of this test is that the data is normally distributed. A p-value below 0.05 led to the rejection of H0, implying that the residuals were not normally distributed. A p-value above 0.05 supported the acceptance of H0, indicating that the residuals were normally distributed. The p-values for each variable are as follows:

Social Media variable

```
        Shapiro-Wilk normality test

data:  data$medsos
W = 0.8024, p-value = 4.47e-10
```

Interest Variable

```
        Shapiro-Wilk normality test

data:  data$Minat
W = 0.75814, p-value = 2.435e-11
```

Time Variable

```
        Shapiro-Wilk normality test

data:  data$Waktu
W = 0.8185, p-value = 1.434e-09
```

Productivity Variable

```
        Shapiro-Wilk normality test

data:  data$Produktivitas
W = 0.83346, p-value = 4.502e-09
```

Fig. 1. Normality test using Shapiro-wilk

Based on the Shapiro-Wilk test results in figure 1, the null hypothesis of normality was rejected for all variables. A p-value of 0.05, being equal to the significance level of 0.05, led to the rejection of the null hypothesis. Therefore, the null hypothesis, which states that the variables are normally distributed, was rejected, indicating non-normal distribution. So, the null hypothesis is rejected, which indicates that the data does not follow a typical distribution pattern.

C. Correlation Test using Pearson's test

The correlation coefficient between the variables is generally assessed in correlation analysis. The Pearson correlation coefficient may be defined as a single value that measures the strength of the linear relationship between two variables. A positive relationship signifies that the two variables increase at the same time while a negative relationship signifies that when one increases the other decreases [18]. The Pearson correlation coefficient is a statistical method employed to accurately measure the linear correlation between two variables, thereby indicating the extent of the relationship between them [19]. This test aims to determine whether there is a linear relationship between these variables and how strong the relationship is. The Pearson correlation test was used to test the null hypothesis of no linear correlation between the variables, using p-values. The decision to reject or fail to reject the null hypothesis is based on comparing the p-value to the significance level (α), which is commonly set at 0.05.

H0 = No correlation between variables

H1 = There is a correlation between social media usage and student engagement

Interest Variable and Social Media

```
        Pearson's product-moment correlation

data:  data$Minat and data$medsos
t = 3.0075, df = 95, p-value = 0.003371
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1013651 0.4668416
sample estimates:
      cor
0.2948495
```

Interest and Time variables

```
        Pearson's product-moment correlation

data:  data$Minat and data$Waktu
t = 0.82698, df = 95, p-value = 0.4103
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.1168733  0.2792790
sample estimates:
      cor
0.08454274
```

Interest and Productivity Variables

```
        Pearson's product-moment correlation

data:  data$Minat and data$Produktivitas
t = 0.49731, df = 95, p-value = 0.6201
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.1500132  0.2478829
sample estimates:
      cor
0.05095678
```

Fig. 2. Pearson's product-moment Correlation

Based on the results of the Pearson correlation test used in figure 2, there were significant correlations between between interest and social media. Still, there is no meaningful relationship between sound and time and productivity.

D. Chi-Square Test

The Chi-Square test is a statistical method utilized to examine the relationship between two categorical variables. Chi-square is a tool to measure the level of statistical relationship between variables. In this study, the chi-square method is used to examine the extent of the relationship between features in the dataset and classes [20] Its primary objective is to determine whether a relationship or association exists between the variables under investigation. During the Chi-Square test, the observed frequency distribution is compared with the theoretically expected frequency distribution. The null hypothesis (H0) for this test assumes no relationship or association between the variables, while the alternative hypothesis (H1) posits the presence of a relationship or association.

H0: There is no difference between the variables.

H1: There is a difference between the variables.

Social Media variable

```
        Chi-squared test for given probabilities

data:  tab
X-squared = 61.144, df = 3, p-value = 3.348e-13
```

Variable Interest

```
        Chi-squared test for given probabilities

data:  tab
X-squared = 136.56, df = 4, p-value < 2.2e-16
```

Time Variable

```
        Chi-squared test for given probabilities

data:  tab
X-squared = 83.155, df = 4, p-value < 2.2e-16
```

Productivity Variable

```
        Chi-squared test for given probabilities

data:  tab
X-squared = 110.9, df = 5, p-value < 2.2e-16
```

Fig. 3. Chi-square Variable test

Based on the Chi-square test above in figure 3, all these variables do not follow the expected distribution. So, this shows a significant difference between the actual distribution of the data and the anticipated distribution of a given probability.

E. Regression Equation

A linear-regression model is used to identify the general underlying pattern connecting independent and dependent variables, prove the relationship between these variables, and predict the dependent variables for a specified value of the independent variables [21].

The social media variable exhibited a statistically significant ($p < 0.01$) positive effect on interest, with a coefficient of 0.40055. This suggests that for every one-unit increase in social media, interest increases by 0.40055 units, controlling for other variables.

The time variable has an estimated coefficient of (-0.04068) with a level of significance of 0.69660. A higher p-value indicates that the service variable does not significantly affect the dependent variable in this model. The Multiple R-squared value (0.1026) shows that 10.26% of the variability in the dependent variable is explained by social media, time, and productivity.

The F-statistic was 3.543, with a corresponding p-value of 0.01759. Since this p-value is less than the significance level ($\alpha = 0.05$), the null hypothesis is rejected. Therefore, at least one independent variable in the regression model has a statistically significant influence on the dependent variable.

F. Residual Analysis Test

1) Durbin-Watson test (DW)

Durbin-Watson test in the time-series analysis is to test if there exists the zero autocorrelation or not, and had have its own critical value table by Durbin and Watson [22]. The Durbin-Watson test is used to detect autocorrelation in the residuals of a regression model, indicating whether the residuals are independent. Where autocorrelation is defined as the relationship between observations at different times in the time series, the value of the DW ranges from 0 to 4. If the DW value is close to 2, indicating the absence of positive autocorrelation, it usually occurs in a time series that fluctuates up and down repeatedly.

Meanwhile, if the DW value is close to 4, it indicates a negative autocorrelation, which generally occurs in time series data that has a cyclical pattern.

H0 = The model's residuals are not autocorrelated
H1 = There is autocorrelation in the residual model

Interest and Social Media Variables

```
        Durbin-Watson test

data:  ha1
DW = 1.8298, p-value = 0.1844
alternative hypothesis: true autocorrelation is greater than 0
```

Interest and time variables

```
        Durbin-Watson test

data:  ha4
DW = 1.9114, p-value = 0.3173
alternative hypothesis: true autocorrelation is greater than 0
```

Interest and Productivity Variables

```
        Durbin-Watson test

data:  ha5
DW = 1.9503, p-value = 0.3936
alternative hypothesis: true autocorrelation is greater than 0
```

Fig. 4. Non-Autocorrelation test

Based on the results of the Durbin-Watson test above in figure 4, there is no significant autocorrelation among interest, social media, time, and productivity. So, this shows no significant linear dependence between observations on these variables.

Regression Equation
Interest = 2.00505 + (0.40055 * Social media) + (-0.04068) * time + (-0.10624) * Productivity

Fig. 5. Regression Equation

2) Breusch-Pagan (BP) test

The Breusch-Pagan test is a statistical test employed to identify the presence of heteroscedasticity in a linear regression model. Heteroscedasticity occurs when the variability of the residuals (errors) is not constant across the range of predictor values (independent variables). Therefore, it is important to perform the Breusch-Pagan test to ensure that the linear regression model follows the assumptions of homoscedasticity. By detecting heteroscedasticity, this test helps determine whether there are varying levels of dispersion in the residuals.

H0 = Homoscedasticity
H1 =heteroscedasticity

Interest and Social Media Variables

```
        studentized Breusch-Pagan test

data:  ha1
BP = 8.2638, df = 1, p-value = 0.004044
```

Interest and Time Variables

```
        studentized Breusch-Pagan test

data:  ha4
BP = 2.3592, df = 1, p-value = 0.1245
```

Interest and Productivity Variables

```
        studentized Breusch-Pagan test

data:  ha5
BP = 0.31602, df = 1, p-value = 0.574
```

Fig. 6. Constant Error Variance Test

Based on the results of the studentized Breusch-pagan test, there is significant heteroscedasticity between interest and social media variables. However, no considerable heteroscedasticity exists between interest and time or productivity variables.

3) Shapiro-Wilk test

The Shapiro-Wilk test is a statistical test employed to assess the normality assumption of a data sample. It tests the null hypothesis that the sample data is drawn from a normally distributed population. In the Shapiro-Wilk test, the null hypothesis (H0) assumes that the data follows a normal distribution, while the alternative hypothesis (H1) suggests otherwise. A p-value is generated by the test to assess the strength of evidence against the null hypothesis, thereby guiding the decision to either reject or fail to reject it[16].

The hypothesis test for Shapiro-Wilk test is :
H0 = Normal Distributed Data
H1 = Data is not normally distributed.

Interest and Social Media Variables

```
        Shapiro-Wilk normality test

data:  bb1
W = 0.90683, p-value = 4.056e-06
```

Interest and Time Variables

```
        Shapiro-Wilk normality test

data:  bb4
W = 0.81362, p-value = 1e-09
```

Interest and Productivity Variables

```
        Shapiro-Wilk normality test

data:  bb5
W = 0.78759, p-value = 1.615e-10
```

Fig. 7. Normality Test using Shapiro Wilk test

Based on the results of the Shapiro-Wilk Normality Test, it can be concluded that the data on these variables are not normally distributed.

4) Wilcoxon test

The data were analyzed using non-parametric test because the data collected were not normally distributed. In this study, a non-parametric Wilcoxon signed rank test was conducted [23] The Wilcoxon test or paired sign is a nonparametric statistical method for comparing two paired conditions or groups, suitable for use when the data do not meet the assumption of normality or when the data is on an ordinal scale.

H0: There is no significant difference between the two conditions/groups tested.
H1: There is a significant difference between the two conditions/groups tested.

Interest and Social Media Variables

```
        Wilcoxon signed rank test with continuity correction

data:  data$Minat and data$medsos
V = 534, p-value = 0.01017
alternative hypothesis: true location shift is not equal to 0
```

Interest and Time Variables

```
        Wilcoxon signed rank test with continuity correction

data:  data$Minat and data$Waktu
V = 842.5, p-value = 0.002366
alternative hypothesis: true location shift is not equal to 0
```

Interest and Productivity Variables

```
        Wilcoxon signed rank test with continuity correction

data:  data$Minat and data$Produktivitas
V = 632, p-value = 0.001373
alternative hypothesis: true location shift is not equal to 0
```

Fig. 8. Non-Parameter Statistic test using Wilcoxon

Based on the results of the Wilcoxon Signed Rank Test with continuity correction, there are significant differences between the variables of interest and social media, interest and time, as well as welfare and productivity. Thus, indicating a substantial relationship between the interest variable and the other two variables.

5) Multicollinearity Test

To evaluate multicollinearity, the Variance Inflation Factor (VIF) was calculated. Significant multicollinearity is indicated by high VIF values. The interpretation of VIF values is as follows:

- VIF = 1: No multicollinearity. This indicates the absence of significant correlation among the independent variables.
- VIF between 1 and 5: Low multicollinearity, there is a correlation between the independent variables, but it is still acceptable in the regression analysis.
- VIF is > 5: High multicollinearity. there is a significant correlation between the independent variables,

This can affect the regression results. A high VIF value indicates that the independent variable strongly depends on other variables.[8]

If the VIF value is more than 5 or 10, it indicates significant multicollinearity and can interfere with the interpretation of the regression results. In cases of high multicollinearity, it is essential to evaluate the independent variables involved and take appropriate action, such as removing unimportant independent variables or performing data transformations.

```
   medsos          Waktu Produktivitas
 1.540797       1.388616       1.453773
```

Fig. 9. Multicolinearity Test

In this study, three results were obtained from the variables above. It can be concluded that in this study, there was little or even almost no significant correlation between the independent variables.

IV. CONCLUSION

A. Conclusion

Based on residual and correlation tests, the analysis revealed some key findings regarding the relationships between student learning interests, social media, study time, and productivity. A significant finding was the presence of heteroscedasticity in the relationship between student learning interests and social media, indicating that social media inconsistently influences student learning interests. However, the residual analysis also confirmed no autocorrelation among the variables of interest and social media, interest and study time, and interest and productivity, meaning there's no significant linear correlation in their residuals over time.

Further analysis using the Variance Inflation Factor (VIF) showed no multicollinearity among the model's predictor variables, confirming their independence and lack of significant influence on each other. Overall, the findings suggest that the learning interests of Undergraduate students Information System, Class of 2022 significantly affect their use of social media. Conversely, student study time and productivity were found to have no substantial impact on social media usage.

B. Recommendations

This study was conducted with a sample size of 97 valid respondents. For future research, it is highly recommended to increase this sample size to ensure a more robust and representative reflection of the entire student population within the Undergraduate students Information System, Batch 2022. A larger sample will significantly enhance the statistical power, validity, and generalizability of the research findings, allowing for more confident conclusions about the broader student body.

Beyond the statistical tests utilized in this study, it is advisable for future research to incorporate a wider array of advanced analytical techniques. Methods such as multiple regression, factor analysis, path analysis, analysis of variance (ANOVA), and various non-parametric tests could provide a more comprehensive and nuanced understanding of the relationships between the variables. This study primarily focused on interest, study time, and productivity as factors influencing social media use; however, based on its findings, it is further recommended to explore and include other potential variables that may also exert an influence on social media usage, thereby enriching the research's scope and depth.

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] "Digital 2024: Indonesia — DataReportal – Global Digital Insights." Accessed: May 24, 2025. [Online]. Available: https://datareportal.com/reports/digital-2024-indonesia?rq=Indonesia

[2] Y. K. Dwivedi, *Social Media: The Good, the Bad, and the Ugly*, vol. 9844. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-45234-0.

[3] P. N. Howard and M. R. Parks, "Social Media and Political Change: Capacity, Constraint, and Consequence," *Journal of Communication*, vol. 62, no. 2, pp. 359–362, Apr. 2012, doi: 10.1111/j.1460-2466.2012.01626.x.

