ISSN 2085-4552

# ULTIMATICS

Jurnal Teknik Informatika



Engineering | Vol. 15 | No. 1 | Page 1 - 78 | June 2023 | ISSN 2085-4552

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

ISSN 2085-4552

# EDITORIAL BOARD




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

# EDITORIAL ADDRESS

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

# 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 2085-4552

# FOREWORD

ULTIMA Greetings!

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

In this June 2023 edition, Ultimatics enters the 1st Edition of Volume 15. In this edition there are ten scientific papers from researchers, academics and practitioners in the fields of Computer Science and Informatics. Some of the topics raised in this journal are: Implementation of AHP and Topsis Algorithm on Web-based Application Design to Determine Prospective Winners of Betta Fish Contest, Estimated Value of Software Developer Productivity at the Software Implementation Stage Using Function Points, Comparative Analysis of Phishing Tools on Social Media Sites, Web-based Writing Learning Application of Basic Hanacaraka Using Convolutional Neural Network Method, Gamification To Study Machine Learning Using Octalysis Framework, Methods of Stealing Personal Data on Android Using a Remote Administration Tool with Social Engineering Techniques, Development of A Mobile-Based Student Grade Processing Application Using the Waterfall Method, Analysis of User-generated Content in Visitor Reviews of Tourist Attractions Using Semantic Similarity, E-Commerce System for Media Group Cooperative, and Comparative Study of Robot Framework and Cucumber as BDD Automated Testing Tools.

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 ultimatics@umn.ac.id and the webpage of our Journal here.

Finally, we would like to thank all contributors to this June 2023 Edition of Ultimatics. We hope that scientific articles from research in this journal can be useful and contribute to the development of research and science in Indonesia.

June 2023,

**M.B.Nugraha, S.T., M.T.**
Editor-in-Chief

ISSN 2085-4552

# TABLE OF CONTENT

ISSN 2085-4552

# Implementation of AHP and Topsis Algorithm on Web-based Application Design to Determine Prospective Winners of Betta Fish Contest

Muhammad Abdurraffi[1], Fenina Adline Twince Tobing[2]
[1,2]Program Studi Informatika, Fakultas Teknik dan Informatika Universitas Multimedia Nusantara
[1]muhammad.abdurraffi@student.umn.ac.id, [2]fenina.tobing@umn.ac.id

Accepted 20 June 2022
Approved 09 May 2023

***Abstract*— The procurement of beauty competitions from ornamental betta fish has been held throughout Indonesia. The concept of judging and assessment carried out in the ornamental betta fish beauty competition still uses manual bookkeeping. The assessment with the national scale (SNI) is still a concern for the competition participants because of the lack of knowledge about the standard for judging ornamental betta fish contests. DSS (Decision Support System) can be an application development solution to determine the candidate for the winner of the ornamental betta fish contest. The criteria used include color value, neatness value, proportion value, mental value, and appearance value. The calculation method used in conducting this research is AHP and TOPSIS methods. The recommendations made will go through the AHP process for weighting based on the SNI assessment standard, while TOPSIS is used to receive alternatives from users. Then the ranking is based on the preference value generated by the TOPSIS method. This study has obtained the results of the application of the AHP and TOPSIS methods to make a recommendation system for contestants for winning ornamental betta fish correctly. Testing the level of user satisfaction obtains satisfaction results measured through the EUCS (End User Computing Satisfaction) dimension with the help of a Likert Scale calculation, so that it gets a satisfaction value of 87.89%.**

***Index Terms*— AHP; Betta Fish; DSS; Web Aplication; TOPSIS**.

## I. INTRODUCTION

Betta fish are freshwater fish originating from various countries or regions in Southeast Asia that have unique shapes and characteristics and are more aggressive in defending their territory. Betta fish are generally divided into three groups, namely ornamental betta, complaint betta, and wild betta [1]. The procurement of beauty competitions from betta fish has been held throughout Indonesia. The scoring system of the betta fish beauty competition uses the SNI (Standar Nasional Indonesia) standard [2] and IBC (International Betta Congress) [3]. However, the judging that is carried out in every betta fish beauty competition is still done manually by each organizer, this makes the assessment time inefficient and tends to be slow to win the ornamental betta fish contest [4]. Then, for contestants who have just joined the contest, they tend to have concerns about the competition scoring system due to lack of knowledge about the standards for judging ornamental betta fish contests [5].

In designing an application to determine the best betta fish candidates, a decision support system is needed to get an accurate evaluation of the winners of the betta fish beauty competition. DSS (Decision Support System) is a system that is used as a problem solving tool to assist decision making that will be considered [6]. Then, we need a method that can convert weights and calculations to get ranking results. The Analytic Hierarchy Process (AHP) method is a decision making method that is objective and subjective and can be used to determine the weight of the criteria [7]. Meanwhile, the TOPSIS method is one of the best methods in making decisions with many alternative choices [8].

AHP and TOPSIS are methods that are widely used in decision making. Previous research concluded that the AHP and TOPSIS methods were able to provide an alternative with an ideal solution distance value. The combination of the two methods can produce a more objective ranking result, so that it has a better recommendation quality [9]. The second study concluded that based on the calculation of the AHP-TOPSIS method on student data in 2015 it had an accuracy of 81% better than the calculation of the PROMETHEE method on student data in 2015 which had an accuracy of 70% [10]. The latest research concludes that the system built can determine the best supplier based on the supplier who has the highest weight, in this study PT Global Fiberindo has the highest weight value of 0.472 [11].

## II. METHODS

### A. *Web Based Application*

Web-based applications are applications that can be accessed via an internet browser, the use of web-based applications has been widely used by people in this century. The advantage of web-based applications is that there is no need to get a license to develop it, this is

because the license is owned by the application service provider. Then, in the development of web-based applications, they tend to only require standard system specifications and do not really require high specifications. Web-based applications can also be accessed anywhere without the need to do the installation process into the user's device. The operating system is also not an obstacle in making web-based applications, because web based applications have no limits on the operating system owned by the user's device. Web-based applications can also be accessed through various electronic devices such as computers, smartphones, laptops, and tablets [13].

### *B. Betta Fish*

Betta fish are freshwater fish originating from various countries or regions in Southeast Asia that have unique shapes and characteristics and are more aggressive in defending their territory. Betta fish are generally divided into three groups, namely ornamental betta, complaint betta, and wild betta. They has a unique shape and characteristics and tends to be aggressive in defending its territory. Among fans, hickeys are usually divided into three categories, namely ornamental hickeys, complaint hickeys, and wild hickeys. In Indonesia itself there are native races of betta fish, one of which is Betta channoides found in Pampang, East Kalimantan [1].

Betta fish consist of 73 species and are divided into 13 groups. Of the 73 species on earth, the betta fish species circulating in the market on average come from the splendens complex group, which consists of betta splendens, betta stiktos, betta mahachai, betta smaragdina and betta imbellis, as well as variants of crosses from these species. the. Betta fish fans divide the betta fish group 3 types, e.g ornamental, complaint, and wild [14].

In the ornamental betta fish contest there are 5 criteria to be assessed, namely color, tidiness, proportion, mentality, and appearance. Color categorization is based on looking at the colors on the betta fish's body parts [2]. Then, the categorization based on the neatness seen from the neatness of all parts of the betta fish body includes, body/head, anal fin, dorsal fin, pelvic fin, and tail fin. Neatness is assessed based on the detail aspects of ornamental betta fish, for example there is no bend in each fin bone, all scales are perfectly fused and nothing is loose, there are no fine hairs on the fish, and so on [2]. Categorization based on proportions seen from the balance between the body, tail, and fins of ornamental betta fish [2]. Categorization based on appearance is based on looking at various aspects, ranging from the size of the betta fish, the health condition of the betta fish, the style/behavior of the fish when testing, overall color, fish mentality, and overall tidiness [2].

Fig. 1 is an example of an ornamental betta fish that won the halfmoon class. This fish belonging to Alexander Chandra managed to get three titles, namely Best of Division, Best of Show, and best in the single tail category. The halfmoon fish has brooders who have won before. Genetic factors are also the cause of betta fish getting good genes from their parents [15].

Fig. 1. Betta Fish Contest Winner for Single Tail Halfmoon Category [15]

### *C. Decision Support System*

The concept of a decision support system is characterized by computer-based interactive systems that help decision makers use data and models to solve unstructured problems. DSS combines data and models into one part and is designed to assist managers in the decision-making process of semi-structural problems. DSS provides support for the manager's judgment not to replace the manager's role. DSS works by combining models and analysis techniques by entering existing data to find the information contained therein. DSS is a computerized system that can collect and analyze data and synthesize it to produce comprehensive information reports. DSS can provide more informed decision making, timely problem solving, and improve the efficiency of problem handling or operations, planning and management [16].

### *D. Analytical Hierarchy Process*

Analytic Hierarchy Process (AHP) was developed in 1980 by Thomas L. Saaty in his book entitled Analytic Hierarchy Process. Analytic Hierarchy Process (AHP) is a decision making process that explains the evaluation factors and weighting factors in multi-factor conditions by doing pairwise comparisons (Pairwise Comparison) [17].

The AHP calculation process begins with defining the problem, determining the solution, and compiling the hierarchy as shown in Fig. 2, in this study the calculation hierarchy only reached level 2 [17].

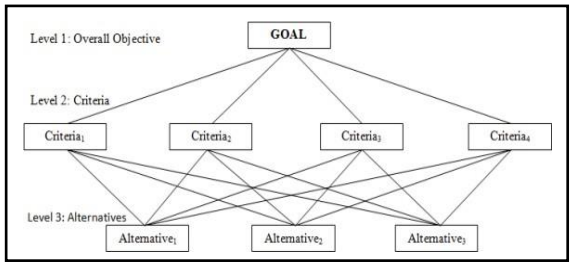


Fig. 2. Hierarchical Structure of AHP [18]

The second step is to determine the priority elements used by making comparisons in pairs, the comparison matrix uses numbers that represent the relative importance of one element to another [17].

TABLE I. PAIRWISE COMPARISON RATING SCALE [17]

| Weight | Distribution |
|---|---|
| 1 | Both element are equally important |
| 3 | One element is slightly more important than the other |
| 5 | One element is clearly more important than the other element |
| 7 | One element is more important than the other element |
| 9 | One element is absolutely important than the other element |
| 2,4,6,8 | Values between two adjacent consideration |
| Opposite | If activity i gets one point compared to activity j then j has the opposite value compared to activity i |

The third step is to determine the priority of the matrix by adding up the values in each column of the pairwise comparison matrix. Then, divide the column value by the total to get the result of matrix normalization. Then, add up the values of each row and then divide by the number of elements to get the average result [17].

The fourth step, multiply each value in the first column by the relative priority of the first element, multiply the value in the second column by the relative priority of the second element, and so on. Then, add up all the rows of the matrix. The number of rows is divided by the corresponding priority element. Add the quotient by the number of elements present and the result is called max [17].

The fifth step is to calculate the Consistency Index (CI). The CI calculation uses the max value minus the number of criteria, then divided by the number of criteria minus 1 as shown in (1) [17].

$$CI = \frac{maks-n}{n-1} \quad (1)$$

The fifth step is to calculate the Consistency Index (CI). The CI calculation uses the max value minus the number of criteria, then divided by the number of criteria minus 1 as shown in (1) [17].

$$CR = \frac{CI}{IR} \quad (2)$$

Index Random (IR) is a random index determined based on Table II. IR represents the value used in the calculation based on the number of criteria used in the study. The size of the matrix is the number of criteria used [17].