[4] S. A. V. Paul (Sr.) and S. A. V. Paul (Jr.), "The Impact of Social Media on Academic Performance Among College Students," *Integrated Journal for Research in Arts and Humanities*, vol. 4, no. 3, pp. 115–117, May 2024, doi: 10.55544/ijrah.4.3.19.

[5] P. Asare Vitenu-Sackey, "The Impact of Social Media on Economic Growth: Empirical Evidence of Facebook, YouTube, Twitter and Pinterest," *International Journal of Business, Economics and Management*, vol. 7, no. 4, pp. 222–238, 2020, doi: 10.18488/journal.62.2020.74.222.238.

[6] P. A. Kirschner and A. C. Karpinski, "Facebook® and academic performance," *Comput Human Behav*, vol. 26, no. 6, pp. 1237–1245, Nov. 2010, doi: 10.1016/j.chb.2010.03.024.

[7] A. O. Safori and N. A. B. A. Rahman, "Higher Education and Social Media: An Empirical Evidence from Jordan and Malaysia," *Advances in Journalism and Communication*, vol. 07, no. 04, pp. 118–128, 2019, doi: 10.4236/ajc.2019.74008.

[8] S. F. Wijaya, J. Wiratama, S. A. Sanjaya, S. Pomalingo, and F. A. Halim, "Developing a User-Centered Design Model for Digital Marketing Adoption in Village MSMEs," *Journal of Logistics, Informatics and Service Science*, vol. 11, no. 6, 2024, doi: 10.33168/JLISS.2024.0621.

[9] D. Bandhu, M. M. Mohan, N. A. P. Nittala, P. Jadhav, A. Bhadauria, and K. K. Saxena, "Theories of motivation: A comprehensive analysis of human behavior drivers," *Acta Psychol (Amst)*, vol. 244, p. 104177, Apr. 2024, doi: 10.1016/j.actpsy.2024.104177.

[10] N. Oni, S. A. Sanjaya, and F. A. Halim, "Design and Development of a Web-Based Inventory System Using Agile Methodology and Hybrid Databases for a Pharmaceutical Company," 2024. doi: 10.1109/ICITACEE62763.2024.10762763.

[11] F. A. Halim, D. I. Djong, and S. F. U. Usman, "Examining the Impact of Platform Usage on Online Purchasing in E-commerce," *IJNMT (International Journal of New Media Technology)*, vol. 10, no. 2, pp. 70–78, Jan. 2024, doi: 10.31937/ijnmt.v10i2.3211.

[12] "Social Media Use 2018: Demographics and Statistics | Pew Research Center." Accessed: May 24, 2025. [Online]. Available: https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/

[13] S. Babones, "Interpretive Quantitative Methods for the Social Sciences," *Sociology*, vol. 50, no. 3, pp. 453–469, Jun. 2016, doi: 10.1177/0038038515583637.

[14] J. Rana, P. L. Gutierrez, and J. C. Oldroyd, "Quantitative Methods," in *Global Encyclopedia of Public Administration, Public Policy, and Governance*, Cham: Springer International Publishing, 2021, pp. 1–6. doi: 10.1007/978-3-319-31816-5_460-1.

[15] P. S. Brenner, *Understanding Survey Methodology*, vol. 4. Cham: Springer International Publishing, 2021. doi: 10.1007/978-3-030-47256-6.

[16] A. Monter-Pozos and E. González-Estrada, "On testing the skew normal distribution by using Shapiro–Wilk test," *J Comput Appl Math*, vol. 440, p. 115649, Apr. 2024, doi: 10.1016/j.cam.2023.115649.

[17] N. Y. Indriyanti, S. Yamtinah, and D. Muawiyah, "An Inquiry into Students' Metacognition and Learning Achievement in a Blended Learning Design," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 21, p. 77, Nov. 2020, doi: 10.3991/ijet.v15i21.12907.

[18] L. Nakawala, A. M. Muwumba, S. P. Ogwang, W. K. Nahamya, and O. Oyesigye, "Correlational Analysis Between Students' Performance in Theory and Practical Summative Examinations," *International Journal of Vocational Education and Training Research*, vol. 11, no. 1, pp. 7–20, Jan. 2025, doi: 10.11648/j.ijvetr.20251101.12.

[19] M. F. Ikhwan, W. Mansor, Z. I. Khan, M. K. Adzhar Mahmood, A. Bujang, and K. Haddadi, "Pearson Correlation and Multiple Correlation Analyses of the Animal Fat S-Parameter," *TEM Journal*, pp. 155–160, Feb. 2024, doi: 10.18421/TEM131-15.

[20] D. Chen Sami, A. Sugiharto, and F. Jie, "Chi Square Feature Selection For Improving Sentiment Analysis of News Data Privacy Treats," *J Theor Appl Inf Technol*, vol. 102, no. 18, 2024, [Online]. Available: www.jatit.org

[21] N. Roustaei, "Application and interpretation of linear-regression analysis," *Medical hypothesis discovery and innovation in ophthalmology*, vol. 13, no. 3, pp. 151–159, Oct. 2024, doi: 10.51329/mehdiophthal1506.

[22] 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.

[23] A. L. Syifadevi and N. A. Rizal, "Performance Analysis of Indonesia's Largest Banks and China Banking Titans: Company Performance under the Wilcoxon Signed Rank Test," 2024, pp. 257–271. doi: 10.2991/978-94-6463-558-4_15.

ISSN 2085-4579

Use of the K-Medoids Algorithm for Food Clustering Using Nutritional Value and Evaluation of the Elbow Method and the Davies Bouldin Index Method

Wildani Eko Nugroho[1], Safar Dwi Kurniawan[1], Yerry Febrian Sabanise[1], Prayoga Alga Vredison[2]
[1] D3 Computer Enginering Program Study, Politeknik Harapan Bersama, Tegal, Indonesia
wild4n1@gmail.com, safar.kurniawan45@gmail.com, yerryfebrian@gmail.com
[2]Informatika, Universitas Pancasakti, Tegal, Indonesia
yogavredizon23@gmail.com

Accepted on June 07, 2025
Approved on June 23, 2025

Abstract—**The four categories of vital nutrients—calories, carbs, protein, and fat—must be present in the food that people eat on a daily basis. Humans require nutrition since it will enable them to do everyday duties and maintain their health. This study applies K-Medoids optimization to the clustering approach. The purpose of this study is to classify foods with comparable nutritional values. Foods with high, medium, and low nutritional levels are grouped into three clusters. Finding the ideal number of clusters is one of the difficulties in data clustering. For this, the Davies-Bouldin Index (DBI) and the elbow technique are frequently employed. The DBI evaluates the quality of clusters by taking into account the distance between clusters and the proximity between points inside the cluster, whereas the Elbow technique plots the intra-cluster variance against the number of clusters to assist in determining the number of clusters. Better clusters are indicated by lower DBI values. The outcomes of the K-Medoids algorithm using the Elbow approach showed a DBI of 0.631 and a cluster value (k) = 3 of 0.046. The better of these two approaches will be used for clustering. According to the comparative results of the two algorithms, the DBI technique yielded a value of 0.631 for the best cluster, while the Elbow approach yielded a number of 0.046 for the best number of clusters. The study's clustering results can be utilized to choose and consume foods that will meet nutritional needs and help delay the onset of food related disorders. For instance, if you wish to gain weight, you can choose foods in cluster 0. Cluster 2 foods can be picked if you wish to diet or lose weight, while Cluster 1 meals can serve as a benchmark if taken in excess, as this can lead to obesity.**

Index Terms—K-Means, K-Medoids, Nutrition, and Clustering

I. INTRODUCTION

One of the primary resources for human survival is food. There are many different types of food from different parts of Indonesia. For instance, the Moluccas have bagea, Makassar has palubasa, Semarang has spring rolls, Medan has bacang, and there are many more varieties. The flavors, textures, and attributes of these dishes vary [1].

A variety of vegetable and animal ingredients make up the actual food. Consequently, food is a rich source of calories, protein, fat, and carbohydrates, among other nutrients. Humans require nutrition in order to perform daily tasks and lead healthy lives [2].

A pertinent strategy in the fields of nutrition research and food technology is the use of the K-Medoids algorithm for food grouping according to nutritional value. Food clustering based on nutritional content might help people choose foods that meet their nutritional demands as the necessity of a balanced diet becomes more widely recognized [3]. This approach not only aids individuals in making informed food choices but also supports researchers and dietitians in developing tailored meal plans that cater to specific dietary needs. By leveraging the K-Medoids algorithm, we can create comprehensive food databases that enhance our understanding of nutritional relationships and promote healthier eating habits [4].

The K-Medoids algorithm has been used in a number of research studies to cluster foods according to their nutritional value. For instance, Prayoga Alga Vredizon et al.'s study divided foods into three clusters high, medium, and low based on their nutritional worth using K-Means [1]. Nevertheless, K-Means and K-Medoids were not specifically compared in this study [5].

Furthermore, a study by Tias Rahmawati et al. used K-Medoids to compare the Silhouette Coefficient and Elbow techniques in clustering Indonesian provinces according to the Human Development Index (HDI) [5]. This study demonstrates the significance of selecting the appropriate evaluation method in figuring out the

ideal number of clusters, despite the fact that food is not the main focus.

Thus, the purpose of this work is to use the K-Medoids algorithm to cluster foods according to their nutritional content and assess the clustering outcomes using the Elbow and Davies-Bouldin Index methodologies. It is intended that the findings of this study will serve as a guide for future research of a similar nature and aid in the selection of foods that meet the nutritional requirements of each individual.

II. METHODS

A. *Clustering*

Clustering is a technique used to separate a set of data into groups based on a desired match. Clustering is the technique of organizing objects or data into clusters in data mining so that the resulting data is nearly equal to the original and can be identified from items in other clusters [6].

Clustering is used to separate data into regions with similar features such that each group has unique attributes. Without the use of preset labels or rules, clustering is a directed, or "unsupervised," data mining technique that arranges and searches data according to similarities between the data [7].

.

B. *K-Medoids Algorithm*

A partition clustering technique for forming clusters from a set of n objects is the K-Medoids algorithm. With this method, a specific item from the data set serves as a cluster center or representative. This algorithm's primary idea is to reduce the disparity between the objects and the selected cluster centers and make sure that the reference points in each cluster match the data that is provided [8].

The steps in the K-Medoids algorithm method are as follows [3]:

1. Calculate how many clusters (k) need to form.
2. Choose at random the initial cluster centers (medoids) for every cluster.
3. Utilizing the Euclidean distance range calculation in Eq., assign each observation to the nearest cluster.
   $d(x,y)\sqrt{\sum_{i-1}^{\pi}(x_i - y_i)^2; i = 1,2,3,\dots n}$ (1).
4. Choose items from each cluster at random to serve as new medoids.
5. Determine the separation between the new medoids and every object in each cluster.
6. Determine the distance value between the new and old total distances to determine the total deviation (S). To create a new set of k objects as medoids, swap the cluster data objects if S is less than 0.
7. Continue Steps 4 through 6 until the medoids remain unchanged, producing a cluster and each member of the cluster.

C. *Elbow Methods*

One popular technique for figuring out how many clusters should be formed during the clustering process is the elbow approach. To maximize or ascertain the ideal number of clusters for the clustering process to be executed, the elbow approach is utilized. Cluster cohesiveness and separation are calculated using the elbow approach [9]. Cluster separation gauges how distinct or disjointed one cluster is from another, and cluster cohesiveness gauges how tightly related the data in the cluster is. The sum of squares error (SSE) can be used to quantify cluster cohesiveness and cluster separation [10].

The following formula can be used to calculate SSE :

$$SSE = \sum_{k=1}^{k} \sum_{xi} |x_i - C_k|^2 \quad (2)$$

Description :
K = Cluster in C
X_i = Distance of the data on the i-th object
Ck = The i-th cluster's center

D. *Davies Bouldin Index (DBI) Methods*

One technique to guarantee the optimal number of clusters following the clustering process is Davies Bouldin Index (DBI) [11]. The DBI technique aims to minimize the distance between objects within a cluster and maximize the difference across clusters. Additionally, the DBI measure is used to evaluate the effectiveness of various clustering techniques; values near zero imply better clusters. The DBI algorithm divides the distance between the cluster centers by the average of the compactness values for each place [12]. An outline of the procedures to determine the DBI value is provided below [13]:

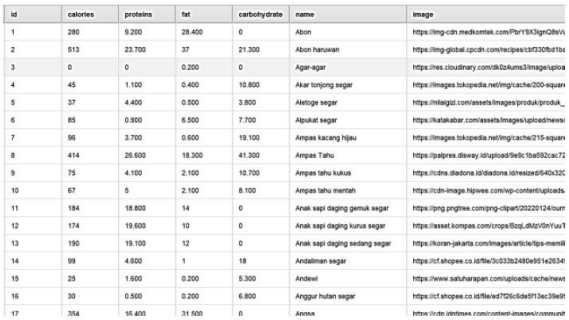
1. Finding the Sum of Squares The attachment between members of a cluster, or how similar they are to one another, is measured by the SSW; the smaller the cluster, the more similar the members are to one another. To find the matrix, cohesion, and homogeneity, SSW is computed. According to the equation formula, cohesion is the relationship between group members in a single cluster.
   $SSW = \frac{1}{m_i}\sum_{j=1}^{m_i} d(x_j \,.\, c_i)$ (3)
   Description :
   m_i = quantity of information in clusterke-i
   x = information within the cluster
   d(x,c) = data's distance from the centroid
   x_i = information within the cluster
   ci = center of cluster i centroid
2. To ascertain separation or heterogeneity, the Sum of Square Between Clusters (SSB), which is a sufficiently high distance between groups to keep them apart from one another, is calculated. The distinction between two groups is separation.
3. The purpose of calculating the ratio is to determine how well one cluster compares to

other clusters. There must be more separations than cohesiveness.