TABLE II. RANDOM CONSISTENCY INDEX LIST (IR) [17]

| Matrix Size | Random Index Value |
|---|---|
| 1,2 | 0.00 |
| 3 | 0.58 |
| 4 | 0.90 |
| 5 | 1.12 |
| 6 | 1.24 |
| 7 | 1.32 |
| 8 | 1.41 |
| 9 | 1.45 |
| 10 | 1.49 |
| 11 | 1.51 |
| 12 | 1.48 |
| 13 | 1.56 |
| 14 | 1.57 |
| 15 | 1.59 |

The final step is to check the consistency of the hierarchy. If the CR value is less than or equal to 0.1, then the calculation results can be said to be consistent [17]. The AHP method provides a solution to a fairly broad and unstructured problem by creating a model [19].

### *E. TOPSIS*

TOPSIS is a decision-making method with many standards, which can help solve various possibilities alternative problems and the best decision-making analysis process. In the analysis has a simple concept that is easy to understand. When performing calculations, this method is quite efficient in measuring the various alternative options available using a form of mathematical calculation that is simple and easy to calculate [8].

The first step is to make a normalization of the decision matrix. Then, weight the normalized matrix. Then, determine the positive and negative ideal solution matrices. Then, Determine the distance between the values of each alternative using a positive and negative ideal solution matrix. Then, determine the preference of the value of each alternative [8]. The calculation of the TOPSIS method for evaluating the performance of each alternative on each standard is normalized, as in (3) [8].

$$r_{ij} \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \quad (3)$$

$$y_{ij} = w_i r_{ij} \quad (4)$$

with i=1,2,…,m; and j=1,2,…,n.

$$A+ = (y1+, y2+, \dots, yn+) \quad (5)$$

$$A- = (y1-, y2-, \dots, yn-) \quad (6)$$

where $\boldsymbol{yj}$ + is the largest value if j is a profit attribute, while $\boldsymbol{yj}$ + is the smallest value if j is a cost

attribute. Then, $\boldsymbol{yj}$ − is the smallest value if j is a profit attribute, while $\boldsymbol{yj}$ − is the largest value if j is a cost attribute [8].

Formula for the distance between alternative $Ai$ and positive ideal solution (7)

$$D_i^+ = \sqrt{\sum_{j=1}^{n}\left(y_i^+ - y_{ij}\right)^2}; \quad i = 1,2,..,m \qquad (7)$$

Formula for the distance between alternative $Ai$ and negative ideal (8).

$$D_i^- = \sqrt{\sum_{j=1}^{n}\left(y_{ij} - y_i^-\right)^2}; \quad i = 1,2,..,m \qquad (8)$$

Formula for the preference value for each alternative ($Vi$) (9).

$$V_i = \frac{D_i^-}{D_i^- + D_i^+} \qquad (9)$$

A larger value of $Vi$ indicates that the alternative of $Ai$ is more used in the selection [8]. The TOPSIS method has an easy-to-understand concept, high computational efficiency, more efficient computational calculations and fast speed [20]. However, the drawback of the TOPSIS method is that there is no priority calculation that has become a standard, where the calculation is useful for increasing the effectiveness of the standard weighting calculation value. Therefore, for this reason, this method can be used in conjunction with, for example, the AHP method to produce maximum results or decisions [21].

### F. Likert Scale

The Likert scale has two forms of questions, namely the form of questions that are used to measure the positive questions of the positive vector table and the form of negative questions that are used to measure the negative vector tables. The scores for the positive questions were 5, 4, 3, 2, and 1; scores for negative questions are 1, 2, 3, 4, and 5. to find the total score and the maximum total score, you can use the calculation formulas as in (10) and (11) [22].

$$TSmax = Respondent \times Highest\ Score \qquad (10)$$

$$Total\ Score = Respondent \times Score \qquad (11)$$

The total score is used to find the score index value to find out the conclusions of the study using a Likert scale, the formula for calculating as in (12).

$$Index = (Total\ Score : Score\ Max) \times 100\% \qquad (12)$$

TABLE III. RATING INTERVAL [22]

| Interval | Description |
|---|---|
| 0%-19,99% | Strongly Disagree |
| 20%-39,99% | Disagree |
| 40%-59,99% | Undecided |
| 60%-79,99% | Agree |
| 80%-100% | Strongly Agree |

### G. EUCS

End User Computational Satisfaction (EUCS) is a method to measure user satisfaction of application systems by comparing expectations and reality of information systems. The EUCS evaluation model was developed by Doll & Torkzadeh. Evaluation using this model can emphasize user satisfaction by evaluating the content, accuracy, format, timing and ease of use of the system [24]. The model uses five criteria that are used to measure user satisfaction.

### H. IBM SPSS Statistics 26

IBM SPSS Statistics is a software capable of performing almost all types of data analysis used in the social sciences, natural sciences, or in the business world. SPSS and IBM have created a program that is user friendly and also powerful in performing statistical calculations. SPSS is capable of performing almost all types of statistical analysis ever used in the social sciences, business, and other sciences [26]. To use the IBM SPSS Statistics software requires experience in the field of statistics or being in such a learning process. Basic knowledge in statistics is important to understand so that the procedures used can produce the output that the user wants [26].

## III. RESULTS AND DISCUSSION

### A. AHP Process

The initial process in the AHP method is to determine the criteria used to perform calculations as in Table IV.

TABLE IV. TABLE CRITERIA

| Code | Description |
|---|---|
| C1 | Color |
| C2 | Neatness |
| C3 | Proportion |
| C4 | Mental |
| C5 | Appearance |

The weights of these criteria are taken to create scores on the pairwise comparison rating scale. The comparison value between criteria using a pairwise comparison rating scale is made based on the difference between the weights of the criteria. The difference used is based on the difference between the criteria held by the SNI assessment weight of the Indonesian Betta Ornamental Society (MCHI) with the smallest difference of 0 points and the largest difference of 40 points. The comparison can be seen in Table V.

TABLE V. DETERMINATION OF VALUES BASED ON PAIRWISE COMPARISON RATING SCALE

| Difference | Scale | Description |
|---|---|---|
| 0 | 1 | Equally important |
| 5 | 2 | Values between equally important and slightly more important |
| 10 | 3 | A little more important |
| 15 | 4 | Value between slightly more important and definitely more important |
| 20 | 5 | Obviously more important |
| 25 | 6 | The intermediate value is clearly more important and more important |
| 30 | 7 | More important |
| 35 | 8 | Value between more important and absolutely important |
| 40 | 9 | Absolute importance |

The results of determining the interests obtained based on user input from the selected class and sub class will produce a scale value determined by the system. The scale values form a pairwise comparison matrix. The matrix is made based on comparisons between criteria as in Table VI.

TABLE VI. CRITERIA PAIRED COMPARISON MATRIX

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| **C1** | 1 | 2 | 2 | 5 | 0,5 |
| **C2** | 0,5 | 1 | 1 | 4 | 0,33 |
| **C3** | 0,5 | 1 | 1 | | 0,33 |
| **C4** | 0,2 | 0,25 | 0,25 | 1 | 0,167 |
| **C5** | 2 | 3 | 3 | 5,988 | 1 |
| **Total** | 4,2 | 7,25 | 7,25 | 19,98 | 2,327 |

Then the matrix normalization process is carried out by adding up each value in each column and dividing it by the value in each column. The calculation will produce a criterion normalization matrix which can be seen in Table VII.

TABLE VII. CRITERIA NORMALIZATION MATRIX

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| **C1** | 0,238095 | 0,275862 | 0,275862 | 0,25015 | 0,214869 |
| **C2** | 0,119408 | 0,137931 | 0,137931 | 0,20012 | 0,141813 |
| **C3** | 0,119408 | 0,137931 | 0,137931 | 0,20012 | 0,141813 |
| **C4** | 0,047619 | 0,034483 | 0,034483 | 0,05003 | 0,071766 |
| **C5** | 0,47619 | 0,413793 | 0,413793 | 0,299581 | 0,429738 |

The results of the search for the value of the criterion weight (W) can be seen in Table VIII.

TABLE VIII. CRITERIA WEIGHT TABLE

| Criteria | W |
|---|---|
| C1 | 0,2510 |
| C2 | 0,1474 |
| C3 | 0,1474 |
| C4 | 0,0477 |
| C5 | 0,4066 |

Then in Table IX is a table of the results of matrix multiplication between pairwise comparison matrices with the value of the criterion weight (W). The result of the multiplication produces the value of max. The value of max is used to find the t value which is the reference for the Consistency Index (CI) value search.

TABLE IX. LAMBDA MAX CALCULATION TABLE

| Criteria | W | Lambda Max |
|---|---|---|
| C1 | 0,2510 | 1,2821 |
| C2 | 0,1474 | 0,7451 |
| C3 | 0,1474 | 0,7451 |
| C4 | 0,0477 | 0,2395 |
| C5 | 0,4066 | 2,0782 |

Then the next step is to get the value of t. The t value is obtained by dividing each column max by each column of criterion weight (W) and adding, then dividing by the number of criteria elements.

$$t = \frac{1}{5} \times \left(\left(\frac{1,2821}{0,2510}\right) + \left(\frac{0,7451}{0,1474}\right) + \left(\frac{0,7451}{0,1474}\right) + \left(\frac{0,2395}{0,0477}\right) + \left(\frac{2,0782}{0,4066}\right)\right) = 5,0709$$

Then the next step is to calculate the Consistency index (CI). The step to get the CI is to subtract the t value with the number of criteria elements, then divide by the number of criteria values minus 1. To get the Consistency Ratio (CR) value is to divide the result of the CI value by the Random Index (IR). The IR used for the number of criteria elements as many as 5 criteria is 1.12.

$$CI = \frac{5,0709-5}{5-1} = 0,017731 \quad (14)$$

$$CR = \frac{0,017731}{1,12} = 0,015832 \quad (15)$$

The CR value obtained is 0.015832, this value is below 0.1 which is a requirement so that the weight of the criteria can be used. Then the weight of the calculated criteria is considered valid. If the CR value is above 0.1, the calculation must be repeated so that the CR value can meet the requirements.

### *B. TOPSIS Process*

TOPSIS calculation is carried out when the validation of the CR value has been fulfilled in the AHP calculation. The first stage is to make a decision matrix based on the input file from the user regarding the assessment of betta fish. The determined rating scale is set with a rating of 1-100, the value is obtained from the user via file uploads in the application. The decision matrix is shown in Table X.

TABLE X. DECISION MATRIX

| Serial Number | Criteria | | | | |
|---|---|---|---|---|---|
| | C1 | C2 | C3 | C4 | C5 |
| A1 | 23 | 34 | 90 | 65 | 54 |
| A2 | 32 | 33 | 87 | 86 | 53 |
| A3 | 43 | 86 | 32 | 42 | 56 |
| A4 | 12 | 57 | 32 | 42 | 67 |
| A5 | 43 | 67 | 64 | 45 | 12 |
| A6 | 23 | 86 | 53 | 42 | 32 |
| A7 | 43 | 97 | 45 | 24 | 46 |
| A8 | 53 | 56 | 32 | 43 | 45 |
| A9 | 44 | 45 | 23 | 42 | 32 |
| A10 | 80 | 34 | 98 | 53 | 74 |

Then the step taken is to find the value of the normalized decision matrix. The calculation is done by

squaring each value, then adding all of them and the result of the sum is rooted, then each data is divided by the result of the rooting. The calculation results can be seen in Table XI.