4. DBI calculation

III. Result and Discussion

Pre-processing, also known as preparation, is the first step in the research process. This involves a number of steps, such as data preprocessing, which includes attribute selection, removing duplicate data, normalizing data, and lowering the dimensions of the dataset used. After undergoing data preprocessing, the final dataset consists of food nutritional values. The K-Medoids technique is used on the final dataset in order to find the optimal clusterization performance value [14]. The K-Medoids Algorithm is used to conduct the evaluation procedure during the testing phase. A screenshot of the food nutrition dataset from the Indonesian Food and Drink Nutrition Dataset, https://www.kaggle.com/datasets/anasfikrihanif/indonesian-food-and-drink-nutrition-dataset, On November 1, 2023, Anas Fikri Hanis submitted 1346 nutrition lists for 100 grams of Indonesian food website is shown in Figure 1.



id	calories	proteins	fat	carbohydrate	name	image
1	280	9.200	28.400	0	Abon	https://img-cdn.medkomtek.com/PbrY9X3IgnQ8sVv
2	513	23.700	37	21.300	Abon haruwan	https://img-global.cpcdn.com/recipes/cbf330fbd1be
3	0	0	0.200	0	Agar-agar	https://res.cloudinary.com/dk0z4ums3/image/uploa
4	45	1.100	0.400	10.800	Akar tonjong segar	https://images.tokopedia.net/img/cache/200-square
5	37	4.400	0.500	3.800	Aletoge segar	https://nilaigizi.com/assets/images/produk/produk_
6	85	0.900	6.500	7.700	Alpukat segar	https://katakabar.com/assets/images/uploadnews/
7	96	3.700	0.600	19.100	Ampas kacang hijau	https://images.tokopedia.net/img/cache/215-square
8	414	26.600	18.300	41.300	Ampas Tahu	https://palpres.disway.id/upload/9e9c1ba592cac72
9	75	4.100	2.100	10.700	Ampas tahu kukus	https://cdns.diadona.id/diadona.id/resized/640x320
10	67	5	2.100	8.100	Ampas tahu mentah	https://cdn-image.hipwee.com/wp-content/uploads
11	184	18.800	14	0	Anak sapi daging gemuk segar	https://png.pngtree.com/png-clipart/20220124/ourn
12	174	19.600	10	0	Anak sapi daging kurus segar	https://asset.kompas.com/crops/BzqLdMzV0nYuuT
13	190	19.100	12	0	Anak sapi daging sedang segar	https://koran-jakarta.com/images/article/tips-memili
14	99	4.600	1	18	Andaliman segar	https://cf.shopee.co.id/file/3c033b2480e951e2634
15	25	1.600	0.200	5.300	Andewi	https://www.satuharapan.com/uploads/cache/news
16	30	0.500	0.200	6.800	Anggur hutan segar	https://cf.shopee.co.id/file/ed7f26c6de5f13ec39e9f
17	354	16.400	31.500	0	Angsa	https://cdn.idntimes.com/content-images/communit

Fig 1. Food Nutrition Dataset on the Kaggle Site

Several preparation stages must be completed before the K-Medoids algorithm is implemented. This is done in order to maximize the outcomes of the K-Medoids algorithm's implementation. Among the preprocessing procedures used in this study are the following ones:

a. Atribut selection

The purpose of attribute selection is to eliminate unnecessary qualities and choose those that will be utilized in the K-Medoids algorithm. Since the K-Medoids method can only accept characteristics with numeric data types, unneeded attributes are removed. Therefore, non-numerical attributes must be chosen first. This is the outcome of choosing the traits that will be utilized later in table 1.

Table 1. Result Of Attribute Selection

Name of the Attribute	Description
Calories	Calories 100 grams/food (cal)
Fat	100 grams of fat every meal (grams)
Protein	100 grams of protein each meal (grams)
Carbohidrat	100 grams of carbohydrates per meal (grams)

An overview of datasets that might be able to give an overview to assess the amount of nutrients that enter the body when ingesting typical Indonesian foods and drinks can be found in the attribute selection table above.

b. Remove Duplicate

By comparing each piece of data one at a time, the Remove Duplicates function eliminates duplicates. One data set from the comparison results will remain after all data with the same value are eliminated one at a time. Since the nutritional value of each product is typically the same, this research removes redundant data to increase the efficiency of data processing.

c. Outlier Detection

Finding data that behaves differently from other typical data requires outlier detection, which will have an impact on the K-Medoids algorithm's data processing procedure. Distance-based outlier identification is the outlier identification method employed in this study. Two points are determined by selecting a distance-based outlier, which is subsequently verified. It will be regarded as an outlier if the neighbor points are near together and then separated. Once this study has been tested, the parameters' findings are derived utilizing the distances of ten neighbors, nine outliers, and the Euclidean distance.

d. Normalization of Data

One of the data normalizing techniques employed in this study is min-max normalization. The original data values are transformed into various values between 0 and 1 using this min-max normalization. Forming data into numbers with the same range is the goal of the normalization process, which speeds up data processing.

e. The reduction of dimensions

Principal Component Analysis (PCA) is a technique used to minimize the dimensions of data containing numerous variables. PCA shrinks the data's dimensionality without appreciably lowering its information content. Based on the data, PCA creates a new set of dimensions that are rated. The principal components of PCA are derived from the breakdown of the covariance matrix's eigenvalues and eigenvectors. Two components and a fixed number of dimensions are the parameters utilized for PCA implementation. Two traits remained after the previous four were eliminate

1. K-Medoids Algorithm

A number of tasks or procedures need to be completed in experiments utilizing the K-Medoids algorithm, including dimension reduction, outlier detection, attribute selection, data standardization, and the elimination of duplicate data. The Elbow method and the Davies-Bouldin Index method are the two techniques used to test the K-Medoids algorithm. A perspective of the K-Medoids experiment is shown in Figure 2.

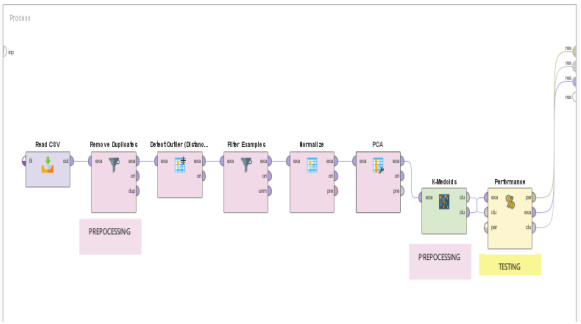


Fig 2. Algorithm Testing Scenario for K-Medoids

The K-Medoids method is tested in the image above. Both the DBI approach and the elbow method are employed in some of these exams. The K-Medoids algorithm is tested to determine its accuracy and success rate. The DBI approach and the elbow method are two methods that can be used to determine the number of clusters.

A. Elbow Method

In order to determine the number of clusters in the K-Medoids algorithm, the foundation of this phase is to establish an ideal K value. Each value in the Cluster Sum of Squares (WCSS) will be computed using the Elbow technique, which runs cluster values from 2 to n. The WCSS represents the sum of squares between each cluster's centroid value and average value [15]. The average distance of the sum of squares between the cluster center values is the sort of average calculation that is computed. Next, using a chartline, choose the spot that creates the most elbow. The outcomes of applying the Elbow approach to the K-Medoids algorithm, as displayed in Table 2, are as follow. Within-cluster sum of squares, K value, and average Figure 3. Graph Calculating How Many Clusters the K-Medoids Algorithm Needs

Table 2. Avg. Within-Cluster Sum Of Square K-Medoids

K	WCSS
2	0,091
3	0,046
4	0,036
5	0,027
6	0,023
7	0,019

The results of the Avg. Within-cluster Sum of Square K-Medoids, which represents the sum of these squared distances for every data point within each cluster, are shown in the table above. In K-Medoids, the ideal number of clusters is found using WCSS. Data within the cluster is more homogeneous, and the cluster is more compact when the WCSS is lower. Three to seven clusters are tested in the above table to determine the ideal number of clusters. Figure 3 shows the Elbow Method cluster experiments in Figures 3-7.

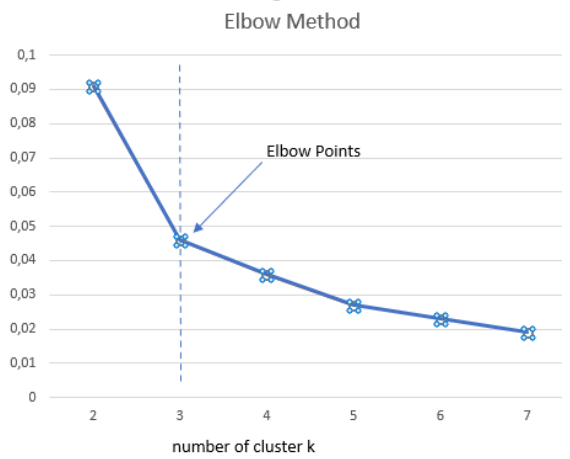


Fig 3. Calculating the K-Medoids Algorithm's Cluster Count

Figure 2 shows the outcome of an experiment to find the ideal number of clusters based on the K-Medoids algorithm's SSE (Sum of Square Error) value using the Elbow approach, which has significantly dropped in this instance. The cluster's quality decreases with increasing SSE value and vice versa. The cluster quality improves with a reduced SSE value. As can be seen in the above image, the SSE value is at its peak when there are two clusters (k = 2), and it dramatically decreases but forms an elbow when there are three clusters (k = 3). The SSE value dropped once more when there were four clusters (k = 4), and so on until there were seven clusters (k = 7). The elbow is found at the number of clusters k = 4 by calculating the average value. This is because the graph shows that the number of clusters that form an elbow is clearly visible when the number of clusters k = 3, and the number of clusters k = 4 to k = 7 is seen to start decreasing as well. sum of squares within the cluster of 0.046.

B. Davies Bouldin Index (DBI) Method

A cluster validation that incorporates separation and cohesion data is the Davies-Bouldin Index (DBI). Cohesion is the degree of similarity between the data and the cluster centroids, whereas separation is the distance between the cluster centroids. The quantity and proximity of the clustered data determine the quality of the clustering results. The cluster outcomes are better when the DBI value is smaller (non-negative => 0) [16]. The outcomes of applying the Davies-Bouldin Index approach to the K-Medoids algorithm, as displayed in table 3. are as follows.

Table 3. Value Of The Davies Bouldin Index K-Medoids

K	DBI
2	0,810
3	0,631
4	0,996
5	1,009
6	0,996
7	1,144

According to the above table 3., cluster k = 3 has the minimum DBI value, which is 0.631. This demonstrates that the outcomes match expectations.

The optimal number of clusters is three, according to the Elbow and Davies-Bouldin Index techniques that have been used. Table 4 below displays these findings:

Table 4. Value Of Wcss K-Medoids

Cluster	Avg. Within Centroid Distance
0	0,091
1	0,046
2	0,036

As can be seen from the above table, the cluster's data are reasonably close to one another, as indicated by the Average Within Centroid Distance value being near zero. The scatter/bubble in the image can be used to help visualize the image.

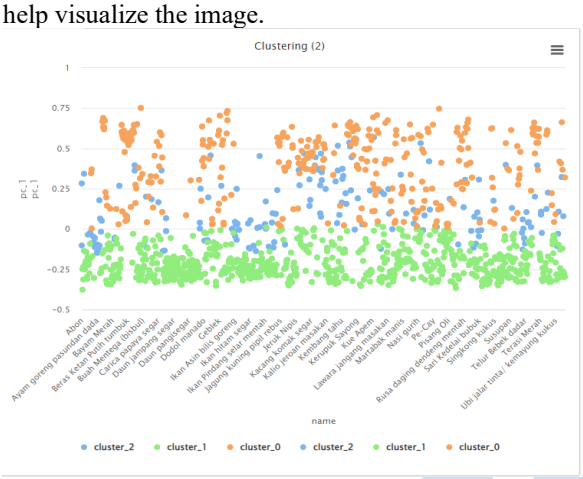


Fig 4. K-Medoids Algorithm Scatter/Bubble Outcome with Three Clusters

Three clusters were created by testing the K-medoids technique, as seen in the image above. Cluster 0 is represented by the color orange, Cluster 1 by the color green, and Cluster 2 by the color blue. Table 6 below displays the detailed clusterization results. However, Table 5 displays the quantity of data for every cluster.

Table 5. The Quantity Of Information For Every K-Medoids Cluster

Cluster	Jumlah Data
0	347
1	744
2	175

The best or most optimal number of clusters is at number three, according to the K-Medoids algorithm's implementation with multiple preprocessing steps, including data selection, duplicate removal, data normalization, and dimensional reduction, as well as cluster optimization using the Elbow Method and Davies-Bouldin Index. Clusters 0, 1, and 2 data are acquired by evaluating K-Medoids using these two techniques. There are 347 food data points in cluster 0, 744 data points in cluster 1, and 175 food data points in cluster 2. Table 5 illustrates this, and Table 6 provides the specific findings for the list of items that fall into each cluster.