TABLE XI. NORMALIZED DECISION MATRIX

| Serial Number | Criteria | | | | |
|---|---|---|---|---|---|
| | C1 | C2 | C3 | C4 | C5 |
| A1 | 0,1672 | 0,1689 | 0,4628 | 0,4036 | 0,3404 |
| A2 | 0,2326 | 0,1639 | 0,4474 | 0,5340 | 0,3341 |
| A3 | 0,3126 | 0,4273 | 0,1645 | 0,2607 | 0,3530 |
| A4 | 0,0872 | 0,2832 | 0,1645 | 0,2607 | 0,4224 |
| A5 | 0,3126 | 0,3329 | 0,3291 | 0,2794 | 0,0756 |
| A6 | 0,1672 | 0,4273 | 0,2725 | 0,2607 | 0,2017 |
| A7 | 0,3126 | 0,4819 | 0,2314 | 0,1490 | 0,2900 |
| A8 | 0,3853 | 0,2782 | 0,1645 | 0,2670 | 0,2837 |
| A9 | 0,3199 | 0,2236 | 0,1182 | 0,2607 | 0,2017 |
| A10 | 0,5816 | 0,1689 | 0,5040 | 0,3290 | 0,4665 |

Then the next process is to calculate the weighted normalized decision matrix. The calculation process uses the criterion weight value (W) which has been multiplied by a normalized decision matrix. The results of the calculation of the weighted normalized decision matrix can be seen in Table XII.

TABLE XII. WEIGHTED NORMALIZED DECISION MATRIX

| Serial Number | Criteria | | | | |
|---|---|---|---|---|---|
| | C1 | C2 | C3 | C4 | C5 |
| A1 | 0,0419 | 0,0248 | 0,0682 | 0,0192 | 0,1384 |
| A2 | 0,0583 | 0,0241 | 0,0659 | 0,0254 | 0,1358 |
| A3 | 0,0784 | 0,0629 | 0,0242 | 0,0124 | 0,1435 |
| A4 | 0,0218 | 0,0417 | 0,0242 | 0,0124 | 0,1717 |
| A5 | 0,0784 | 0,0490 | 0,0485 | 0,0133 | 0,0307 |
| A6 | 0,0419 | 0,0629 | 0,0401 | 0,0124 | 0,0820 |
| A7 | 0,0784 | 0,0710 | 0,0341 | 0,0071 | 0,1179 |
| A8 | 0,0967 | 0,0410 | 0,0242 | 0,0127 | 0,1156 |
| A9 | 0,0802 | 0,0329 | 0,0174 | 0,0124 | 0,0820 |
| A10 | 0,1459 | 0,0248 | 0,0742 | 0,0156 | 0,1897 |

Next is to find a positive ideal solution and a negative ideal solution. The calculation in finding these two values is to find the minimum and maximum values in each column of the weighted normalized decision matrix. The matrix of positive and negative ideal solutions can be seen in Table XIII.

TABLE XIII. MATRIX OF POSITIVE IDEAL SOLUTION AND NEGATIVE IDEAL SOLUTION

| | C1 | C2 | C3 | C4 | C5 |
|---|---|---|---|---|---|
| y+ | 0,1459 | 0,2510 | 0,0742 | 0,0254 | 0,1897 |
| y- | 0,0218 | 0,1474 | 0,0174 | 0,0071 | 0,0307 |

The next calculation is to find the distance of the positive ideal solution and the distance of the negative ideal solution. The process is done by reducing the value of the ideal solution by the value of each criterion in the weighted normalized decision matrix. The calculation results are then squared and then added together. The result of the sum is rooted to get the value of the distance of the positive ideal solution and the distance of the negative ideal solution. The calculation results can be seen in Table XIV.

TABLE XIV. IDEAL SOLUTION DISTANCE TABLE

| Serial Number | Ideal Solution Distance | |
|---|---|---|
| | C1 | C2 |
| A1 | 0,125098 | 0,121335 |
| A2 | 0,113291 | 0,122755 |
| A3 | 0,097078 | 0,132301 |
| A4 | 0,138734 | 0,14235 |
| A5 | 0,176391 | 0,069453 |
| A6 | 0,154297 | 0,071297 |
| A7 | 0,107985 | 0,115199 |
| A8 | 0,10733 | 0,114521 |
| A9 | 0,144078 | 0,078381 |
| A10 | 0,047156 | 0,209672 |

Then the next step is to find the preference value. The calculation process is to add up the negative ideal solution distance (S-) with the positive ideal solution distance (S+). Then each S- is divided by the sum. The calculation results to get the preference value can be seen in Table XV.

TABLE XV. PREFERENCE VALUE MATRIX

| Serial Number | Preference Value |
|---|---|
| A1 | 0,4923 |
| A2 | 0,5200 |
| A3 | 0,5767 |
| A4 | 0,5064 |
| A5 | 0,2825 |
| A6 | 0,3160 |
| A7 | 0,5161 |
| A8 | 0,5162 |
| A9 | 0,3523 |
| A10 | 0,8163 |

Based on Table XV, the preference value of the highest ornamental betta fish contest is the participant with serial number A10. It can be concluded that the winner of the ornamental betta beauty contest is the fish with serial number A10 with a preference value of 0.8163. In Fig. 3 is a ranking display that has been sorted by preference value which is used as a percentage of the assessment. All preference values are made in the form of a percentage so as to produce a percentage rating. The value of the application is slightly different from the manual calculation due to the rounding performed on the application. It can be seen and concluded that the AHP and TOPSIS calculations in Fig. 3 are not much different from Table XV.

| Appraisal Presentation |
|---|
| 81.67% |
| 57.73% |
| 52.06% |
| 51.63% |
| 51.63% |
| 50.74% |
| 49.30% |
| 35.22% |
| 31.61% |
| 28.18% |

Fig. 3. System Ranking Calculation Results

### C. System Evaluation Results

System evaluation is done by asking respondents who have participated in an ornamental betta fish competition to try the system. Then, respondents were asked to fill out a questionnaire consisting of 13 questions with 5 scales, namely strongly agree, agree, hesitate, disagree, and strongly disagree. The questions used are the End User Computing Satisfaction (EUCS) method. In the EUCS method, there are 5 dimensions of questions, namely, the dimensions of content (content), dimensions of display (format), dimensions of accuracy (accuracy), dimensions of timeliness (timelines), and dimensions of user ease (ease of use).

Based on the calculation of the five dimensions in the EUCS, the average percentage score can be calculated. The calculation can be seen as follows:

$$\boldsymbol{Final\ Percentage} = (88) + (88{,}89) + (84{,}89) + (87{,}67) + (90)5 = 87{,}89\%$$

The results of the final percentage calculation obtained 87.89%. It can be concluded that the respondents strongly agree with the satisfaction of the application to determine the winner of this betta fish contest.

### D. Questionnaire Validity Test

The validity test of the questionnaire is a test to test the results of the questionnaire on user satisfaction in the EUCS method. Questionnaire validation is used to test the level of validity of the questionnaire that has been filled out by the respondent. This validity test was carried out using IBM SPSS Statistics 26. The calculation results for the questionnaire validity test can be seen in Table XVI.

TABLE XVI. VALIDITY TEST CALCULATION RESULT

| Question | Score $r_{table}$ | Score $r_{count}$ | Description |
|---|---|---|---|
| 1 | 0,374 | 0,832 | Valid |
| 2 | 0,374 | 0,856 | Valid |
| 3 | 0,374 | 0,835 | Valid |
| 4 | 0,374 | 0,838 | Valid |
| 5 | 0,374 | 0,833 | Valid |
| 6 | 0,374 | 0,880 | Valid |
| 7 | 0,374 | 0,799 | Valid |
| 8 | 0,374 | 0,854 | Valid |
| 9 | 0,374 | 0,860 | Valid |
| 10 | 0,374 | 0,906 | Valid |
| 11 | 0,374 | 0,891 | Valid |
| 12 | 0,374 | 0,860 | Valid |
| 13 | 0,374 | 0,883 | Valid |

Each value of r_count is compared to the value of r_table. Then, if the value of r_count > the value of r_table, then the data is considered valid. On the other hand, if the value of r_count < the value of r_table, then the data is considered invalid. Based on the results of the calculations in Table XVI, it can be concluded that each item in the questionnaire is considered valid.

### E. Interface System Implementation

Figure 4 is part of Main Page and How to Use page where the user is given information regarding the first steps that must be taken to start running the application.

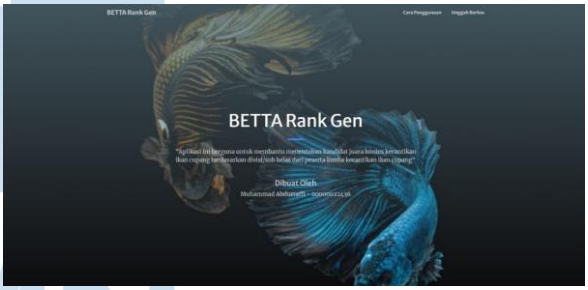


Fig. 4. Main Page

Figure 5 is the File Upload page for input betta fish class and their subclass. The betta fish class functions to provide information regarding which class is ranked, while the betta fish subclass functions to determine the weight of the betta fish species that are contested. The user must select both inputs to be able to proceed to the next stage.

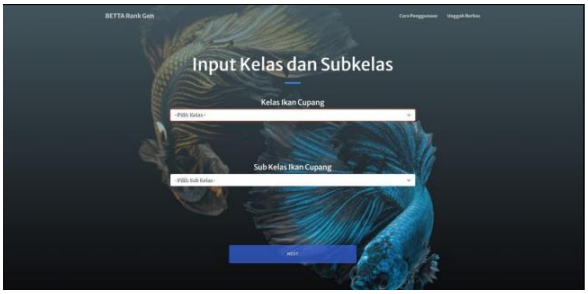


Fig. 5. Page File Upload

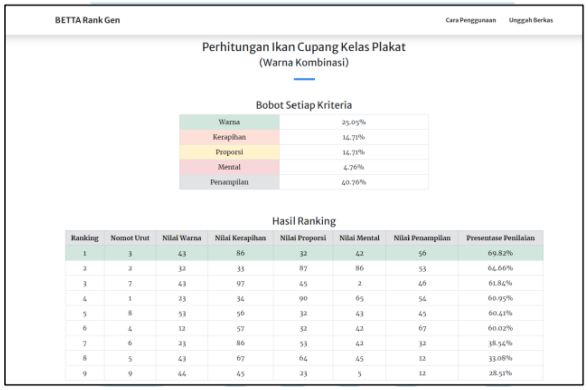


Fig. 6. Table of Calculation Results for Betta Fish Contest Champion Candidates

## IV. CONCLUSIONS

The decision support system for determining the winner of the betta fish contest using the web-based AHP and TOPSIS methods has been successfully created. The system can provide recommendations for betta fish contest winners based on five criteria determined by the system, namely color value, neatness value, proportion value, mental value, and appearance value. The data provided by the user in the form of betta fish values from each criterion can be processed properly by the system. The survey has been conducted on 30 respondents. The method used in the survey is End User Computing Satisfaction (EUCS) to conclude user satisfaction with the application. The results of the calculations carried out reached 87.89% as the final percentage of testing the EUCS method. The calculation is assisted by the Likert scale method which produces a very good percentage. Test the validity of the questionnaire on 13 questions that have been filled out by respondents. Calculations were carried out using IBM SPSS Statistics 26. The calculation results stated that all 13 questions were valid.

## REFERENCES

[1] KKP, "Badan Karantina Ikan, Pengendalian Mutu dan Keamanan Hasil Perikanan," 2011. [Online]. Available: http://www.bkipm.kkp.go.id/bkipmnew/ias/ias_dtl/2. [Accessed 19 Februari 2021].

[2] Masyarakat Cupang Hias Indonesia, "Penilaian Kontes Ikan Cupang Hias," STANDARD NASIONAL INDONESIA, pp. 1-47, 2012.

[3] Dewan Juri IBC, "Standard Penilaian Kontes Cupang berdasarkan International Betta Congress," STANDARD IBC, pp. 1-94, 2012.

[4] A. Y. Nugroho, "Sistem Pendukung Keputusan Dalam Menentukan Pemenang Kontes Ikan Cupang Hias Menggunakan Metode Gap Profile Matching," IT Journal, vol. 4, no. 1, pp. 99-108, 2016.

[5] M. A. Ashari, "Penentuan Juara Kontes Ikan Cupang Dengan Menggunakan Metode PROMETHEE," Artikel Skripsi Universitas Nusantara PGRI Kediri, pp. 2-8, 2019. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[6] Basri, "Metode Weighted Product (WP) Dalam Sistem Pendukung Keputusan Penerimaan Beasiswa Prestasi," Jurnal INSYPRO, pp. 1-6, 2017.

[7] E. Sulistiyani, M. I. H. Amir, Y. K. R, Nasrullah and D. Injarwanto, "Implementasi Metode Analytical Hierarchy Process (AHP) Sebagai Solusi Alternatif Dalam Pemilihan Supplier Bahan Baku Apel Di PT.Mannasatria Kusumajaya," Technology Science and Engineering Journal, vol. 1, no. 2, pp. 1-15, 2017.

[8] S. Murdowo, "Simulasi Sederhana Perhitungan Sistem Pendukung Keputusan Menggunakan Metode TOPSIS," Jurnal INFOKAM, vol. XVI, no. 2, pp. 129-134, 2020.

[9] P. P. Santika and I. P. S. Handika, "Sistem Pendukung Keputusan Penerimaan Karyawan Dengan Metode AHP TOPSIS (Studi Kasus: PT. Global Retailindo Pratama)," SINTECH JOURNAL, vol. 1, no. 1, pp. 1-9, 2018.

[10] Venendhie, R. Saptono and H. Setiadi, "Comparative Analysis Of AHP-TOPSIS Method and PROMETHEE Method In Determining Kartu Indonesia Pintar Receiver," ITSMART: Jurnal Ilmiah Teknologi dan Informasi, vol. 7, no. 2, pp. 94-101, 2018.

[11] F. A. T. Tobing, M. I. Dzulhaq and R. F. Sidiq, "Penerapan Metode Fuzzy AHP Untuk Sistem Pendukung Keputusan Pemilihan Pemasok Terbaik," ULTIMA Computing, vol. XI, no. 2, pp. 90-94, 2019.

[12] S. Enjelina and E. Insannudin, "PERANCANGAN APLIKASI BERBASIS WEB INTERAKTIF HALOAPP BERBASIS ANDROID DAN iOS," Jurnal Aplikasi Berbasis Web, vol. 10, pp. 1- 6, 2016.

[13] A. Hatmoko, "Pengertian Aplikasi Berbasis Web," 2019. [Online]. Available: https://www.kompasiana.com/komentar/andihatmoko/5b7eb431bde5 7534805e0e52/pengertian-aplikasi-berbasis-web. [Accessed 12 Agustus 2021].

[14] M. Sari, Hasanuddin and R. Aditya, "Pengenalan Ikan Cupang (Betta Fish) Menggunakan Augmented Reality," JTIULM, vol. 1, no. 1, pp. 26-36, 2018.

[15] D. Ginanjar, "Maaf, Cupang Halfmoon Marble Milik Alexander Chandra Not for Sale!," 2020. [Online]. Available: https://www.jawapos.com/hobi-kesenangan/08/11/2020/maaf_cupang-halfmoon-marble-milik-alexander-chandra-not-for-sale/. [Accessed 10 Oktober 2021].

[16] T. Limbong, Muttaqin, A. Iskandar, A. Windarto, J. Simarmata, Mesran, O. Sulaiman, D. Siregar, D. Nofriansyah, D. Napitupulu and A. Wanto, Sistem Pendukung Keputusan: Metode & Implementasi, 1 ed., Medan: Yayasan Kita Menulis, 2020.

[17] L. M. Yulyantari and I. P. Wijaya, Manajemen Model Pada Sistem Pendukung Keputusan, Yogyakarta: Andi, 2019.

[18] P. Agrawal, V. Mishra, M. Bag, M. Sahai and V. Singh, "Supplier Selection in Dynamic Environment using Analytic Hierarchy Process," International Journal of Information Engineering and Electronic Business, vol. 4, no. 10, pp. 20-26, 2014.

[19] R. G. U. Panuntun, "Analisis Pengambilan Keputusan Prioritisasi Proyek Pengambangan Blok Minyak Alih Kelola XYZ Menggunakan Metode AHP," Jurnal Universitas Pertamina, pp. 1-62, 2021.

[20] A. Mulio, "Perbandingan Metode TOPSIS dan Weighting Product Pada Sistem Pendukung Keputusan Seleksi karyawan Baru pada PT.Lotte Mart," Jurnal Universitas Potensi Utama, pp. 11-96, 2016.

[21] L. Mutmainah and Yunita, "Penerapan Metode TOPSIS Dalam Pemilihan Jasa Ekspedisi," Jurnal SISFOKOM (Sistem Informasi dan Komputer), vol. 10, no. 01, pp. 86-92, 2021.

[22] D. Kho, "Teknik Elektronika," 2020. [Online]. Available: https://teknikelektronika.com/pengertian-skala-likert-likert-scale_menggunakan-skala-likert/. [Accessed 24 Februari 2021].

[23] V. H. Pranatawijaya, Widiatry, R. Priskila and P. B. A. A. Putra, "Pengembangan Aplikasi Kuesioner Survey Berbasis

Web Menggunakan Skala Likert dan Guttman," Jurnal Sains dan Informatika, vol. 5, no. 2, pp. 128-137, 2019.

[24] N. Dalimunthe and C. Ismiati, "Analisis Tingkat Kepuasan Pengguna Online Public Access Catalog (OPAC) Dengan Metode EUCS (Studi Kasus: Perpustakaan UIN SUSKA Riau)," Jurnal Rekayasa dan Manajemen Sistem Informasi, vol. 2, no. 1, pp. 71-75, 2016.

[25] A. S. Restanti, E. Y. Astuti, Munjiati, U. Nurwijayanti and S. Widianingtias, "Analisis End-User Computing Satisfaction pada Online Public Access Catalogue “Izylib” di Lingkungan Universitas Jenderal Soedirman," PUSTABIBLIA: Journal of Library and Information Science, vol. 1, no. 1, pp. 223-238, 2017.

[26] D. George and P. Mallery, IBM SPSS Statistics 26 Step by Step: A Simple Guide and Reference, 6th ed., New York: Routledge, 2019.

ISSN 2085-4552

# Estimated Value of Software Developer Productivity at the Software Implementation Stage Using Function Points

Asri Maspupah[1], Lukmannul Hakim Firdaus[2], Wendi Wirasta[2]
[1,2,3] Jurusan Teknik Komputer dan Informatika, Politeknik Negeri Bandung, Bandung, Indonesia
[1] asri.maspupah@polban.ac.id, [2] lukmannul.hakim@polban.ac.id, [3] wendi.wirasta@polban.ac.id

Accepted 10 October 2022
Approved 28 November 2022

***Abstract*— Most Software Development Processes (SDP) project failures occur due to errors in estimating the cost, time, and effort during the planning stage. This happened because the planning still relied on the intuition and experience of the programmer. One approach that can be taken to plan the right SDP is to know the value of SDP productivity. The focus of this research was to determine the value of productivity based on the differences in programmers' skills. This case study was conducted to determine the productivity value of the web-based software that has been built, namely McDelivery. The productivity value was calculated based on the ratio of software size to effort. In this case, the software size was obtained by calculating the Application Function Point Count (APFC). Meanwhile, the effort was obtained from expert judgment to determine the time needed by the development team at the junior, middle, and senior software developers to implement software functionality in person-day to the form of program code. The result showed that the productivity value of SDP was directly proportional to the level of ability of the programmers. These productivity values can be used as a solution option to calculate the estimated time, cost, and even the availability of programmers that were adjusted to the conditions faced in planning software development.**

***Index Terms*— application function point count; function point analysis; productivity metrics; software development process; software developer**.

## I. INTRODUCTION

The accuracy of activity planning determines the success factor of the software development process (SDP) project at the beginning of a development which is directly proportional to the realization of the implementation of activities at the end of time [1]. This accuracy is indicated by the software meeting the requirements within a period and incurring reasonable and planned costs [2]. Most SDP project failures occurred due to incompatibility of planning with actual implementation. Generally, the cause of the discrepancy lies in the estimation of cost, time, and effort [3]. Meanwhile, the estimated effort is used to plan and calculate the software development costs.

Based on the research conducted by Usman in 2015 concerning the measurement of the accuracy of planning estimates with the implementation in agile software development through a survey of 60 companies showed that approximately 78.33% of the companies stated that there was an inaccuracy estimation between the implementation and the planning [4]. The inaccuracy estimation occurs when the implementation of software development exceeds the planning estimate (60%) or solves faster than the planning estimate (18.33%). One of the contributing factors is that software development planners need to think about the best scenario based on the development team's ability to deal with software complexity. In addition, planners also need to pay more attention to the software development effort.

Proper SDP planning can be done by considering realistic costs, time, and effort [5]. However, as SDP planners, IT project managers have difficulty in making accurate effort estimates [6]. In general, the calculation of each functional software's development effort depends on intuitive experience, which is a subjective assessment of the SDP planner. In this case, development effort is defined as person-hour working on several SDP activities [7]. At the same time, the determination of person-hours depends on the software developer's ability level and the work's difficulty [8].

The accuracy of SDP planning is related to productivity metrics because the value of productivity can measure the development process's effectiveness at the project's end [9]. In general, productivity is defined as the ratio between input and output. Input is a resource to produce output. This definition is very suitable for the manufacturing industry because it clearly shows the quality standard of input and output measurement units [8]. In the field of software engineering (SE), the term productivity refers to the effectiveness of development project efforts measured by the output rate per unit of software [10] [7].

The calculation of the SDP productivity value is generally done twice; those are at the beginning and the end of the project. The difference in productivity values shows the inaccurate prediction of project planning effort estimates. However, using historical SDP

productivity data from previous projects in similar software developments can increase the accuracy of project planning parameter estimates (effort, time, and cost) [9].

On the other hand, most IT projects have used productivity metrics to measure SDP productivity based on the comparison between software size (output) and effort (input) required in developing software [11]. In this case, the accuracy of the productivity value of the productivity metric must consider three criteria: (1) the scope of the resource; (2) the scope of input; and (3) the scope of the output to be calculated [12]. Thus, practitioners and academics must carefully know the scope of effort and software size that will be useful for measuring the SDP productivity.

Most previous researchers used time as the definition of effort in measuring SDP productivity. However, the definition of productivity in the SE concept should focus on the level of complexity of work done by each person [8]. Using time as a criterion for measuring productivity inputs can lead to the question, what time should be used? In this case, the time is the duration spent by the person doing a job or time paid (contract) person within the worked hour range. On the other hand, each person has a different level of ability, so the level of productivity in completing the work is also different. In the context of SE, software developers' ability can affect the productivity level in software development [8]. Meanwhile, software metrics can quantify software size as a productivity output criterion [2]. Software metrics as software measurement standards aim to get a value for the size of software complexity [13].