Table 6. Some Food Information From The Clustering Findings Of K-Medoids

K	List Of Food
0	Asam payak segar, Bagea kelapa asin, Bagea kelapa manis, Bagea kenari asin, Bagea kenari manis, Bagea kw 1, Bagea kw 2, Bakpia kue, Bakwan, Bantal, Batatas tali ubi rebus, Beef burger, Belitung talas kukus, Beras ganyong, Beras Giling, Beras Giling masak (nasi), Beras ladang mentah, Beras Menir, Beras menir mentah, Beras Merah tumbuk, Bihun, Bihun goreng instan, Bihun Jagung mentah, Bihun mentah, Biji nangka/Biji salak, Bika ambon, Bika Ambon, Biskuit, Bonggol Pisang kering, Brem, Brondong, Buah kom segar, Bubur sagu, Bulung Jajan, Bulung Sangu, Bungkil Biji Karet, Bungkil Kelapa, Burung sarang segar, Cabai merah kering, Cake tape, Cantel mentah, Cengkeh kering, Centel, Ceriping getuk singkong, Coklat bubuk, Coklat Manis batang, Coklat Susu batang, Combro, Daun ndusuk segar, Daun salam bubuk, Deblo, Djibokum masakan, Dodol, Dodol bali, Dodol BanjarmasinEmping komak, Enting-enting gepuk hello kity, Es Krim (Coconut milk), Es Mambo, Gadeng/Gadung kering, Gaplek, Gatep segar, Gatot, Geblek, Gelatine, Gemblong, Gula putih, Gurandil, Havermout, Intip goreng, Jagung grontol, Jagung Kuning giling, Jagung Kuning pipil baru, Jagung uning pipil kering mentah, Jagung pipil var. harapan kering, Jagung pipil var. metro kering, Jagung Rebus, Jagung titi, Jali, Jam selai, Jampang huma mentah, Jamur encik, Jamur kuping kering, Japilus, Jawawut, Jenang, Jengkol segar, Kabuto, Kacang Andong, Kacang Arab, Kacang babi kering, Kacang Bogor goreng, Kacang tunis kering, Kacang uci kering, Kacang urei kering, Kambose, Katul Beras, Katul Jagung, Kecimpring singkong goreng, Kelepon, Kentang Hitam, Keremes, Keripik gadung, Kerupuk Aci, Kerupuk cumi goreng, Kerupuk Ikan berpati, Kerupuk kemplang (ikan) mentah, Kerupuk kemplang goreng, Ketumbar, Ketupat ketan, Kopi bagian yang larut, Kopi bubuk instant, Koro Benguk biji, Koro Roway biji, Koro Wedus biji, Koya, Koya mirasa, Kranji segar, Kue ali, Kue Apem, Kue bangket, Lopis, Lupis ketan, Madu, Maizena tepung, Makaroni, Markisa squash, Martabak manis, Martabak mesir, Martabak Telur, Masekat, Melase, Merica, Mie Bendo, Mie Goreng, Mie kering, Mie Sagu, Misoa, Misro, Nangka biji, Nasi, Nasi beras merah, Nasi Goreng, Nasi jagung, Nasi minyak, Nopia special, Oncom ampas kacang hijau, Onde-onde, Oyek, Oyek (dari singkong), Pastel, Pati Singkong (tapioka), Pempek belida, Pempek kulit, Pempek tenggiri, Pepaya lodeh, Permen, Petis Udang, Petis udang kering, Petis udang pasta, Pisang ketip segar, Pisang Mas, Pisang Oli, Pulut, Putu Mayang, Rarawuan, Rasbi (Beras Ubi), Rasi (Beras Singkong), Rebung laut mentah, Rempeyek kacang uci, Renggi goreng, Risoles, Roti boong, Roti Gambang, Roti Putih, Roti warna sawo matang, Sagu Ambon, Sale pisang siam. Santan (kelapa saja), Sapi abon, Sarimuka, Semur Jengkol, Sente talas segar, Serbuk Coklat, Serimping talas kebumen, Setrup sirup, Singkong Goreng, Singkong kukus, Singkong tape, Srikaya, Susu asam untuk bayi bubuk, Susu Kental Manis, Susu skim bubuk, Suwir-suwir, Talas pontianak segar, Tapai singkong, Tauco cap DAS cake, Tauji cap singa, Teh, Teh hijau daun kering, Teh melati daun kering, Tekokak kering, Tempe Sayur, Wajit camilan, Widaran, Wingko babat, Yangko
1	Bekasam, Bekasang, Belibis daging segar, Belimbing, Belut segar, Belutlaut segar, Bengkuang, Bentul (Komba) talas segar, Bentul talas kukus, Betok wadi masakan, Bihun Goreng, Bir (4% alkohol), Bit, Biwah segar, Buah ruruhi segar, Buah tuppa segar, Bubur tinotuan (Manado), Bulgogi masakan, Buncis, Buncis asam, Buncis rebus, Bunga pepaya segar, Bunga turi segar, Buntil, Buntil daun talas, Buras, Cabai gembor merah segar, Cakalang asar (asap papua), Cammetutu,

K	List Of Food
	Cap cai sayur, Carica papaya segar, Cempedak, Dideh darah ayam, Dideh darah sapi, Dodonkol, Domba daging kurus segar, Domba ginjal segar, Duku, Durian, Duwet segar, Eceng Gondok, Embacang, Encung asam segar, Erbis, Es krim, Es Sirup, Fillet o-fish, Gado-gado, Gadung kukus, Gadung mentah, Gadung ubi kukus, Gadung ubi segar, Gambas (Oyong), Gambas lodeh, Gandaria, Ganyong kukus, Ganyong mentah, Ganyong segar, Gembili, Gembili ubi segar, Genjer segar, Gete kuah asam masakan, Ginjal Babi, Ginjal Domba, Ginjal Sapi, Gudeg, Gudeg sayur, Gulai asam keueung masakan, Hangop, Hati Babi, Hati Sapi, Hofa/Ubi hutan segar, Ikan Patin segar, Kentang, Keong, Kepala Susu (Krim), Kepiting, Kerang, Kerbau daging segar, Keribang ubi segar, Kerokot segar, Kesemek, Ketimun, Ketimun madura segar, Ketoprak, Ketupat Tahu, Kokosan, Komak polong segar, Koro wedus polong, Kotiu hinela tawang nggole, Krokot, Kucai, Kucai Muda (Lokio), Kuda daging segar, Kue Koya, Kue sus, Kulit melinjo, Kundur segar, Kunyit, Kura-kura, Langsat, Lantar segar, Lawar babi masakan, Lawar penyu masakan, Lawara jangang masakan, Lawara penjah masakan, Lema/ Rebung asam, Lemon segar, Lemon Squasih, Lemonade, Leunca buah, Lilin bungkus gedi, Lobak, Lokan segar, Lontar segar, Lumai/Leunca segar, Mangga benggala segar, Oramu ninahu ndawa olaho masakan, Otak, Paku hinela wulelenggapaya, Pala daging, Papais, Papeda, Parede baleh masakan, Paria (Pare), Paria Putih kukus, Paria putih segar, Pe-Cay, Peda Ikan Banjar, Pelecing kangkong, Pelepah manuk masakan, Pempek adaan, Pempek kapal selam, Pempek kelesan, Penyu serapah masakan, Pepare ular segar, Pepaya Muda, Pepaya segar, Pucuk lumai/daun leunca segar, Punai daging goreng, Pundut nasi, Purundawa, Purut segar, Putri malu segar, Putri selat, Ragi, Rajungan segar, Rambutan, Rambutan Aceh, Rambutan binjai segar, Rambutan sinyonya, Rawon masakan, Rebon (udang kecil segar), Rebung, Rendang sapi masakan, Rimbang segar, Rujak cingur, Rumput laut, Rusip, RW anjing masakan, Uceng/ bunga melinjo segar, Udang besar segar, Udang galah segar, Udang segar, Ulat sagu segar, Umbut rotan, Usus Sapi, Uwi, Waluh balamak, Woku ubi, Wortel Segar, Wortel kukus, Wortel rebus, Yoghurt.
2	Abon, Abon haruwan, Ampas Tahu, Angsa, Arwan sirsir, Ayam, Ayam goreng Kentucky sayap, Babi hutan masak rica masakan, Bebek (itik), Bebek alabio daging segar, Bebek daging goreng, Belut goreng, Buaya daging dendeng mentah, Bubur, Bungkil Kacang Tanah, Cassavastick, Coklat Pahit batang, Corned Beef, Cumi-cumi goreng, Daging Babi Gemuk, Daging Babi Kurus, Dendeng belut goreng, Dendeng Daging Sapi, Dendeng mujahir goreng, Domba daging gemuk segar, Empal (daging) Goreng masakan, Empal Goreng, Enting-enting wijen, Gendar goreng, Ham, Itik alabio daging dendeng mentah, Jambal goreng, Kacang atom, Kacang belimbing (kecipir) kering, Kacang goyang, Kacang Kedelai basah, Kaholeo masakan, Kecipir biji, Keju, Keju Kacang Tanah, Kelapa hutan kering, Kerupuk urat, Kluwek, Kripik Tempe Goreng, Kue kelapa, Kwaci, Kwark (Quark), Lamtoro var. gung tanpa kulit, Pala biji, Paniki masak santan masakan, Pencok lele masakan, Pisang Siam goreng, Rebon kering, Rusa daging dendeng mentah, Saga biji tanpa kulit, Santan murni, Sapi daging dendeng mentah, Sapi daging gemuk segar, Sardines dalam kaleng, Sari Kedelai bubuk, Sate Bandeng, Sate pusut masakan, Sie reuboh masakan, Susu bubuk, Tahu telur, Teri balado masakan, Teripang dendeng mentah, Terong + Oncom makanan, Tinoransak masakan, Toge -Tahu makanan, Udang kering, Udang kering mentah, Udang papay mentah, Ulat sagu panggang, Wijen, Worst (sosis daging).

Table 7. An Explanation Of Every Cluster

Cluster	Description
0	From 100 grams of food, this cluster found that the food had a high calorie content (60–720 calories), low protein and fat content (average <50 mg), and high carbohydrate content (>300 mg)
1	This cluster discovered that, out of 80 grams of food, the average protein and fat content was between 2 and 50 mg, the calorie content was between 100 and 500 kcal, and the carbohydrate level was less than 5 mg.
2	This cluster discovered that the dietary data had enough carbohydrate levels (average of over 1 mg) from 90 grams of food, protein levels below 60 mg, fat levels below 40 mg, and calorie levels between 4 kcal and 300.

Three clusters with disparate outcomes are derived from the K-Medoids algorithm's implementation results.

Foods with high calorie and carbohydrate content are found in Cluster 0. They have minimal quantities of fat and protein. According to this cluster, food data has a high calorie content (60–720 calories), low protein and fat content (average <50 mg), and high carbohydrate content (>300 mg) per 100 grams of food.

Foods having a high calorie content and relatively low protein, fat, and carbohydrate contents are also abundant in cluster 1. According to this cluster, 80 grams of food had a high calorie level (above 100 kcal to 500 kcal), an average protein and fat content (above 2 mg to 50 mg), and a carbohydrate content (below 5 mg).

Foods with very high calorie content and relatively low levels of protein, fat, and carbohydrates make up the final cluster, cluster 2. According to this cluster, dietary data had nutritional levels ranging from 4 to 300 calories, less than 60 mg of protein, less than 40 mg of fat, and more than 1 mg of carbs on average from 90 grams of food.

In K-Medoids clustering, the elbow approach visualizes an "elbow" graph made from the SSE (sum of squared errors) values against different numbers of potential clusters in order to discover the ideal number of clusters. The elbow points on this graph show how many clusters are best suited to the data. Alternatively, the K-Medoids algorithm can be used to assess the quality of the clustering it produces by using the DBI (Davies-Bouldin Index) approach. DBI is used to assess how well data lies within a cluster (intra-cluster) and how well data separates clusters (inter-cluster). The higher the quality of clustering generated by K-Medoids, the lower the DBI score.

IV. CONCLUSION

The ideal number of clusters is three, according to research using K-Medoids to categorize meals according to nutritional values from the selection phase to testing. The Davies-Bouldin Index Cluster and the Elbow Method are used in the cluster optimization process to calculate this number. Foods with high calorie and carbohydrate content are found in Cluster 0.

They have minimal quantities of fat and protein. Foods with high calorie content and relatively low protein, fat, and carbohydrate content are also prevalent in cluster 1. Foods with a high calorie content and relatively low proportions of protein, fat, and carbohydrates make up the final cluster, cluster 2.

The Elbow and Davies-Bouldin Index methods are both employed in the K-Medoids algorithm. The Elbow method yields a value of 0.046, which is utilized as an elbow point, while the Davies-Bouldin Index approach yields the smallest value, 0.631.

REFERENCES

[1] P. Alga Vredizon, H. Firmansyah, N. Shafira Salsabila, and W. Eko Nugroho, "Penerapan Algoritma K-Means Untuk Mengelompokkan Makanan Berdasarkan Nilai Nutrisi," *J. Technol. Informatics*, vol. 5, no. 2, pp. 108–115, 2024, doi: 10.37802/joti.v5i2.577.

[2] A. Pratama, "Noodle Grouping Based On Nutritional Similarity With Hierarchical Cluster Analysis Method," *Anal. Kekuatan Kontruksi Rangka pada Peranc. Belt Conveyor Menggunakan Ansys Work.*, vol. 2, no. 2, pp. 58–71, 2023.

[3] I. W. Septiani, A. C. Fauzan, and M. M. Huda, "Implementasi Algoritma K-Medoids Dengan Evaluasi Davies-Bouldin-Index Untuk Klasterisasi Harapan Hidup Pasca Operasi Pada Pasien Penderita Kanker Paru-Paru," *J. Sist. Komput. dan Inform.*, vol. 3, no. 4, p. 556, 2022, doi: 10.30865/json.v3i4.4055.

[4] A. Hoerunnisa, G. Dwilestari, F. Dikananda, H. Sunana, and D. Pratama, "Komparasi Algoritma K-Means Dan K-Medoids Dalam Analisis Pengelompokan Daerah Rawan Kriminalitas Di Indonesia," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 103–110, 2024, doi: 10.36040/jati.v8i1.8249.

[5] B. Indikator and I. P. M. Dengan, "Analisis perbandingan silhouette coefficient dan metode elbow pada pengelompokkan provinsi di indonesia berdasarkan indikator ipm dengan k-medoids 1,2,3," vol. 13, pp. 13–24, 2024, doi: 10.14710/j.gauss.13.1.13-24.

[6] E. A. Saputra and Y. Nataliani, "Analisis Pengelompokan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode Clustering K-Means," *J. Inf. Syst. Informatics*, vol. 3, no. 3, pp. 424–439, 2021, doi: 10.51519/journalisi.v3i3.164.

[7] H. S. Kasman, "Perbandingan Efektifitas Algoritma K-Means Clustering -Topsis dan K-Medoids Clustering -Topsis dalam Menentukan Karyawan IT dengan Kinerja Terbaik A . Peneliti Terdahulu K-Means merupakan metode clustering non-hierarkis yang berusaha membagi data ke dalam," vol. 4, no. 2, 2024.

[8] B. A. Setiawan and Sulastri, "Perbandingan Clustering Optimalisasi Stok Barang Menggunakan Algoritma K – Means Dan Algoritma K – Medoids," pp. 978–979, 2021, [Online]. Available: https://unisbank.ac.id/ojs/index.php/sendi_u/article/view/8645

[9] P. Vania and B. Nurina Sari, "Perbandingan Metode Elbow dan Silhouette untuk Penentuan Jumlah Klaster yang Optimal pada Clustering Produksi Padi menggunakan Algoritma K-Means," *J. Ilm. Wahana Pendidik.*, vol. 9, no. 21, pp. 547–558, 2023, [Online]. Available: https://doi.org/10.5281/zenodo.10081332

[10] N. A. Maori and E. Evanita, "Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering," *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023, doi: 10.24176/simet.v14i2.9630.