In previous research, the researchers proposed several software metrics, including line of code (LOC), constructive cost model (COCOMO), and function point analysis (FPA). LOC is the most straightforward software metric using the actual line code as a criterion [2]. However, LOC is so dependent on programming languages and development technologies that it cannot be used to measure the productivity of non-technical activities and is challenging to be measured at the beginning of development [14][2]. Furthermore, COCOMO uses a mathematical formula to determine software development efforts. COCOMO involves line code information and justification of development efforts by domain experts [15].

Meanwhile, FPA uses the requirement specification functionality as the basis for measurement. The FPA calculation uses the standard method to measure software engineering based on the scope of the software. Thus, the FPA calculation, regardless of technology and programming language, is more straightforward and meaningful from the end user's point of view [16] [15].

Among the various variations of software metrics, FPA is the most commonly used approach [5]. Alan J. Albrecht introduced FPA from IBM in 1979 [3]. At the end of the FPA measurement process, the software has a function point (FP) value. In this case, FP shows the value of software functionality as the basis for successful product delivery to end users [8]. Thus, FP is a unit of software size for software development analysis. For example, realistic cost estimation, measurement of SDP productivity value based on the ratio of effort spent on each FP, and measurement of software quality based on the ratio of the number of defects found in each FP [17].

Based on the explanation above, this study aimed to obtain the value of software developer productivity at the software implementation stage. The focus of the study was the analysis of effort and software size from the side of software developers with different levels of programming ability in the case of software development from similar applications that already exist. In this case, the SDP productivity measurement parameters were Effort and software size. The definition of the software developer's programming ability specifications is through the justification of domain experts, namely software developer experts with 12 years of experience as a team leader of software developers. Furthermore, the value of software development productivity is used as an indicator of time and cost estimation so that SDP planning becomes realistic. In addition, the productivity measurement parameters used time-person as input and FP as output. At the same time, the selection of FP is a unit of software size because it is more appropriate to use functional measurement software independent of technology to be used as a parameter for calculating SDP planning.

## II. METHODS

The research stage started with the analysis of problems in software coding productivity. The next stage was selecting a productivity measurement method—finally, the measurement of productivity in this case was carried out through an experiment. Figure 1 shows the research flow.

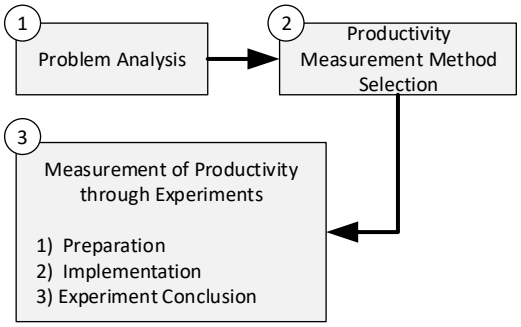


Fig. 1. The Research Stage

The productivity measurement method used productivity metrics. Then, the input and output terminology were mapped into the software

development process. Furthermore, the product output calculation used FPA, while software coding input used calculating effort (person-day) required by the developer. The productivity measurement experiment consisted of 3 stages. First, experimental preparation including preparing the case studies of products to be built and mapping the developer programming skills into three levels (junior, middle, and senior). Second, the implementation of experiments to calculate the input, output, and productivity of the implementation software stage. Third, the conclusion of the experimental results, which was to analyze the productivity value at each developer level.

### A. *Function Point Analysis (FPA)*

A measure of software complexity is the output produced during software development. Furthermore, the output was used as a parameter to generate productivity values by comparing the effort when carrying out activities. Software complexity becomes a number through software measurements. FPA is a method used to measure the complexity and functionality of software in SDP projects [2] [18].

The method used in FPA was dividing the size of the software into smaller components so that it was easier to analyze [3]. The critical value in the FPA measurement is the calculation of software function based on five standard functions set by the International Function Point Users Group (IPPUG), namely internal logical files (ILF), external logical files (EIF), external input (EI), external output (EO), and external inquiries (EQ) [5]. In this case, the FP calculation process started from a high level by analyzing the software functionality specifications. It consisted of six steps, namely (1) determining the type of FP count; (2) identifying the software scope; (3) weighing the software based on standard function; (4) calculating the unadjusted function point (UFP); (5) justifying the value adjustment factor (VAF), and (6) calculating the adjusted function point (AFP) [19] [3].

*Determining the type of FP Count*

The calculation of software size depends on the purpose of the type of software to be analyzed. In 2010, IFPUG divided the types of FP calculations into three categories. First, the development project function point count (DPFPC) is the type of FP count intended for software developed for the first time and released to end users. Second, enhancement project function count (EPFC) is the type of FP count intended for software developed in adaptive maintenance projects. Adaptive maintenance projects aim to improve performance and implement change requirements from end users that must be matched in the first stage of development. Third, application function point count (AFPC) is the type of FP count intended for software developed on existing software products. However, adding new functionality are needed when there is changing requirements from end users.

This study used APFC to calculate software size with the formula shown in equation 1.

$$AFP = ADD \times VAF \quad (1)$$

Information:

- AFP : *application project function point count*
- ADD : *unadjusted function point count from application functionality added to an existing application*
- VAF : *value adjustment factor*

*1) Identifying the scope and limitations of the software*

The scope and limitations of the software form the basis of software development. The scope defines a set of software functionality that includes data, screens, and reports. At the same time, the limitation of software as an interface between software and end users defines things outside the software's scope. Thus, determining the scope and limitations of the software can provide information on the size of the software based on functionality at the end of development so that the FP count can be carried out [20]. The steps for determining the scope and limitations of the software are:

*a)* defining a set of sub-processes within the scope of the software.

*b)* understanding the purpose of measuring the FP count on the software.

*c)* defining the software process flow in managing data into information.

*d)* defining business areas to support each process in the software.

*e)* defining logical data both within the scope of the software and logical data originating from outside the scope of the software.

*2) Identifying the scope and limitations of the software*

Unadjusted function points show the value of software complexity by weighing each functional requirement based on five standard functions [3] [2]. The weighting of software complexity consists of three categories, namely low (L), average (A), and high (H) [20]. Meanwhile, the five standard functions are divided into data and transactional [20] [5].

The data function shows software functionality on internal and external data storage requirements. Data functions include logical data in software (ILF) and external interface files that connect data from outside with internal software (EIF). The transactional function shows the software's functionality in processing data when there is an interaction between the software and the end user. The transaction process includes transactions receiving input data (EI), displaying output (EO), and querying data (EQ) [20].

The weighting of complexity in each standard function is conducted based on the functional software's RET, DET, and FTR values. In this case, RET (record

element type) is a subgroup of data elements in data storage, DET (data element type) is an attribute in data storage and application, and FTR (file type reference) is a type of data files read/managed in a transaction. The data function weighting matrix involves RET and DET values, while the transactional function weighting matrix involves FTR and DET values [5].

Based on the explanation above, the calculation of the unadjusted function point is divided into two parts: the standard weighting of the function and the calculation of all weight values.

*3) Weighing the Standard Function*

*a) External Input (EI)*

EI is an elementary software process related to receiving data from outside the system so that changes in software behavior and changes in ILF data occur. Examples of EI are input data from end users or other systems. Table 1 shows the EI complexity matrix and the weight value for each level of complexity.

TABLE I. EI COMPLEXITY MATRIX

| Number of File Type Reference (FTR) | Number of Data Element Type (DET) | | |
|---|---|---|---|
| | 1-4 | 5-15 | > 16 |
| 0 - 1 | Low (3) | Low (3) | Average (4) |
| 2 | Low (3) | Average (4) | High (6) |
| > 2 | Average (4) | High (6) | High (6) |

*b) External Output (EO)*

EO is an elementary software process that sends data from inside to outside the system. The logical process of retrieving data from within the system contains at least one of the processes between the mathematical calculation process, the process of making derived data, and changing the data of one or more ILFs. An example of EO is creating an output file sent to another system [20]. Table 2 shows the EO complexity matrix and the weight value for each level of complexity.

TABLE II. EO COMPLEXITY MATRIX

| Number of File Type Reference (FTR) | Number of Data Element Type (DET) | | |
|---|---|---|---|
| | 1-5 | 6-19 | > 19 |
| 0 – 1 | Low (4) | Low (4) | Average (5) |
| 2 – 3 | Low (4) | Average (5) | High (7) |
| > 3 | Average (5) | High (7) | High (7) |

*c) External Inquiries (EQ)*

EQ is an elementary software process that sends data from inside to outside the system. The difference between EQ and EO lies in data collection. EO does not create data from the mathematical calculation and derived data process [20]. Table 3 shows the EQ complexity matrix and the weight value for each level of complexity.

TABLE III. EQ COMPLEXITY MATRIX

| Number of File Type Reference (FTR) | Number of Data Element Type (DET) | | |
|---|---|---|---|
| | 1-5 | 6-19 | > 19 |
| 0 – 1 | Low (3) | Low (3) | Average (4) |
| 2 – 3 | Low (3) | Average (4) | High (6) |
| > 3 | Average (4) | High (6) | High (6) |

*d) Internal Logical File (ILF)*

ILF is a logical group of corresponding data within the software scope and managed by one or more leading software processes. An example of an ILF is a table in a relational database and a collection of files stored in an application [20]. Table 4 shows the ILF complexity matrix and the weight values for each level of complexity.

TABLE IV. ILF COMPLEXITY MATRIX

| Number of File Type Reference (FTR) | Number of Data Element Type (DET) | | |
|---|---|---|---|
| | 1 - 19 | 20 - 50 | > 50 |
| 1 | Low (7) | Low (7) | Average (10) |
| 2 – 5 | Low (7) | Average (10) | High (15) |
| > 5 | Average (10) | High (15) | High (15) |

*e) External Output (EO)*

EIF is a logical group of interrelated data from outside the scope of the software and managed by one or more of the leading software processes. Logical data EIF is a source of reference data by the software being measured [20]. Table 5 shows the EIF complexity matrix and the weight values for each level of complexity.

TABLE V. EIF COMPLEXITY MATRIX

| Number of File Type Reference (FTR) | Number of Data Element Type (DET) | | |
|---|---|---|---|
| | 1 - 19 | 20 - 50 | > 50 |
| 1 | Low (5) | Low (5) | Average (5) |
| 2 – 5 | Low (5) | Average (7) | High (10) |
| > 5 | Average (7) | High (10) | High (10) |

*4) Calculating Unadjusted Function Point (UFP)*

The calculation of the UFP value is carried out by adding the weights of EI, EO, EQ, ILF, and EIF, which are calculated based on the complexity value of the software functionality [3]. Table 6 shows the process of calculating UFP.

*5) Justifying the Value Adjustment Factor (VAF)*

Value adjustment factor (VAF) is a set of factors that affect software complexity [20]. VAF uses standardized questions of general system characteristics (GSCs) to assess general characteristics of software functionality. GSCs have 14 characteristics that reflect the degree of influence of requirements on functional software. The VAF value is calculated based on the justification of the domain expert who knows the software domain by giving weight to each characteristic between the ranges of 0 (not essential) to d. 5 (very important) [3]. Table 7 shows a list of VAF questions.

Furthermore, the TDI value was used in calculating the VAF value using Equation 2.

$$VAF = 0.65 + (TDI \times 0.01) \quad (2)$$

- VAF : *value adjustment factor*
- TDI : *total degree of influence*

Information:

TABLE VI. UNADJUSTED FUNCTION POINT (UFP) CALCULATION [3]

| Standard Function | Software Complexity | | | |
|---|---|---|---|---|
| | Low (L) | Average (A) | High (H) | Total |
| External Input (EI) | __ x 3 = | __ x 4 = | __ x 6 = | Total wight of EI |
| Enternal Output (EO) | __ x 4 = | __ x 5 = | __ x 7 = | Total wight of EO |
| External Inquries (EQ) | __ x 3 = | __ x 4 = | __ x 6 = | Total wight of EQ |
| Internal Logical Files (ILF) | __ x 7 = | __ x 10 = | __ x 15 = | Total wight of ILF |
| External Interface File (EIF) | __ x 5 = | __ x 7 = | __ x 10 = | Total wight of EIF |
| Unadjusted Function Point (UFP) | | | | The sum of weight EI EO, EQ, ILF, EIF |

TABLE VII. UNADJUSTED FUNCTION POINT (UFP) CALCULATION [3]

| No. | Characteristics | Question | Degree of Influence (DI) *) |
|---|---|---|---|
| 1. | Data communications | How many communication facilities are there to aid the transfer or exchange of information with the application or system? | _____ |
| 2. | Distributed data processing | How are distributed data and processing functions handled? | _____ |
| 3. | Performance | Did the user require response time or throughput? | _____ |
| 4. | Heavily used configuration | How heavily used is the current hardware platform where the application will be executed? | _____ |
| 5. | Transaction rate | How frequently are transactions executed daily, weekly, monthly, etc.? | _____ |
| 6. | On-Line data entry | What percentage of the information is entered On-Line? | _____ |
| 7. | End-user efficiency | Was the application designed for end-user efficient? | _____ |
| 8. | On-Line update | How many ILF's are updated by On-Line transaction? | _____ |
| 9. | Complex processing | Does the application have extensive logical or mathematical processing? | _____ |
| 10. | Reusability | Was the application developed to meet one or many user's needs? | _____ |
| 11. | Installation ease | How difficult is conversion and installation? | _____ |
| 12. | Operational ease | How effective and/or automated are start-up, back up, and recovery procedures? | _____ |
| 13. | Multiple sites | Was the application specifically designed, developed, and supported to be installed at multiple sites for multiple organizations? | _____ |
| 14. | Facilitate change | Was the application specifically designed, developed, and supported to facilitate change? | _____ |
| Total Degree of Influence (TDI) | | | $\Sigma DI_{1-14}$ |

*6) Calculating Adjusted Function Point (AFP)*

Adjusted function point (AFP) is the final value of FP as the value of software complexity calculated based on the type of software [20]. AFP calculation was done by using equation 1 formula with APFC software type.

## B. Software Developer Specification

Software developers are experts who are engaged in developing software. A software developer career in software engineering consists of three levels based on programming skills: junior, middle, and senior [21]. These levels reflect the specifications of the software developer, who can show the responsibilities, qualifications, and amount of take-home pay, as well as the level of productivity in program coding.

- A junior software developer is a developer who has experience developing software for 1-3 years and is familiar with 1 or 2 programming languages/ development frameworks as well as basic programs, such as programming structures, ACID attributes (atomicity, consistency, isolation, and durability) databases, data transactions in databases, and the basis of database design [21][22].
- A middle software developer is a developer who has experience in developing software for 3-5 years, mastering 2 or 3 programming languages/ framework development, programming with reasonably high complexity, able to work as a problem solver, and able to perform proper debugging but has not been able to make appropriate technology decisions [21] [22]. At this level, the developer is suitable for software development, has enough experience working in the field of software development, and is usually quite proficient at being a full-stack developer (backend, frontend, and database).
- Senior software developers have experience in developing software for at least five years, master new programming languages, adapt quickly, and work as problem solvers by providing the best solutions [21] [22]. Senior software developers generally analyze problems that have not occurred, then take preventive measures by preparing the right technology during software development. At this

level, the developer has experience as an expert in the world of work.

The software developer specifications were further used as the research object to determine the productivity value during software implementation in SDP. Furthermore, the programming skills in software development at each level of software developer were mapped through domain expert interviews.

### C. Analytical Hierarchy Process

SDP productivity measures software developer performance by calculating the comparison ratio between software size, the product produced, and the effort spent producing the product (Adrián, 2015). Figure 1 shows an illustration of the SDP productivity model.

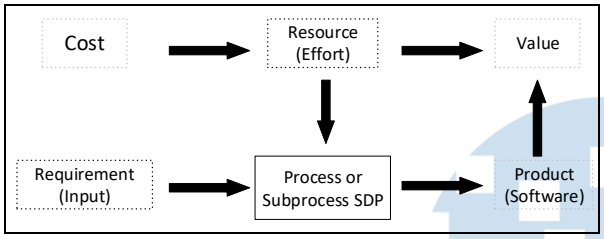


Fig. 2. Productivity Model [9]

Figure 1 is a productivity diagram that shows that there is effort as a resource needed to produce software on SDP. Thus, the measurement of SDP productivity used the Equation 3.

$$Productivity = \frac{Software\ Size}{effort} \quad (3)$$

Software size is a software size calculated based on the value of software complexity. Meanwhile, effort is the time a developer takes to produce software. However, equation 3 still used general parameters because the software size and SDP effort values can be calculated using various approaches, as shown in Tables 8 and 9.

TABLE VIII. SOFTWARE SIZE MEASUREMENT [10]

| Size | Parameter | Units of measurement |
|---|---|---|
| Task | Number of: Classes, Modifications, Modifications Request, Module Modifications, Modules, Work Items, Pages, Requirements | #Classes, #Modifications, #Modules, #WorkItems, #Pages, #Requirements |
| EP | FP, CFP, EFP, S, Code Size, OOmFPWeb, UFP, OOFP, SM | function points |
| LOC | LOC, KLOC, KSLOC, SLOC, ELOC, NLOC, AvgLOC, WSDI, SLC, KNCSS, LOC added, S, SL L, CP, Size, Total Churn, NCLOC, Code Contribution | lines of code |

TABLE IX. INPUT MEASUREMENT [10]

| Size | Parameter | Units of measurement |
|---|---|---|
| person | developer | person |
| cost | C, man-cost | person-cost |
| time | hour, T, time, minute, dav-time, time-month, month, cycle-time, year | hour, minute, month, year |
| effort | developer-hour, developer-quarter, developer-year, E, Eft, engineering-month, H, man-day, man-hour, man-month, effort, PD, PH, man-quarter, PM, man-project-time, SM, person-days, person-month, staff-hour, Staff-month | person-hour, person-day, person-month, person-quarter, person-year |

Based on the table above, the SDP productivity parameters used function points as the value of software complexity and person-day as a unit of effort. The selection of person-day as an effort parameter because the FP that is done at one time is easier to analyze in units of days. In addition, planning a software implementation schedule by a domain expert on a feature with low complexity must be done at least one day before making the program. It is necessary to understand software functionality so that there is time allocation for unexpected conditions.

Function points were calculated through FPA, while person-days were determined based on the software developer's programming ability in software development. Thus, the calculation of the SDP productivity value used Equation 4.

$$Productivity\ SDP = \frac{FP}{Person-Day} \quad (4)$$

Description:

- FP : *function point*, a measure of software complexity
- Person-Day : total working time (days) per software developer in implementing each functional software.

The SDP productivity measured was the programming ability of software developers at different levels in software implementation, from design to program code. Furthermore, the productivity value was compared to get the percentage level of speed of software implementation on the same FP software.

## III. RESULTS AND DISCUSSION

### A. Case Study

The focus of the study was FP calculations on software development from similar applications, namely the McDelivery application, which can be accessed at the link https://www.mcdelivery.co.id/id/. McDelivery is a web-based application used for ordering food and paying for restaurants. The application has simple software functionality, such as user authentication, viewing data, inserting data, updating data, deleting data, and validating transactions. In addition, services come from outside

the system, such as payment gateways during payment processing and location coordinates from the Google API. These two characteristics are the basis for choosing McDelivery as a case study describing FP as software size calculation. The scope of the McDelivery application is divided into two parts, namely the ordering process and the payment process with the software functionality shown in Table 10.

TABLE X. MCDELIVERY FUNCTIONALITY SOFTWARE

| Feature Code | Software Functionality | Specification Details |
|---|---|---|
| FR01 | Homepage | The interface displays general information about the ordering system. |
| FR02 | The page starts ordering with a page in the form of a pop-up order box | The order registration process interface consists of three ways:<br>• Method 1. Log in for customers who have registered a user.<br>• Method 2. Registration for customers who do not have an account and are visiting the website for the first time. Orders are made using a personal identity stored on the website.<br>• Method 3. Order food with guest status for customers who want to place an order without register. |
| FR03 | Registration page | The account registration process on the website. |
| FR04 | Delivery address input page | The interface for filling out the order delivery address form which includes the process:<br>a. filling in the order field.<br>b. displaying the delivery location map. |
| FR05 | Menu package list page | The interface displays a food menu catalog, and there is a process for adding menus to the order cart. |
| FR06 | Early message pop up | The order process ahead of delivery time. |
| FR07 | Food ordering page | The food ordering process includes the process:<br>a. displaying menu list.<br>b. viewing order list information.<br>c. completing the order. |
| FR08 | Order details page | The food ordering detail process interface includes the process:<br>a. viewing detailed menu information.<br>b. placing an order by inputting the number of orders on the selected menu and entering unique request data.<br>c. calculating the total price of the order.<br>d. adding order to cart. |
| FR09 | Order overview page | The interface displays an order summary which includes the process:<br>a. displaying a list of order details.<br>b. entering unique record data.<br>c. displaying a list of payment bills.<br>d. entering captcha code. |
| FR10 | Payment page | The interface displays the payment type for processing, including<br>a. Entering the payment method by selecting the available payment types.<br>b. Entering the delivery contact.<br>c. showing a list of bills including order code.<br>d. confirmation of order data. |
| FR11 | Payment processing page | The order bill payment process, including the process:<br>a. displaying the billing list and buyer contact.<br>b. displaying a choice of payment methods (debit or gopay).<br>c. processing the payment using the debit or gopay method according to the selected payment method. |
| FR12 | Order confirmation page | The successful order confirmation to the customer includes the process:<br>a. displaying information on successful payment and display shipping address.<br>b. customer can track order.<br>c. customer can add order to favorite list.<br>d. sending an email to provide information about successfully placing an order and displaying the details of the order list. |
| FR13 | Order tracking page | The interface displays the order status after the order is received by McDelivery. There are four order statuses, namely orders received, in process, being delivered, and sent. |
| FR14 | Order page failed | The interface displays an order failed message when the booking time exceeds the order limit of 30 minutes. The failed booking page contains the following order failed information dan cancel order button. The system displays the close application page when the customer presses the cancel order button. |
| FR15 | Great offers page | The interface displays a list of promos offered by McDelivery at the time of food/beverage purchases. |
| FR16 | Website headers | Navigation links are located in the header section at the top of each website page. |
| FR17 | Website Footer | Navigation links are located in the footer at the bottom of each website page. |
| FR18 | Order sidebar | The navigation link on the side is for displaying the menu list category of the food/beverage ordering page. |
| FR19 | Terms and conditions page | The interface displays a list of terms and conditions for placing an order from the menu offered by McDelivery. |
| FR20 | Privacy policy page | The interface displays policies and privacy as long as the customer makes a menu order on the system. |
| FR21 | Question and answer page | The interface displays a list of questions and answers about how to order menus on the McDelivery information system. |

*B. Software Size Calculation*

Software size was calculated using the FPA method, which consisted of 3 stages. FP calculations on a McDelivery system are described below.