[11] F. Irhamni, F. Damayanti, B. Khusnul K, and M. A, "Optimalisasi Pengelompokan Kecamatan Berdasarkan Indikator Pendidikan Menggunakan Metode Clustering dan Davies Bouldin Index," *Semin. Nas. dan Teknol. UMJ*, no. 11, pp. 1–5, 2014.

[12] I. T. Umagapi, B. Umaternate, S. Komputer, P. Pasca Sarjana Universitas Handayani, B. Kepegawaian Daerah Kabupaten Pulau Morotai, and B. Riset dan Inovasi, "Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa," *Semin. Nas. Sist. Inf. dan Teknol.*, vol. 7, no. 1, pp. 303–308, 2023.

[13] I. F. Fauzi, M. Gito Resmi, and T. I. Hermanto, "Penentuan Jumlah Cluster Optimal Menggunakan Davies Bouldin Index pada Algoritma K-Means untuk Menentukan Kelompok Penyakit," vol. 7, no. 2, pp. 2598–8069, 2023.

[14] A. Atira and B. Nurina Sari, "Penerapan Silhouette Coefficient, Elbow Method dan Gap Statistics untuk Penentuan Cluster Optimum dalam Pengelompokkan Provinsi di Indonesia Berdasarkan Indeks Kebahagiaan," *J. Ilm. Wahana Pendidik.*, vol. 9, no. 17, pp. 76–86, 2023, [Online]. Available: https://doi.org/10.5281/zenodo.8282638

[15] A. R. Aruadi, F. Andreas, and B. N. Sari, "Analisis Klaster K-Means pada Data Rata-Rata Konsumsi Kalori dan Protein Menurut Provinsi dengan Metode Davies Bouldin Index," *J. Pendidik. dan Konseling*, vol. 4, no. 4, pp. 5652–5661, 2022.

[16] D. Nasution, D. N. Sirait, I. Wardani, and Dwiyanto, "Optimasi Jumlah Cluster Metode K-Medoids Berdasarkan Nilai DBI Pada Pengelompokkan Data Luas Tanaman Dan Produksi Kelapa Sawit Di Sumatera Utara," *Kumpul. J. Ilmu Komput.*, vol. 9, no. 2, p. 381, 2022.

ISSN 2085-4579

Project Management and Cost Optimization in QR Code-Based Parking Reservation System Development: A PERT, Monte Carlo and PMBOK Approach

Yostian Ari Sujarwo[1], Samidi[1]
[1] Master of Computer Science, Faculty of Information Technology, Budi Luhur University
2211601915@student.budiluhur.ac.id
samidi@budiluhur.ac.id

Accepted on May 16, 2025
Approved on June 30, 2025

***Abstract*—This study investigates project cost optimization through a quantitative case study on the development of an in-house parking reservation system using QR code technology, utilizing a descriptive-quantitative method based on the PMBOK (Project Management Body of Knowledge) framework. Project activities were systematically organized using a Work Breakdown Structure (WBS) to enhance clarity and task accountability. To improve cost efficiency and support data-driven decision-making, PERT (Program Evaluation and Review Technique) cost analysis was applied to refine cost estimates based on optimistic, most likely, and pessimistic scenarios. Following this, Monte Carlo simulation was employed to assess cost uncertainty and variability, providing a probabilistic range of possible outcomes. The simulation results revealed a significant reduction in projected human resource costs from IDR 748 million to IDR 403 million, with a 95% probability that the actual total project cost would fall within this range. These findings offer a more realistic and risk-aware estimation for project budgeting and planning. Furthermore, robust risk management strategies and proactive stakeholder engagement contributed to informed decision-making, enhanced transparency, and overall project success.**

***Index Terms*—Project Management; Qr Code Technology; Parking Reservation System; PERT Cost Analysis; Monte Carlo Simulation**

I. INTRODUCTION

The rapid advancement of technology has significantly transformed various sectors, including urban infrastructure and management. The growing demand for efficient parking solutions in urban areas has led to the need for more sophisticated management systems. Traditional parking reservation methods, often characterized by manual processes and limited accessibility, are becoming inadequate to meet modern requirements. The adoption of QR code technology offers a streamlined approach to parking reservations, enabling users to book and manage their parking spaces with ease and precision. However, the successful implementation of such a system requires careful project management to address various challenges, including technical integration, resource allocation, risk management, and stakeholder coordination.

QR Code is a method that converts textual information into two-dimensional codes, which are then printed on a more compact medium. QR Code is a form of bar code that can be scanned and interpreted using a mobile phone camera [1].

As concluded by Farooq et al. “Many studies have indicated that insufficient skills, a lack of understanding in key knowledge areas, or the inability to effectively use modern tools by the project manager primarily cause project failures and delays. Consequently, these project managers struggle to effectively apply project management methods throughout the project life-cycle” [2].

To address these challenges effectively, it is crucial to implement robust project management practices that encompass all phases of the project life-cycle. This includes thorough planning, clear definition of project objectives, detailed scheduling, efficient resource management, and proactive risk management. In addition, ensuring that all stakeholders are engaged and their needs are addressed is vital for the successful deployment of the QR code-based parking reservation system.

Numerous strategies have been suggested to assist in human resource scheduling and allocation using models that incorporate triangular constraints. These approaches include optimizing project duration, reducing project costs, enhancing revenue, or improving efficiency [3]. In the human resource management, study says that “Strategic human resource planning is crucial for effective construction

project management. Proper planning and allocation of human resources significantly enhance overall project performance" [4]. Therefore, managers must focus more on how their team operates as a cohesive unit, while assessing collaboration and communication skills, as these factors directly affect teamwork performance.

In addition to human resource management, project cost management also plays a crucial role in the success of a project. Successful project cost management demands precise budget forecasts, continuous oversight, and the flexibility to modify plans as needed. Without adequate control, cost overruns and delays frequently impede the smooth execution of the project. Numerous study says that poor budget estimation can cause cost overrun. According to statistics, about 25% of all software projects that are started are completed on time, another 25% are cancelled, and about 50% of projects are completed over budget or late [5]. Another study says "Project success is described through three main factors: cost, quality, and project implementation time" [6]. Foroutan Mirhosseini et al. (2022) identified factors that influence cost increases in the planning phase, including: Inaccurate estimation, ineffective project management, technological advancements, and changes or variations in staffing [7]. But not all of the study says cost overrun makes project delayed or canceled. Ning Nan & Harter. (2009) says "development team get motivated and can catch up with the schedule when the project comes to get cost overrun" [8].

Effective project planning and scheduling can greatly enhance project performance in every industry. By dedicating more effort to planning, scheduling, and control, project performance can be significantly improved and also can avoid project become over budgeting.

In the context of this project, the application of QR code technology aims to simplify the reservation process and enhance user convenience by allowing real-time booking and management of parking spaces. The integration of this technology into an in-house system demands meticulous attention to technical details and coordination among various teams. Effective project management is not only essential for overcoming technical hurdles, but also for ensuring that the system meets the desired operational standards and delivers value to both users and stakeholders.

II. METHOD

A. Project Management Body Of Knowledge

The Project Management Institute (PMI) has significantly contributed to the establishment of standards and documentation of project management practices. Their widely accepted Project Management Body of Knowledge (PMBOK) is a comprehensive resource detailing standards, guidelines, and best practices for project management. The PMBOK, published by PMI, encompasses a range of globally recognized processes, tools, techniques, and methodologies designed to aid project managers in effectively overseeing projects. The guide is structured around several key knowledge areas, including scope, time, cost, quality, resource, communication, risk, procurement, and stakeholder management.

The PMBOK acts as an essential resource for achieving Project Management Professional (PMP) certification and is widely adopted by organizations to enhance project success and consistency in project management practices. The PMBOK is structured around several key knowledge areas essential for successful project management:

- Scope Management: Defines and controls what is included and excluded in the project.
- Time Management: Involves planning and controlling the schedule to ensure timely completion.
- Cost Management: Encompasses budgeting and controlling project costs to stay within financial constraints.
- Quality Management: Guarantees that the project deliverables meet the necessary standards and satisfy stakeholder expectations.
- Resource Management: Focuses on the efficient allocation and utilization of human and material resources.
- Communication Management: Involves the effective dissemination of information among project stakeholders.
- Risk Management: Recognizes, evaluates, and addresses risks that could affect the project.
- Procurement Management: Manages the acquisition of goods and services from external sources.
- Stakeholder Management: Involves engaging and managing relationships with all parties affected by the project.

In working on a project, it is important to follow a structured foundation. Figure 1 describes the project life cycle, which serves as the general framework for project execution.

This study focuses on optimizing project cost and improving efficiency during the development of an in-house parking reservation system using QR code technology. To guide the investigation, the following research questions are proposed:

- How can structured project management approaches improve project cost efficiency?

- How effective are PERT and Monte Carlo methods in estimating and controlling project costs?

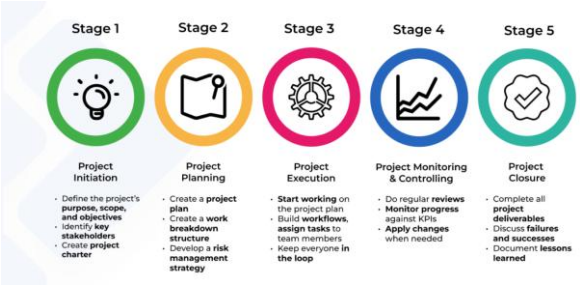


Fig. 1. Project Life Cycle

B. *Project Initiation*

Project initiation is the first phase in the project life-cycle, designed to lay a strong foundation for the project. This stage encompasses several crucial steps to ensure the project has a clear objective and receives support from all stakeholders. In a traditional organizational structure, project initiation involves the mobilization of resources across various organizational units to achieve objectives within a defined timeframe, under the supervision of a project manager [9]. The initial step in starting a project is to identify the problem or potential benefit that the project seeks to address. This involves a thorough examination of the problem or opportunity, coupled with consultations with stakeholders to gain a clear understanding of their requirements.

After identifying the need or opportunity, the next step is to create a business case. This document outlines the project's benefits, costs, risks, and objectives, providing a rationale for why the project should proceed and how it will benefit the organization. The business case is then evaluated through a feasibility study to assess whether the project is viable from technical, economic, and operational standpoints. This study ensures that risks can be managed and that necessary resources are available.

C. *Project Planning*

Project planning is the second stage in the project life-cycle, focusing on detailing how the project will be executed and managed. This stage is crucial as it ensures that all project aspects are thoroughly planned, thereby minimizing risks and enhancing the likelihood of success. A key component of project planning is developing a comprehensive project plan, which addresses various elements such as project objectives, timeline, funding, standards, assets, and potential issues, communication, and procurement.

The project scope is meticulously defined to identify and document all necessary tasks, including the development of a Work Breakdown Structure (WBS) that divides the project into more detailed, actionable tasks. WBS is a crucial tool and technique in scope management within project management [10]. Another essential element is project scheduling, which lists all tasks in chronological order and establishes their interrelationships, resulting in a project schedule that indicates when each task should start and finish, and identifies the risky path that determines the project's overall duration.

Communication is a vital tool in project management, and a project's success depends significantly on how effectively its communication system operates. As time goes by, communication becomes increasingly important and central to all management processes [11]. The project communication plan describes how information will be distributed to stakeholders and the project team, including the frequency and methods of communication to be used, ensuring that all parties receive timely and transparent information. Procurement planning identifies the need for purchases or contracts with third parties and outlines the process for acquiring and managing these goods or services.

By creating a thorough and structured project plan, the planning phase ensures that all project aspects are thoughtfully considered and planned. This approach not only facilitates more effective project management, but also provides clear guidance for the project team in executing their tasks. Effective planning is crucial for achieving successful project outcomes, as it helps identify potential issues early on and provides strategies to address them before they escalate. But it is also necessary to know the focus on project planning varies according to the project's size and the type of product being developed. However, regardless of the project type, short-term planning for sprints is consistently conducted [12].

D. *Project Execution*

Project Execution is the third stage in the project life-cycle, concentrating on implementing the project plan developed during the planning phase. During this phase, all planned activities are carried out to produce the project's deliverable. It is also the phase where most of the budget and resources are utilized, making effective management crucial to keep the project on track and achieve its objectives.

The execution phase involves coordinating between the project team and stakeholders to carry out the tasks outlined in the project schedule. The project manager has a vital role in leading the team, managing resource allocation, and overseeing the implementation of day-to-day activities. Effective communication is crucial at this stage to keep all team members and stakeholders updated on project progress, changes, and any emerging issues. Regular meetings, status reports, and online collaboration tools are commonly used to facilitate smooth communication.

The primary goal of the execution phase is to produce project outputs that meet the defined objectives and specifications. The success of this phase largely depends on the project team's collaboration, the project manager's ability to manage resources and resolve issues, and the adaptability of the project plan in response to changes and challenges. With effective management, the execution phase can ensure the project progresses smoothly and deliverable align with stakeholder expectations.

E. Project Monitoring And Controlling

Project Monitoring and Controlling is a stage in the project life-cycle that operates alongside the execution phase. This stage aims to ensure the project is proceeding as planned, meeting its objectives, and adhering to quality standards. It involves ongoing monitoring of project performance, evaluating work outcomes, and make the required adjustments to keep the project on track.

A key aspect of monitoring and controlling is measuring project performance. Project managers utilize various methods and strategies to collect performance information, such as earned value analysis (EVA), key performance metric (KPM), and status reports. The data is subsequently matched against the project plan to detect any discrepancies or variances.. For instance, if the project encounters delays or exceeds its budget, corrective actions can be implemented to resolve the issue.

Project control also encompasses change management. Changes in a project are often unavoidable, whether they pertain to scope, schedule, cost, or quality. The change control process ensures that we thoroughly evaluate all proposed changes, obtain approval from relevant stakeholders, and document them properly. This helps maintain project stability and prevents changes from negatively impacting the project's overall success.

Additionally, risk management is a crucial component of monitoring and controlling. Risks identified during the planning stage are continuously monitored, and mitigating actions are implemented if they occur. Risk monitoring also involves identifying new risks that may arise during project execution, and assessing their impact on the project. Effective risk management helps minimize potential obstacles and ensures the project's successful completion.

F. Project Closure

Project Closure is the final stage in the project life-cycle, signaling the formal completion of all project activities. The main goal of this stage is to confirm that all deliverable have been completed and delivered according to the specifications and that all project objectives have been met. Additionally, this phase aims to formally conclude all project activities and ensure that all resources utilized during the project are released and reassigned as necessary within the organization.

One of the initial steps in project closure is the verification and acceptance of deliverable. The project team collaborates with stakeholders to make sure that all project results meet the defined conditions and standards. This process often includes final testing, inspections, and obtaining approval signatures from relevant stakeholders. Once confirmation is received that deliverable have been accepted, the project is considered formally completed.

The next step involves preparing project closure documentation. This includes a final project report that summarizes the project's performance, results achieved, and lessons learned throughout the project. The report typically features variance analysis, highlighting differences between the initial plans and actual outcomes, and offers recommendations for future projects. Additionally, all project documents, such as contracts, financial statements, and correspondence, are archived for future reference.

Project evaluation is also a crucial component of project closure. This involves a final evaluation of the project's outcomes as well as the effectiveness of the project team. This evaluation is often conducted through a closing session where the project team and stakeholders review what went well and identify areas for improvement. This session helps to capture lessons learned that can be used to enhance performance and results in future projects.

The project closure phase occurs when the achieved requirements and goals align with the agreement between the product owner and the team. At this point, the final version of the product is prepared for release and distribution. Additionally, all necessary documentation and user manuals are completed and ready [13].

G. PERT

The PERT (Program Evaluation and Review Technique) method is a project management approach that utilizes statistical analysis to predict task completion times. It serves as a fundamental framework for scheduling and cost estimation. However, since construction projects are inherently uncertain, advanced modeling techniques are necessary to precisely estimate the Most Likely (M), Pessimistic (P), and Optimistic (O) values. These estimates are essential for effective resource allocation and project planning [14]. The most likely estimate represents a projection based on the assumption that both advantageous and disadvantageous conditions will occur in a balanced and realistic manner. A pessimistic estimate refers to a projection made under the assumption that all adverse conditions prevail, where every potential threat materializes and no opportunities are realized. An optimistic estimate refers to a projection made under the assumption that

all favorable conditions prevail, wherein every opportunity is realized and no adverse events or threats occur [15].

Although PERT is commonly applied in project time management to evaluate and review project durations, it can also be extended to project cost management. In this context, what is originally referred to as the optimistic estimate assuming all favorable conditions will occur translates into the lowest possible project cost. Similarly, the most likely and pessimistic estimates represent, respectively, the expected cost under balanced conditions and the highest potential cost when all unfavorable factors are assumed to take place.

PERT, also referred to as the Back Research Technique, utilizes time as a key variable in planning, scheduling, organizing, coordinating, and controlling uncertain activities, while also defining performance specifications [16]. A central aspect of PERT is its ability to account for uncertainties in the duration of activities during the analysis [17].

Using PERT in project management, project manager can identify potential risks, optimize workflow and improve decision-making for the project. The PERT method also promotes a more structured and analytical approach to project execution. By breaking down complex project into smaller and manageable tasks. Besides that, the method improve coordination between stakeholders with valuable insights into potential risks and uncertainties, ensuring the project are completed within the expected timeframe and budget.

The formula used for PERT-based cost estimation is as follows:

$$\text{PERT Cost Estimate} = \frac{O + 4M + P}{6} \quad (1)$$

Where:

- O = Optimistic cost which can be interpreted as the lowest possible cost
- M = Most likely which means normal cost
- P = Pesimistic cost which means highest possible cost

This formula enables project managers to obtain a weighted average of the estimated project cost, giving greater emphasis to the most likely cost while still accounting for best-case and worst-case scenarios. By incorporating uncertainty into the estimation process, the PERT method enhances the accuracy of budget forecasts and supports more effective cost planning and risk management throughout the project life-cycle.

H. Monte Carlo Simulation

The Monte Carlo method is a problem-solving approach that deals with stochastic characteristics by utilizing repeated statistical experiments. This method relies on a probabilistic model, aligning with the processes defined by the model, to generate simulation test results as an estimated solution to the problem [18]. Monte Carlo method is widely used in various field, including finance and engineering. As Abu Bakar and Rosbi did in their journal *Monte Carlo Simulation for Data Volatility Analysis of Stock Prices in Islamic Finance for Malaysia Composite Index* [19]. They use monte carlo simulation method to assess the fluctuation rate of Sharia-compliant companies in the Malaysia Stock Exchange.

Monte carlo used to analyze complex system, optimize decision-making and assess risk by simulating numerous possible outcomes. By running a large number of simulations, the monte carlo method allows project manager to understand uncertainties. Additionally, this method allows businesses, scientists, or anyone who used it to explore different scenarios, evaluate potential risks, and develop robust strategies based on probability distributions rather than fixed assumptions. As Tokdemir et al. state in their paper, Monte Carlo simulation generates a probability distribution of potential outcome values after multiple iterations and Monte Carlo simulation allows decision-makers to estimate the likelihood of meeting project objectives and to analyze the differences between the base case scenario and both the best- and worst-case outcomes [20]. As shown in [21], the Monte Carlo method can yield more accurate results than other uncertainty analysis models while using less sample data.

III. RESULT AND DISCUSSION

In this section, we examine the outcomes of implementing project management in the development of QR code-based parking reservation system.

A. Conceptual Framework

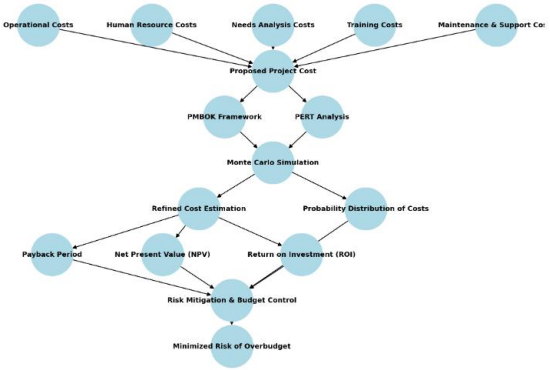


Fig. 2. Conceptual Framework

In the conceptual framework used describes how the risk of cost overrun in a project can be reduced

with the PMBOK Framework, PERT and Monte Carlo Simulation approach. Here is an explanation of each part:

- *Cost Factor*

There are several factors affect the Proposed Project Cost, including Operational Costs, Human Resource Costs, Needs Analysis Costs, Training Costs and Maintenance & Support Costs. This factors are the main elements that determine the initial cost of Proposed Project Costs. Precise of initial costs is crucial, as errors at this stage can cause to budget inconsistencies, potentially resulting in cost overrun or insufficient funding.

- *Cost Management Method*

Once a project has an initial cost estimate, the PMBOK framework and PERT Analysis methods are used to improve the accuracy of planning cost. PMBOK framework provides a systematic approach to project cost management and PERT Analysis is used for estimating project duration and cost by considering optimistic, pessimistic and realistic scenarios.

- *Monte Carlo Simulation*

The result of PERT Analysis are used in monte carlo simulation which will simulate possible cost distributions based on various scenarios, generated refined cost estimation and provide probability distribution of costs to understand cost risk statistically. And generate Payback Period, NPV and Return of Investment.

- *Risk Mitigation and Budget Control*

Based on monte carlo simulation result, we can implement risk mitigation and budget control to reduce uncertainty in project costs and improve the accuracy of budget control.

B. Project Charter

TABLE I. PROJECT CHARTER TABLE

Project Title	: :	Sistem Reservasi Parkir Inap
Project Information	: :	Start Date : February 20, 2020 Complete Date : February 20, 2021 Project Manager : Yostian Ari Sujarwo Key Stakeholders : PT. Sanggraha Daksamitra
Project Objective	:	Developing an overnight parking system using Qr Code
Project Scope	:	Building web and mobile based applications for end users Building desktop based applications for parking gate admin Integrating applications with parking gates
Project Limitation	:	This system is only built for PT. Sanggraha Daksamitra
Project Milestones	:	Web, mobile and desktop UI/UX design completed May 20, 2020 Backend development completed October 20, 2020 Backend integration with web, mobile and desktop UI completed December 2020 System testing completed February 14, 2021 Beta Release February 17, 2021 Final Release February 20, 2021
Project Budget	:	The total budget for this project is IDR 900,000,000 which includes the cost of application development and infrastructure development, testing, training, and also socialization.
Main Risks	:	Delay in project completion Integration issues with parking gates
Agreement	:	Yostian Ari Sujarwo (Jakarta, Februari 2020)

C. Project Planning

- *Requirement*

The project requirements document provides guidance on various aspects important for the implementation and success of the project. In this project it will be divided into 2 specifications, namely: a) Functional Specification; b) Non Functional Specification. Both spesification will be describe in table II and table III.

TABLE II. FUNCTIONAL SPESIFICATION

Function	Description	Requirement
Dashboard	Displaying parking lot statistics	Every user and admin in the system can see the availability of parking lots.
User Management	Displaying user information	Each user can change information about themselves. Admin can disable problematic users
Reservation	Displaying reservation information	Each user can see the reservation information that is ongoing or has occurred. Each staff can see the reservations that the user made
Shuttle Bus	Displays shuttle bus schedule information	Every user can see the shuttle bus schedule. Admin can change the shuttle bus schedule
Report	Displays information about financial reports	Each staff can view financial reports and print financial reports.

TABLE III. NON FUNCTIONAL SPESIFICATION

Category	Specification	Detail
Performance	Response Time	
	Home Page	< 1 second for standard internet
	Parking Reservation	< 3 seconds for parking reservation
Scalability	Horizontal Scaling	

	Infrastructure	Supports dynamic addition of additional servers
	Vertical Scaling	
	Server	Ability to increase CPU, RAM and storage capacity
Security	Authentication	
	Authentication Method	Token-based authentication and 2FA
	Session Management	User sessions expire automatically after 10 minutes of inactivity.
	Encryption	
	Data	Encryption of sensitive data in storage and transmission
	SSL Certificate	Use of SSL certificates for secure communications
Availability	Operational Hours	
	SLA	99.9% within one year; planned downtime < 1 hour per month
	Backup	
	Backup Frequency	Automatic backup every night
	Recovery	Data recovery within < 1 hour after failure
Maintenance	Updates	
	Update Frequency	Automatic software and security update
	Compatibility	Updates are compatible with supported hardware and software.
	Monitoring	
	System Log	Activity and error logging
	Monitoring Tools	Real-time performance and availability monitoring

- *Work Breakdown Structure*

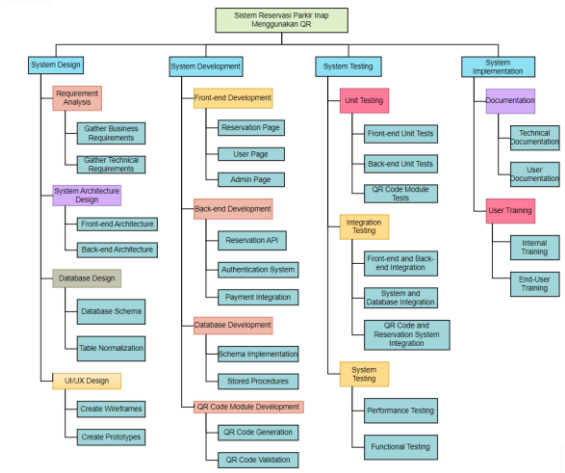


Fig. 3. Work Breakdown Structure

Fig 3 shows the work breakdown structure developed by the author based on the actual project tasks. The tasks are divided into four main categories: system design, system development, system testing, and system implementation.

D. Project Schedule Management

- *Management Approach*

A management approach is a systematic framework designed to help organizations manage projects, processes, or daily operations effectively and efficiently. It starts with establishing SMART (specific, measurable, achievable, relevant, and time-bound) targets and aims. Following goal-setting, strategic and operational planning is conducted to develop both long-term and short-term plans that include a vision, mission, strategy, and resource allocation. This phase also involves risk management to identify, analyze, and plan for potential risks during implementation.

Next, the approach focuses on organizing and allocating resources by setting up an effective organizational structure and managing assets such as personnel, funds, equipment, and technology. During the implementation phase, project management techniques are used to lead the team, manage activities according to the schedule, budget, and quality standards, and ensure effective communication with stakeholders.

The management approach also emphasizes full documentation and reporting, including performance reports and lessons learned. It supports continuous improvement through the Kaizen principle, where feedback from evaluations is used to refine future processes and approaches. By following this structured and systematic approach, organizations can efficiently achieve their goals, manage resources effectively, and mitigate risks during project or operational execution.

- *Scheduling Methodology*

Scheduling methodology is a structured approach for planning, organizing, and managing time and resources needed to complete a project efficiently. The process begins with determining the tasks and activities required to accomplish the project's objectives. Each task is detailed, including its duration, order, and dependencies.

Next, the project schedule is developed using tools like Gantt charts or network diagrams, along with methods like the Critical Path Method (CPM) or the Program Evaluation and Review Technique (PERT). Gantt charts offer a visual overview of the project timeline, detailing the tasks to be completed and the time assigned to each. CPM and PERT assist in identifying the critical path—the series of tasks that influence the project's total duration and cannot be postponed without impacting the project's completion.

- *Critical Path Method (CPM)*

Critical Path Method (CPM) is a project management method used for plan, schedule, and control project activities. According to [22], CPM is a method of analyzing project networks that focuses on optimizing the cost or budget of a project by reducing or accelerating the project's completion time. CPM identifies the longest sequence of dependent tasks in a project, referred to as the critical path, which governs

the overall duration of the project. Tasks on this critical path must be finished on time to ensure that the project meets its deadline.

By pinpointing these critical activities, the project manager can allocate resources and focus appropriately to keep the project on schedule. CPM involves constructing a network diagram to illustrate the sequence and dependencies of tasks, and determining the start and finish times for each activity. It also helps identify float or slack time, which is the allowable delay in non-critical tasks without affecting the entire project timeline. Consequently, CPM is a valuable tool for effective project planning and management, assisting in achieving project goals within the planned schedule. The following is a CPM table for a parking reservation system.