*1) Unadjusted Function Point (UFP) Calculation*

UFP was used to see software complexity by weighing the software functionality against five standard functions based on process logic design. Table 12 shows the software's weighting results, while Table 13 shows the results of the calculation of UFP.

*2) Value Adjustment Factor (VAF) Justification*

VAF justification was carried out by a senior developer with 12 years of experience developing web-based software who knows the logic's complexity in implementing software functionality into a program code. Table 11 shows the results of VAF justification for GSCs by domain experts.

TABLE XI. NORMALIZED DECISION MATRIX

| Serial Number | Criteria | | | | |
|---|---|---|---|---|---|
| | C1 | C2 | C3 | C4 | C5 |
| A1 | 0,1672 | 0,1689 | 0,4628 | 0,4036 | 0,3404 |
| A2 | 0,2326 | 0,1639 | 0,4474 | 0,5340 | 0,3341 |
| A3 | 0,3126 | 0,4273 | 0,1645 | 0,2607 | 0,3530 |
| A4 | 0,0872 | 0,2832 | 0,1645 | 0,2607 | 0,4224 |
| A5 | 0,3126 | 0,3329 | 0,3291 | 0,2794 | 0,0756 |
| A6 | 0,1672 | 0,4273 | 0,2725 | 0,2607 | 0,2017 |
| A7 | 0,3126 | 0,4819 | 0,2314 | 0,1490 | 0,2900 |
| A8 | 0,3853 | 0,2782 | 0,1645 | 0,2670 | 0,2837 |
| A9 | 0,3199 | 0,2236 | 0,1182 | 0,2607 | 0,2017 |
| A10 | 0,5816 | 0,1689 | 0,5040 | 0,3290 | 0,4665 |

Furthermore, the calculation of VAF used the equation 2.

$$VAF = 0.65 + (TDI \times 0.01)$$

$$= 0.65 + (42 \times 0.01) = 1.07$$

*3) Adjusted Function Point (AFP) Calculation*

The FP value used the equation on the type of APFC software development using the 1 equation.

$$AFP = ADD^{*} \times VAF$$

$$AFP = 496 \times 107 = 530{,}72$$

*) The ADD value was taken from the results of the UFP value in Table 12

*C. Software Size Calculation*

Determination of software developer specifications on programming skills used interview techniques to domain software development experts who have positions as senior software developers. The interview technique aimed to justify the effort as SDP productivity input parameters have a domain scope consistent with developer specifications at each level.

The domain expert has experience developing software on various software functionalities, such as software development on multi-platforms (web, mobile, and desktop), data retrieval into excel and pdf files, coding with scheduled running (schedulers), application programming interfaces (API), and creation of user interfaces in web and mobile form. In addition, the domain expert can map human resources (HR) into software development projects based on the level of software complexity and software developer programming skills. Table 14 shows the results of the analysis of software developer specifications at each level of programming ability.

TABLE XII. WEIGHTING THE FUNCTIONALITY OF MCDELIVERY SOFTWARE

| Software Functionality | EI | | | EO | | | EQ | | | ILF | | | EIF | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | FTR | DET | W | FTR | DET | W | FTR | DET | W | FTR | DET | W | FTR | DET | W |
| FR01 | 3 | 49 | H | 0 | 0 | L | 3 | 3 | L | 6 | 16 | A | 0 | 0 | L |
| FR02 | 1 | 12 | L | 0 | 0 | L | 1 | 2 | L | 1 | 2 | L | 0 | 0 | L |
| FR03 | 1 | 31 | A | 1 | 4 | L | 1 | 28 | A | 4 | 13 | L | 0 | 0 | L |
| FR04 | 2 | 13 | A | 0 | 0 | L | 2 | 10 | A | 1 | 8 | L | 1 | 6 | L |
| FR05 | 1 | 15 | L | 0 | 2 | L | 1 | 8 | L | 3 | 7 | L | 0 | 0 | L |
| FR06 | 1 | 7 | L | 2 | 5 | L | 1 | 9 | L | 2 | 6 | L | 0 | 0 | L |
| FR07 | 1 | 17 | A | 1 | 5 | L | 2 | 35 | H | 2 | 21 | A | 0 | 0 | L |
| FR08 | 1 | 14 | L | 2 | 6 | A | 1 | 32 | A | 3 | 15 | L | 0 | 0 | L |
| FR09 | 2 | 14 | A | 2 | 5 | L | 2 | 16 | A | 1 | 19 | L | 0 | 0 | L |
| FR10 | 2 | 14 | A | 2 | 2 | L | 2 | 13 | A | 2 | 21 | A | 0 | 0 | L |
| FR11 | 2 | 27 | H | 0 | 0 | L | 2 | 57 | H | 4 | 15 | L | 1 | 1 | L |
| FR12 | 1 | 5 | L | 0 | 0 | L | 1 | 22 | A | 4 | 14 | L | 0 | 0 | L |
| FR13 | 1 | 2 | L | 0 | 0 | L | 1 | 8 | L | 1 | 6 | L | 0 | 0 | L |
| FR14 | 0 | 2 | L | 0 | 0 | L | 0 | 4 | L | 0 | 0 | L | 0 | 0 | L |
| FR15 | 1 | 4 | L | 0 | 0 | L | 1 | 11 | L | 1 | 6 | L | 0 | 0 | L |
| FR16 | 1 | 9 | L | 0 | 0 | L | 1 | 7 | L | 2 | 4 | L | 0 | 0 | L |
| FR17 | 1 | 17 | A | 0 | 0 | L | 1 | 8 | L | 3 | 7 | L | 0 | 0 | L |
| FR18 | 1 | 3 | L | 0 | 0 | L | 0 | 11 | L | 2 | 7 | L | 0 | 0 | L |
| FR19 | 1 | 4 | L | 0 | 0 | L | 0 | 4 | L | 2 | 5 | L | 0 | 0 | L |
| FR20 | 1 | 4 | L | 0 | 0 | L | 1 | 4 | L | 2 | 5 | L | 0 | 0 | L |
| FR21 | 1 | 7 | L | 0 | 0 | L | 1 | 10 | L | 3 | 6 | L | 0 | 0 | L |

*) W = Weight of Level Complexity Software

TABLE XIII. UFP MCDELIVERY CALCULATION RESULTS

| Standard Function | Software Complexity Value | | | |
|---|---|---|---|---|
| | Low (L) | Average (A) | High (H) | Total |
| External Input (EI) | 13 x 3 = 39 | 6 x 4 = 24 | 2 x 6 = 12 | 75 |
| Enternal Output (EO) | 20 x 4 = 80 | 1 x 5 = 5 | 0 x 7 = 0 | 85 |
| External Inquries (EQ) | 13 x 3 = 39 | 6 x 4 = 24 | 2 x 6 =12 | 75 |
| Internal Logical Files (ILF) | 18 x 7 = 126 | 3 x 10 = 30 | 0 x 15 = 0 | 156 |
| External Interface File (EIF) | 21 x 5 = 105 | 0 x 7 = 0 | 0 x 10 = 0 | 105 |
| Unadjusted Function Point (UFP) | | | | 496 |

TABLE XIV. SOFTWARE DEVELOPER PROGRAMMING SPECIFICATIONS

| Level | Category | Programming Ability |
|---|---|---|
| Junior | Algorithm understanding | • basic validation logic, namely mandatory, field format, data type, and alignment.<br>• regular operating business.<br>• basic algorithm structure, namely sequence, selection, and repetition.<br>• complex algorithm structure, namely nested if with two levels, nested repetition with 2-3 levels, and combination of if and repetition with two levels. |
| | Coding | • reading the source code process flow.<br>• creating program code according to software functionality.<br>• proprietary bug fixing program code.<br>• implementation of functions according to the development framework.<br>• understand aspects of clean code. |
| | Query database | managing database with data definition language and data manipulation language, and retrieve data with the complexity of two tables. |
| | Technology exploration | installing and adding plugin tools. |
| | Software testing | self-testing, unit testing, and code quality checker. |
| Middle | Algorithm understanding | middle developers have all the algorithm understanding abilities of junior-level software developers, business validation, and the use of algorithm structures with a complexity of 3 to 5 levels. |
| | Coding | middle developers have all the coding skills of a junior-level software developer, bug fixing in other people's code, and can create effective code. |
| | Query database | middle developers have all the skills to make database queries owned by junior-level software developers, manage databases with additional data control languages, master PL/SQL and retrieve data with a complexity of 3 to 5 tables. |
| | Technology exploration | installing, adding plugin tools, and modifying tools. |
| | Software testing | self-testing, unit testing, and code quality checker. |
| Senior | Algorithm understanding | senior developers have all the algorithm understanding abilities of medium-level software developers and use algorithm structures with six levels of complexity to infinity. |
| | Coding | senior developers have all the coding skills of medium-level software developers, review the creation of effective program code structures following the software development framework, and can create development frameworks. |
| | Query database | senior developers have all the capabilities to make database queries owned by medium software developers and retrieve data with six levels of complexity to infinity. |
| | Technology exploration | senior developers have all the capabilities of a medium-level software developer and decide on the right technology according to the problem domain. |
| | Software testing | self-testing, unit testing, and code quality checker. |

### *D. SDP Productivity Calculation*

The effort to calculate the SDP productivity value is the time required for software developers to complete program coding of software functionality in person-day units. It was assumed that work effort for one day is 8 hours. The determination of effort is known through justification by domain experts based on the software developer's level of software complexity and programming ability.

Based on the detailed specifications of the software functionality in Table 10, the SDP productivity value for each software developer is shown in Table 15. The effort value of each software developer can be determined from the programming ability and work attitude so that each developer at the same level of programming ability has a different effort value. However, in this study, the effort justification process only involved programming skills, while the ability-to-work attitude was considered the same.

TABLE XV. SOFTWARE DEVELOPMENT EFFORT

| Feature Code | Software Developer Effort (person-day) | | |
|---|---|---|---|
| | Junior | Middle | Senior |
| FR01 | 3 | 2 | 2 |
| FR02 | 5 | 3 | 3 |
| FR03 | 4 | 3 | 3 |
| FR04 | 4 | 3 | 2 |
| FR05 | 2 | 1 | 1 |
| FR06 | 1 | 1 | 1 |
| FR07 | 3 | 2 | 2 |
| FR08 | 5 | 4 | 3 |
| FR09 | 3 | 2 | 2 |
| FR10 | 2 | 1 | 1 |
| FR11 | 5 | 3 | 3 |
| FR12 | 2 | 2 | 2 |
| FR13 | 2 | 1 | 1 |
| FR14 | 3 | 2 | 2 |
| FR15 | 3 | 2 | 2 |
| FR16 | 3 | 2 | 2 |
| FR17 | 1 | 1 | 1 |
| FR18 | 2 | 1 | 1 |
| FR19 | 1 | 1 | 1 |
| FR20 | 1 | 1 | 1 |
| FR21 | 1 | 1 | 1 |
| Total Effort | 56 | 40 | 37 |

Furthermore, SDP productivity was calculated using equation 4 in each software developer specification. The software size parameter uses FP, worth 530.72, while the effort value is based on the total effort per person-day. The value of software developer productivity on SDP is the number of function points that can be worked on for one day to create programs based on requirements specifications and software design. Table 16 shows the results of the calculation of productivity.

TABLE XVI. EFFORT SOFTWARE DEVELOPMENT IN CODING STAGE

| Software Developer Level | Effort (person-day) | Productivity (FP/person-day) |
|---|---|---|
| Junior | 56 | 9.87 |
| Middle | 40 | 13.27 |
| Senior | 37 | 14.34 |

Table 16 illustrates that junior software developers can implement 9 FPs in one day, middle software developers implement 13 FPs in one day, and senior software developers implement 14 FPs per day in the case of the McDelivery ordering system. Based on the effort given in each software developer specification in the case of APFC software development, it can be concluded that the production value is directly proportional to the level of programming ability. This is in line with the visualization of Figure 3.