TABLE IV. CRITICAL PATH METHOD

Task	Duration(week)	Dependency(week)
Needs Analysis	1	-
System Design	3	1
Frontend Development	9	2
Backend Development	24	4
QR Code Integration	5	1
System Testing	8	1
User Training	1	2
System Launch	1	1

- *Program Evaluation and Review Technique (PERT)*

PERT (Program Evaluation and Review Technique) is akin to CPM (Critical Path Method) but focuses more on addressing uncertainties in estimating activity durations. PERT uses three different duration estimates: (a) -optimistic duration (b) -pessimistic duration, and most probable duration. This method creates a network model that incorporates these uncertainties and enables probability analysis of project completion times.

Fig. 4 presents the PERT diagram created based on project tasks, illustrating the sequence of activities and the estimated duration range for each. This diagram supported better planning and risk mitigation, especially in identifying activities on the critical path.

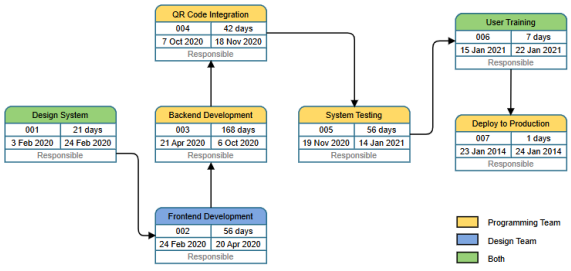


Fig. 4. PERT Diagram

E. Project Cost Management

- *Project Cost*

Project cost refers to the total amount of expenditure needed to finish a project. This comprises several components like personnel, supplies, tools, and other operational costs. Additionally, it accounts for unforeseen costs, including risk and contingency expenses. Effective cost management requires meticulous planning, continuous monitoring, and required modifications to keep the project on budget while maintaining quality and achieving the desired outcomes.

TABLE V. PROPOSED PROJECT COST

No	Cost	Year 0	Year 1	Year 2	Year 3
1	***Operational Cost***				
	Promotion	Rp. 10,000,000	Rp. 5,000,000	Rp. 3,000,000	-
	Server	Rp. 100,000,000	Rp. 100,000,000	Rp. 100,000,000	Rp. 100,000,000
	Total Operating Costs	Rp. 110,000,000	Rp. 105,000,000	Rp. 103,000,000	Rp. 100,000,000
2	***HR Costs***				
	Project Manager	Rp. 180,000,000	-	-	-
	Programmer	Rp. 324,000,000	-	-	-
	System Analyst	Rp. 120,000,000	-	-	-
	Business Analyst	Rp. 100,000,000	-	-	-
	Tester	Rp. 24,000,000	-	-	-
	Total HR Cost	Rp. 748,000,000	-	-	-
3	***Needs Analysis Cost***				
	Analysis of user and stakeholder needs	Rp. 5,000,000	-	-	-
	Total Cost of Needs Analysis	Rp. 5,000,000	-	-	-
4	***Training Cost***				
	User training	Rp. 5,000,000	-	-	-
	Total Training Cost	Rp. 5,000,000	-	-	-
5	***Maintenance and support cost***				
	System maintenance and support	-	Rp. 10,000,000	Rp. 15,000,000	Rp. 20,000,000
	Parking gate maintenance	-	Rp. 6,000,000	Rp. 6,000,000	Rp. 8,000,000
	Total Cost of Maintenance and Support	-	Rp. 16,000,000	Rp. 21,000,000	Rp. 28,000,000

	Total cost	Rp. 868,000 ,000	Rp. 121,000 ,000	Rp. 124,000 ,000	Rp. 128,000 ,000

- *Pert Analysis*

After we have activity pert diagram and proposed project cost, the next step is to perform a detailed breakdown of tasks and resources. This involves estimating the cost of each activity, allocating necessary resources, and aligning the cost estimates with the project schedule to ensure efficient planning and execution. First of all, we need to breakdown the HR costs.

TABLE VI. HR COSTS

No	Position	Salary(million)	Period(month)
1	Project Manager	15	12
2	System Analyst	12	10
3	Business Analyst	10	10
4	Frontend Programmer	13	12
5	Backend Programmer	14	12
6	Tester	8	3

The table above explains the component of human resource costs required in this project. Each row represents one position involved in the project, along with the estimated monthly salary and period of involvement of the personel. It is also important to note that not all roles are engaged for the entire duration of the project. Key positions such as the Project Manager, Frontend Programmer, and Backend Programmer are involved throughout the full implementation period (12 months), ensuring continuous oversight and development. In contrast, roles such as the Tester are only required for specific phases particularly in the final stages of the project resulting in a shorter engagement period of only 3 months.

The next step is to conduct a PERT analysis to ensure that the planned costs correspond accurately with the project's scheduled activities and overall duration.

TABLE VII. MAPPING COST TABLE

No	Activity	Optimistic (Million)	Most Likely (Million)	Pesimistic (Million)	Position
1	Design System	21	31.08	50.4	System Analyst, Business Analyst, Project Manager
2	Frontend Development	38.08	62.72	94.08	Frontend Developer, Project Manager
3	Backend Development	117.6	194.88	288.96	Backend Developer, Project Manager
4	QR Code Integration	29.4	48.72	72.24	Backend Developer, Project Manager
5	System Testing	33.6	51.52	82.88	Tester, Project Manager
6	User Training	2.8	4.2	7	Project Manager
7	Live Production	0.98	1.68	2.4	Frontend Developer, Backend Developer, Project Manager
	Total	**243.96**	**394.80**	**597.96**	

The table above explains optimistic, most likely and pesimistic cost from activity pert diagram, Right now some, key position like frontend programmer and backend programmer are not involved the full implementation period, only project manager that involved the full implementation period. It's because when divided into several acitivities, only project manager that involved in the full implementation period. The table also explains about salary that given to each position has been change, from salary that pay per month to salary pay per day. The next step is to implement pert to every aspect in mapping cost table.

TABLE VIII. PERT ESTIMATION COST

No	Activity	Estimation Cost(Million)
1	Design System	32.94
2	Frontend Development	63.61
3	Backend Development	199.68
4	QR Code Integration	49.26
5	System Testing	53.10
6	User Training	4.48
7	Live Production	1.68
	Total	404.75

The table above shows the estimated cost of each project activity based on PERT method. PERT is used to calculate the cost estimate by considering three possible scenarios. Through this approach, a weighted average cost estimate is obtained for each activity, which provides more realistic cost planning result than using only one estimated value.

- *Monte Carlo Simulation*

Monte Carlo simulation is used to simulate various possible total project cost based on PERT cost estimates of each activity. The result of this simulation

provide a probabilistic distribution of total project cost.

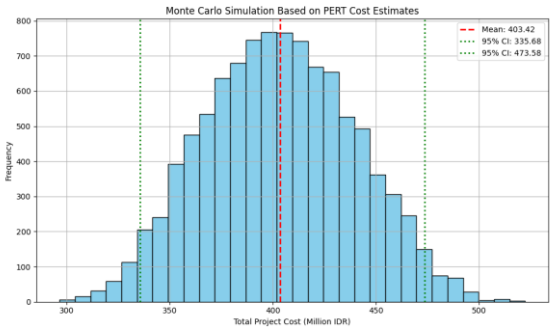


Fig. 5. Monte Carlo Simulation For Total Cost

After performing 10,000 simulations, the Monte Carlo simulation resulted in a mean estimated project cost of 403.42 million IDR, with a 95% confidence interval ranging from 335.68 to 473.58 million IDR. This indicates that there is a 95% probability that the actual total project cost will fall within this range, providing a more realistic and risk-aware cost estimation for project planning. Fig 6 will show cost for every activity in pert cost estimates.

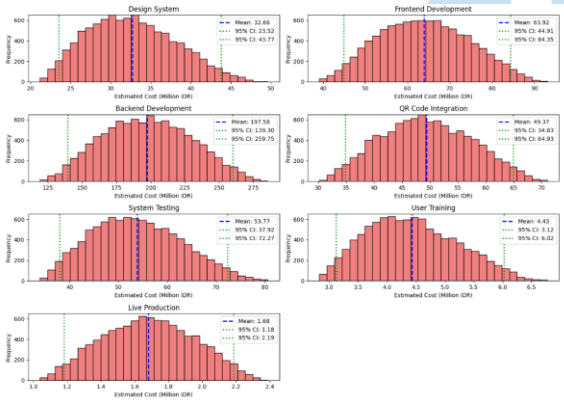


Fig. 6. Monte Carlo Simulation For Every Activity

Fig 6 illustrates the cost distribution for each individual project activity based on 10,000 Monte Carlo simulations using PERT estimates. Each histogram shows the variability and uncertainty of the cost for a specific activity, along with the mean (blue dashed line) and the 95% confidence interval (green dotted lines). The visual representation highlights which activities have higher cost uncertainties and can help identify components with the greatest impact on overall project risk and budget planning.

- *Refined Cost Estimation*

Refined cost estimation is performed to obtain more realistic and accurate cost of the development of the qr parking system. In this process, the PERT approach and monte carlo simulation are used specifically of the hr cost component. Other cost, such as operational costs, need analysis, training, and system maintenance and support are still referred to the initial plan because they have a lower level of uncertainty and have a fixed cost.

TABLE IX. REFINED PROJECT COST

No	Cost	Year 0	Year 1	Year 2	Year 3
1	***Operational Cost***				
	Promotion	Rp. 10,000,000	Rp. 5,000,000	Rp. 3,000,000	-
	Server	Rp. 100,000,000	Rp. 100,000,000	Rp. 100,000,000	Rp. 100,000,000
	Total Operating Costs	Rp. 110,000,000	Rp. 105,000,000	Rp. 103,000,000	Rp. 100,000,000
2	***HR Costs***				
	Project Manager	Rp. 199,990,000	-	-	-
	Programmer	Rp. 157,800,000	-	-	-
	System Analyst	Rp. 10,590,000	-	-	-
	Business Analyst	Rp. 10,590,000	-	-	-
	Tester	Rp. 26,450,000	-	-	-
	Total HR Cost	Rp. 403,420,000	-	-	-
3	***Needs Analysis Cost***				
	Analysis of user and stakeholder needs	Rp. 5,000,000	-	-	-
	Total Cost of Needs Analysis	Rp. 5,000,000	-	-	-
4	***Training Cost***				
	User training	Rp. 5,000,000	-	-	-
	Total Training Cost	Rp. 5,000,000	-	-	-
5	***Maintenance and support cost***				
	System maintenance and support	-	Rp. 10,000,000	Rp. 15,000,000	Rp. 20,000,000
	Parking gate maintenance	-	Rp. 6,000,000	Rp. 6,000,000	Rp. 8,000,000
	Total Cost of Maintenance and Support	-	Rp. 16,000,000	Rp. 21,000,000	Rp. 28,000,000
	Total cost	Rp. 523,420,000	Rp. 121,000,000	Rp. 124,000,000	Rp. 128,000,000

- *S-Curve*

A S-curve is a graphical tool commonly applied in project management to illustrate the progress of work or the usage of resources over time. It is named for its characteristic "S" shape, which reflects the typical progression pattern of a project. Initially, progress is slow as the project starts with planning and setup. As core tasks are executed and resources are heavily engaged, the rate of progress accelerates. Finally, as the project approaches completion, the rate of progress

decelerates, with the focus shifting to finishing touches and final inspections.

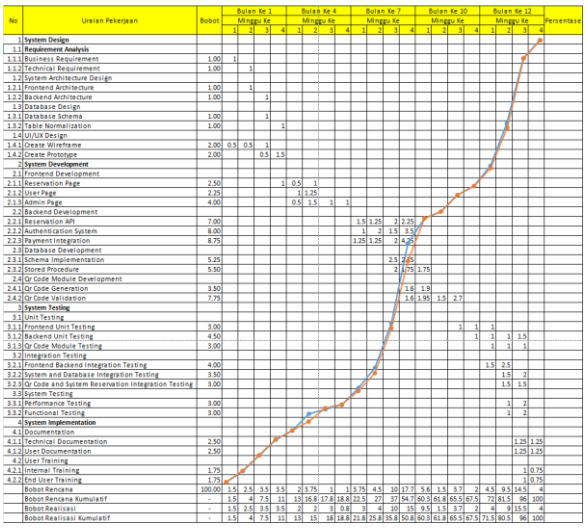


Fig. 7. S-Curve

- *Payback Period*

The payback period is the duration needed to recoup the initial investment in a project or investment through the generated cash inflows. It is determined by summing the annual cash inflows until the cumulative amount matches or surpasses the initial investment. A shorter payback period indicates a quicker recovery of the investment, which is typically preferred due to reduced risk and uncertainty.

TABLE X. PAYBACK PERIOD

Year	Cash Inflow(million)	Cash Outflow(million)
0	0	523.42
1	200	121
2	275	124
3	340	128
4	500	133

From the table above, we can see that the initial investment of 523.42 million will be returned between Year 2 and Year 3. The detailed calculation for Year 2 is as follows:

a) Cumulative total after Year 2 = 475 million.

b) The remaining amount needed to recover the initial investment of 523.42 million after Year 2 is 523.42 million - 475 million = 48.42 million.

c) The payback period occurs in Year 3, with an estimate of 2 years and 1.7 months.

- *Net Present Value (NPV)*

Net Present Value (NPV) is a financial analysis method used to determine the present value of a project's, investment's or accounting future cash flows for the initial investment cost and the discount rate. NPV helps assess whether a project or investment will yield a positive return after considering the cost of capital and associated risks. Below is the NPV table for a parking system.

TABLE XI. NPV

Year	Cas Inflow(million)	Discount Factor(10%)	Discounted Cash Flow(million)
0	0	1.00	-523.420
1	200	0.909	71.71
2	275	0.826	124.68
3	340	0.751	159.47
4	500	0.683	250.75
Total			618.54
NPV			95.12

With NPV value > 0, which is Rp. Rp 95,120,000 the investment for the parking reservation system project is declared feasible.

- *Return Of Investment (ROI)*

Return on Investment (ROI) is a metric that evaluates the performance of an investment or compares the efficiency of different investments. It calculates net profit in relation to the initial investment cost.

Description	Amount(million)
Total Cash Inflow	815
Initial Investment Cost	523.42
Net profit	291.58
ROI	55.7%

F. Project Human Resource Management

Project Human Resource Management is an important aspect of project management that focuses on managing human resources in an effective manner to ensure the project team works efficiently and achieves its goals. The main aim is to guarantee that every team member possesses the essential skills, knowledge, and motivation to execute the project as planned. For the parking reservation system project, the following considerations are essential:

a) Team Organization: The project requires a team comprising at least 1 Project Manager, 1 System Analyst, 3 Backend Developers, and 1 Frontend Developer.

b) Selection: Choosing the best candidates based on their skills, experience, and fit with the project's needs.

c) Motivation: Identifying what motivates team members and fostering a supportive work environment.

d) Effective Communication: Ensuring clear and open communication channels among all team members and stakeholders.

e) Conflict Management: Addressing and resolving any conflicts within the team in a constructive way.

G. Project Communication Management

Project Communication Management is a vital aspect of project management that focuses on planning, executing, and controlling effective communication among all parties involved in the project. The primary aim is to ensure that crucial

information is conveyed clearly, timely, and to the appropriate recipients. Effective communication enhances collaboration, minimizes the risk of misinformation, and ensures that all stakeholders have a consistent understanding of the project's status and objectives. For this project, the communication methods utilized by the team include WhatsApp, Google Meet, and offline meetings, with all communications occurring within a single chat group that includes 6 team members and 1 stakeholder.

H. Project Risk Management

Project Risk Management in a parking reservation system using QR includes the process of identifying, analyzing, and responding to risks that can affect the success of the implementation and operation of the system. With effective risk management, the project can minimize negative impacts and maximize the chances of success. Here are some stages in the parking reservation system:

- *Risk Identification*
  a) Technology: Risks associated with the QR system, including potential failures in QR scanning, connectivity issues, or device malfunctions.
  b) Security: Risks concerning data security, such as potential breaches of user information, cyberattacks, or misuse of data.
  c) Operational: Risks in daily operations, including reservation errors, irregular queues, or system failure under high user volumes.
  d) Stakeholders: Risks affecting user satisfaction, whether drivers using the parking service or parking management personnel.
  e) Finance: Financial risks, such as exceeding budgeted implementation costs, generating lower-than-expected revenue, or failing to meet financial targets.
  f) Regulatory: Risks related to compliance with relevant regulations and standards, including traffic laws, data protection, and information security.
- *Risk Analysis*
  a) Qualitative: Evaluating and ranking risks according to their possible impact and probability of happening. For instance, a data security risk might be categorized as high if user data is extremely sensitive.
  b) Quantitative: Evaluating the financial or operational consequences of specific risks, such as estimating the potential financial loss from a data breach.
- *Risk Response Planning*
  a) Avoid: Eliminating the risk, such as by strengthening security measures to prevent data breaches.
  b) Reduce: Minimizing the impact or probability of the risk, for example, by performing thorough system testing before deployment to reduce the chance of technical failures.
  c) Accept: Accepting the risk if its impact is minor or the cost of mitigating it outweighs the potential impact.
  d) Transfer: Shifting the risk to another party, such as through insurance or a contract with a dependable technology provider.
- *Risk Monitoring And Control*
  a) Continuous Monitoring: Keeping track of identified risks and staying alert for any new risks that may arise during the project.
  b) Corrective Action: Taking immediate steps to address the risk if it materializes, including adjusting the project plan or applying pre-established mitigation strategies.
  c) Risk Reporting: Providing regular updates on risk status to project management and stakeholders.
- *Implementation of Project Risk Management in QR Parking Reservation System:*

TABLE XII. PROJECT RISK MANAGEMENT

Risk Type	Identification	Analysis	Response	Monitoring
Technology Risk	Qr Scanning failure	High impact, medium likelihood	Implementing redundancy systems and extensive testing on various devices	Regularly checking system performance and ensuring technical support availability
Security Risk	Leakage of user data	Very high impact, medium probability	Employing data encryption, firewalls, and performing regular security audits	Monitoring system activities for suspicious behavior and providing regular reports
Operational Risk	Errors in reservation processes	Medium impact, high probability	Designing an intuitive user interface and training staff to manage errors effectively	Conducting user surveys to identify and resolve operational issues
Financial	Cost	High	Use	Revalidate

Risk	overruns	impact, medium probability	PERT Cost Analysis and Monte Carlo Simulation to develop a refined cost plan and estimate uncertainties	cost assumptions regularly and review simulation outputs to adjust financial plans

I. Project Stakeholder Management

Project Stakeholder Management involves identifying, analyzing, and managing interactions with all individuals and groups affected by or involved in a project. For a QR parking reservation system, this process aims to ensure that the needs and expectations of all stakeholders are clearly understood and addressed to achieve the project's success. Implementation of Project Stakeholder Management in a QR Parking Reservation System:

a) Project Owner: PT. Sanggraha Daksamitra.

b) Project Manager: Yostian Ari Sujarwo

c) End User: (Driver)

d) Parking Manager: PT. Sanggraha Daksamitra.

e) Technology Vendor: PT. Gama Solusi Internasional.

J. Project Closure

Project closure is the concluding phase of the project management cycle, ensuring that every aspect of the project is finalized, deliverable are approved by stakeholders, and all documentation is completed and archived appropriately. The detailed steps for closing a QR reservation system project are as follows:

- *Project Work Finalization*

a) Task Completion Verification: Confirm that all project tasks and deliverable have been completed as outlined in the project plan.

b) Final System Test: Perform final testing to ensure the QR parking reservation system meets all specified requirements.

- *Project Acceptance*

a) Stakeholder Approval: Secure final approval from all key stakeholders, confirming that the project has achieved the established success criteria.

b) Final Documentation: Prepare and finalize documentation, including the test report, user manual, and technical documentation.

- *Project Handover*

a) Training and Knowledge Transfer: Provide training to end users, such as parking operators and users, ensuring they understand how to use the system effectively.

b) System Handover: Officially transfer the QR parking reservation system to the responsible operational department or parking operator.

- *Project Evaluation*

a) Project Performance Review: Assess the project's performance by making comparisons achieved results with the original plan and objectives.

b) Lessons Learned: Document lessons learned from the project, including both successes and areas for improvement, for future reference.

- *Documentation and Filing*

a) Documentation Completion: Ensure that all project documents, such as contracts, progress reports, and meeting minutes, are complete and properly filed.

b) Final Project Report: Compile a final report summarizing project activities, results, and evaluations.

- *Project Team Disbandment*

a) Rewards and Recognition: Acknowledge and reward project team members for their contributions to the project.

b) Team Member Transitions: Facilitate the transition of project team members to new roles or back to their previous positions.

- *Closure Communication*

a) Project Closure Announcement: Officially announce the completion of the project to all stakeholders.

b) Final Report Distribution: Distribute the final project report to relevant stakeholders.

IV. CONCLUSION

This study showed that using a structured project management approach based on the PMBOK framework improved cost efficiency by providing clear task definitions, scheduling, and proactive risk control. These elements helped the project stay on track and reduced resource waste.

PERT and Monte Carlo Simulation effectively handled cost uncertainty by offering probabilistic estimates rather than fixed values. Their application led to a significant reduction in projected human resource costs from Rp748 million to Rp403 million with 95% confidence. These findings align with recent research [23], [24], which demonstrated that

simulation-based methods enhance cost estimation and risk management in complex projects.

As a recommendation for future researchers or project managers, it would be beneficial to apply PERT Cost Analysis and Monte Carlo Simulation to the entire project cost, not just human resource expenses. This broader application could offer even more accurate financial planning and reduce the risk of budget overruns across all project components.

REFERENCES

[1] Y. A. Sujarwo and A. Ratnasari, "Aplikasi Reservasi Parkir Inap Menggunakan Metode Fishbone Diagram dan QR-Code," Jurnal Sisfokom (Sistem Informasi dan Komputer), vol. 9, no. 3, pp. 302–309, Aug. 2020, doi: https://doi.org/10.32736/sisfokom.v9i3.808.

[2] M. S. Farooq, A. Hamid, A. Alvi, and U. Omer, "A review on Blended Learning Models, Curricula, and Gamification in Project Management Education," IEEE Access, pp. 1–1, 2022, doi: https://doi.org/10.1109/access.2022.3180355.

[3] C. T. Cerdeiral and G. Santos, "Software project management in high maturity: A systematic literature mapping," Journal of Systems and Software, vol. 148, pp. 56–87, Feb. 2019, doi: https://doi.org/10.1016/j.jss.2018.10.002.

[4] A. Abujraiban and G. J. Assaf, "Effect of Strategic Planning of Human Resources in Management Performance," Civil Engineering Journal, vol. 8, no. 8, pp. 1725–1738, Aug. 2022, doi: https://doi.org/10.28991/cej-2022-08-08-014.

[5] Lebedeva, A. and Guseva, A. (2020) 'Managing it projects and evaluating their cost and complexity: State of the problem', 2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), pp. 677–680. doi:10.1109/summa50634.2020.9280649.

[6] Aranyossy, M., Blaskovics, B. and Horváth, Á.A. (2017) 'How universal are IT project success and failure factors? evidence from Hungary', Information Systems Management, 35(1), pp. 15–28. doi:10.1080/10580530.2017.1416943.

[7] Foroutan Mirhosseini, A. et al. (2022) 'Sustainable Project Management: Reducing the risk of cost inaccuracy using a PLS-SEM approach', Sustainability, 14(2), p. 960. doi:10.3390/su14020960.

[8] Ning Nan and Harter, D.E. (2009) 'Impact of budget and schedule pressure on software development cycle time and effort', IEEE Transactions on Software Engineering, 35(5), pp. 624–637. doi:10.1109/tse.2009.18.

[9] K. P. Sulej, "Organizational Culture and Project Management methodology: Research in the Financial Industry," International Journal of Managing Projects in Business, vol. 14, no. 6, Mar. 2021, doi: https://doi.org/10.1108/ijmpb-08-2020-0252.

[10] [A. Shimoda, P. Wilairath, and T. Kounosu, "Project Planning from the viewpoint of Project Management and Systems Engineering," 2019 IEEE International Conference on Engineering, Technology and Education (TALE), Dec. 2019. doi:10.1109/tale48000.2019.9226007

[11] H. Y. Chiang and B. M. Lin, "A decision model for Human Resource Allocation in project management of software development," IEEE Access, vol. 8, pp. 38073–38081, 2020. doi:10.1109/access.2020.2975829

[12] F. C. Azenha, D. A. Reis, and A. L. Fleury, "The Role and Characteristics of Hybrid Approaches to Project Management in the Development of Technology-Based Products and Services," Project Management Journal, vol. 52, no. 1, p. 875697282095688, Oct. 2020, doi: https://doi.org/10.1177/8756972820956884.

[13] S. Alsaqqa, S. Sawalha, and H. Abdel-Nabi, "Agile Software Development: Methodologies and Trends," International Journal of Interactive Mobile Technologies, vol. 14, no. 11, pp. 246–270, Jul. 2020, Available: https://pdfs.semanticscholar.org/2fef/154748093288894dbd0b98db1b9b54731c71.pdf

[14] K, D.K. et al. (2024) 'A novel LSTM-based approach to PERT analysis in construction project management', 2024 Second International Conference on Advances in Information Technology (ICAIT), pp. 1–6. doi:10.1109/icait61638.2024.10690559.

[15] Ahraz Javeed Reshi, "Analysis and Evaluation of (beta) Beta Distribution or Three Point Estimation for PERT Technique in Project Management," International Journal of Science and Research (IJSR), vol. 12, no. 9, pp. 1736–1738, Sep. 2023, doi: https://doi.org/10.21275/sr23922120800.

[16] Orumie Ukamaka, C. (2020) 'Implementation of Project Evaluation and Review Technique (PERT) and critical path method (CPM): A comparative study', International Journal of Industrial and Operations Research, 3(1). doi:10.35840/2633-8947/6504.

[17] Akpan, N.P. and Agadaga, G.O. (2020) 'Modelling building renovation using pert', Asian Research Journal of Mathematics, pp. 25–38. doi:10.9734/arjom/2020/v16i430184.

[18] Zhang, S. and Jin, L. (2020) 'Research on Software Project Schedule Management Method based on Monte Carlo Simulation', 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), pp. 1605–1608. doi:10.1109/itoec49072.2020.9141570.

[19] Bakar, N.A. (2019) 'Monte Carlo simulation for data volatility analysis of stock prices in Islamic Finance for Malaysia composite index', International Journal of Advanced Engineering Research and Science, 6(3), pp. 6–12. doi:10.22161/ijaers.6.3.2.

[20] Tokdemir, O.B., Erol, H. and Dikmen, I. (2019) 'Delay risk assessment of repetitive construction projects using line-of-balance scheduling and Monte Carlo Simulation', Journal of Construction Engineering and Management, 145(2). doi:10.1061/(asce)co.1943-7862.0001595.

[21] Panqing, Y. et al. (2023) 'Human health-risk assessment of heavy metal–contaminated soil based on Monte Carlo Simulation', Scientific Reports, 13(1). doi:10.1038/s41598-023-33986-3.

[22] M. R. Bintang, K. Rossa Sungkono, and R. Sarno, "Time and cost optimization in feasibility test of CCTV project using CPM and pert," 2019 International Conference on Information and Communications Technology (ICOIACT), Jul. 2019. doi:10.1109/icoiact46704.2019.8938466

[23] J. Huang, A. Ji, Z. Xiao, and L. Zhang, Schedule-cost optimization in high-rise buildings considering uncertainty, Engineering, Construction and Architectural Management, Jun. 2024. doi: 10.1108/ECAM-12-2023-1217.

[24] M. Hu, E. K. Nippala, K. Kallioharju, and S. Pelsmakers, "Monte Carlo simulation approach to understand the cost variance for energy retrofit projects: comparative study of Finland and the United States," Construction Management and Economics, vol. 40, no. 6, pp. 207–222, Feb. 2022, doi: 10.1080/01446193.2022.2034906.

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, Ultima Infosys : Jurnal Ilmu Sistem Informasi 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: **ultimainfosys@umn.ac.id**

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

ISSN 2085-4579

Universitas Multimedia Nusantara
Scientia Garden Jl. Boulevard Gading Serpong, Tangerang
Telp. (021) 5422 0808 | Fax. (021) 5422 0800