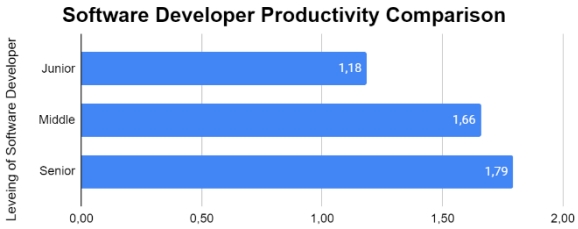


Fig. 3. Software Developer Productivity Comparison

Figure 3 compares software developer productivity at each level in developing software—the developers' ability at every level goes with their experience in software development. The productivity of software developers depends on the software's complexity and the software developer's programming skills. So Figure 3 relates to Table 14 regarding the specification of the software developer's programming abilities for four abilities: algorithms, coding, database queries, technology exploration, and software testing. Junior developers have fewer abilities compared to the two levels of developers above. The middle developer status is between junior and senior, while the senior developer has the highest ability.

The higher the understanding of programming skills when implementing software, the higher the level of productivity. For example, in Table 16 or Figure 3, senior software developers have higher productivity scores and programming skills as problem solvers. Thus, calculating software developer productivity in making program code on SDP using equation 4 can describe productivity quantitatively.

The productivity value at each level of a software developer can be used as an alternative solution to calculate the estimated time, cost, and availability of human resources in planning software development for similar applications that already exist. The estimation is calculated based on the software size, the required software developer specifications, and the productivity value at each developer level who can implement several FPs in one day in the form of program code. Each software developer has a different productivity value based on the level of programming ability and software complexity.

This productivity value is very likely to be used for all software development projects, mainly if the project has limited resources (time, cost, and human resources). For example, suppose a software development project has a time limit that must be completed immediately. In that case, the selection of a software developer is based on the least effort by prioritizing the productivity of highly qualified human resources. Meanwhile, suppose a software development project has limited human resources with senior software developer conditions already mapped out on other software development

projects. In that case, the alternative is to choose a software developer with slightly lower productivity—for example, the selection of HR with medium or low qualifications.

However, the estimated cost of software development calculation needs to be studied deeper, to know whether the increase in productivity is inversely proportional to the cost. For example, the higher the value of HR productivity, the lower the development costs or vice versa. On the other hand, increased productivity may be directly proportional to development costs; the higher the value of HR productivity, the more expensive development costs. Development costs are increasing because of the need for highly qualified human resources with higher salaries.

## IV. CONCLUSIONS

Information on software developer productivity at every level of programming ability is the primary key to making a more realistic SDP plan, namely determining the duration of software development based on the number of FP/day each software developer can develop. For example, senior software developers have less development time than junior and middle developers. This statement is evidenced by the productivity value of senior developers being higher than that of middle and junior developers, namely 9.87 for junior developers, 13.27 for middle developers, and 14.34 for senior developers.

The estimated productivity value is calculated based on the level of programming ability at the software implementation stage to the complexity of the software functionality. The use of function points in calculating the productivity of specific projects allows it to be used as a comparison with other projects in similar problem domains. The number of FP/day implemented by software developers utilizes previous productivity data so that the productivity measure of software developers can be used as an estimated parameter for software development planning and a better estimate of the budget for new projects. This is because the calculation of development effort in planning, which is initially based on the subjective assessment of the SDP planner, can be replaced with an objective assessment by utilizing the productivity value of each software developer's programming ability calculated quantitatively.

Suggestions for further research include adding parameters to calculate software developer effort. For example, the specification of programming skills and work attitude skills, so it is necessary to design a case study on a software development project. In addition, the scope of productivity calculations is not only at the software implementation stage. However, it can involve other development stages, such as functionality specification analysis, design, testing, or software maintenance so that function point calculations can use the DPFPC and EPFC software types.

## ACKNOWLEDGMENT

Researchers would like to thank the Center for Research and Community Service (PPPM) Bandung State Polytechnic (POLBAN), which has supported the implementation of this research, for the financial assistance that has been provided.

## REFERENCES

[1] N. Rachmat and Saparudin, "Estimasi Ukuran Perangkat Lunak Menggunakan Function Point Analysis-Studi Kasus Aplikasi Pengujian dan Pembelajaran Berbasis Web," in *Prosiding Annual Research Seminar*, 2017, pp. 3–5.

[2] H. Rohayani, F. L. Gaol, B. Soewito, and H. L. Hendrie, "Estimated Measurement Quality Software On Structural Model Academic System With Function Point Analysis," in *International Conference on Applied Computer and Communication Technologies (ComCom)*, May 2017.

[3] A. Y. P. Putri and A. P. Subriadi, "Software Cost Estimation Using Function Point Analysis," in *The 4th International Seminar on Science and Technology*, Aug. 2018, vol. 79, pp. 79–83.

[4] M. Usman, E. Mendes, and J. Börstler, "Effort estimation in Agile software development: A survey on the state of the practice," in *ACM International Conference Proceeding Series*, Apr. 2015, vol. 27-29-April-2015. doi: 10.1145/2745802.2745813.

[5] M. F. Hillman and A. P. Subriadi, "40 Years Journey of Function Point Analysis Against Real-time and Multimedia Applications," in *The fifth Information System International Conference* , 2019, pp. 266–274.

[6] B. Prakash and V. Viswanathan, "A survey on software estimation techniques in traditional and agile development models," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 7, no. 3, pp. 867–876, Sep. 2017, doi: 10.11591/ijeecs.v7.i3.pp867-876.

[7] S. Wagner, "Defining Productivity in Software Engineering," in *Rethinking Productivity in Software Engineering*, First Edition., C. Sadowski and T. Zimmermann, Eds. Apress Oppen, 2019. doi: 10.1007/978-1-4842-4221-6.

[8] A. H. López, R. C. Palacios, P. S. Acosta, and C. C. Lumberas, "Productivity measurement in software engineering: A study of the inputs and the outputs," *International Journal of Information Technologies and Systems Approach*, vol. 8, no. 1, pp. 45–67, Jan. 2015, doi: 10.4018/IJITSA.2015010103.

[9] E. A. de Oliveira and R. C. Noya, "Using Productivity Measure and Function Points to Improve the Software Development Process," *Computer Science*, 2013.

[10] E. Oliveira, D. Viana, M. Cristo, and T. Conte, "How have software engineering researchers been measuring software productivity?: A systematic mapping study," in *ICEIS 2017 - Proceedings of the 19th International Conference on Enterprise Information Systems*, 2017, vol. 2, pp. 76–87. doi: 10.5220/0006314400760087.

[11] A. H. López, R. C. Palacios, Á. G. Crespo, and F. C. Isla, "Software Engineering Productivity: Concepts, Issues and Challenges," *International Journal Information Technology Project Managemet*, vol. 2, pp. 37–41, 2011.

[12] S. I. Mohamed, "Software development productivity impact from an industrial perspective," *Int J Sci Eng Res*, vol. 6, no. 2, pp. 1333–1342, Feb. 2015, [Online]. Available: http://www.ijser.org.

[13] J. Rashid, T. Mahmood, and W. M. Nisar, "A Study on Software Metrics and its Impact on Software Quality,"

*Technical Journal, University of Engineering and Technology (UET)*, vol. 24, no. 1, pp. 1–14, 2019.

[14] N. Choursiya and R. Yadav, "An Enhanced Function Point Analysis (FPA) Method for Software Size Estimation," *International Journal of Computer Science and Information Technologies IJCSIT*, vol. 6, no. 3, pp. 2797–2799, 2015.

[15] S. M. R. Chirra and H. Reza, "A Survey on Software Cost Estimation Techniques," *Journal of Software Engineering and Applications*, vol. 12, no. 06, pp. 226–248, 2019, doi: 10.4236/jsea.2019.126014.

[16] P. Vickers, "An Introduction to Function Point Analysis," 2003. [Online]. Available: www.paulvickers.com/northumbria.

[17] D. Garmus and D. Herron, *Function Point Analysis: Measurement Practices for Successful Software Project*. Addison Willey Professional, 2000.

[18] J. Shah, N. Kama, and S. A. Ismail, "An empirical study with function point analysis for software development phase method," in *ACM International Conference Proceeding Series*, May 2018, pp. 7–11. doi: 10.1145/3220267.3220268.

[19] A. Alexander, "How to Determine Your Application Size Using Function Points," 2004. [Online]. Available: http://ifpug.org.

[20] IFPUG, "Function Point Counting Practices Manual," 2010.

[21] AltexSoft Team, "Software Engineer Qualification Levels: Junior, Middle, and Senior," *Altexsoft*, Sep. 23, 2018. https://www.altexsoft.com/blog/business/software-engineer-qualification-levels-junior-middle-and-senior/ (accessed Sep. 27, 2022).

[22] K. Anderson, "Junior vs. Mid vs. Senior software engineers – experience, skills, & expectations," *DEPT*. https://www.deptagency.com/en-us/insight/junior-vs-mid-vs-senior-software-engineers-experience-skills-expectations/ (accessed Sep. 27, 2022).

ISSN 2085-4552

# Comparative Analysis of Phishing Tools on Social Media Sites

Putu Candra Ariani[1], I Gede Bagastia Widi Atmaja[2], Kadek Sukma Jayanti[3], I Gusti Agung Ayu Ananda Dewi[4], Gede Arna Jude Saskara[5], I Made Edy Listartha[6]

[1,2,3,4,5,6]Faculty of Engineering and Vocational, Ganesha University of Education, Singaraja, Indonesia
[1]candra.ariani@undiksha.ac.id, [2]bagastia.widi@undiksha.ac.id, [3]sukma.jayanti@undiksha.ac.id, [4]agung.ayu.ananda@undiksha.ac.id, [5]jude.saskara@undiksha.ac.id , [6]listartha@gmail.com

Accepted 09 December 2022
Approved 29 December 2022

***Abstract*— Social networks, often referred to as social media are a form of information technology development. Social networks are used by society in obtaining information, being a means of long-distance communication, as well as distributing information. However, behind the development of social media which has experienced significant developments, there are problems related to information security. Vulnerable to leakage of credential data and fraud becomes a negative impact due to the development of social media, one of the scams that often occurs on social media is phishing. Such scams are attempts to obtain, steal, or dig into someone's data through emails, text messages, and social media posts. Social media accounts are targeted by phishing criminals because they consist of sensitive data of social media users, one of the social media accounts that can be affected by phishing attacks is Facebook. The problem studied in this study is the way criminals use three different types of phishing tools to carry out attacks. This study aims to compare three phishing tools used in committing information crimes on social media sites in terms of features, accuracy, and ease of installation each of these phishing tools. With this research, readers can understand the comparison of these phishing tools used by hackers to access social media accounts.**

***Index Terms*— accounts; information security; phishing; phishing tools; social media**.

## I. INTRODUCTION

Social media is an important part of people's daily life. In general, information needs to encourage people to use social media as a means of communication, obtaining information, and entertainment. The most common examples of social media are Instagram, TikTok, Facebook, Twitter, and YouTube [1]. Social media is a medium that allows a person to socialize by sharing content, news, and photos with others or can be a tool to promote as well as work [2]. Web dictionaries define social media (nouns) simply as, "Websites and applications used for social networks". It is noted that a total of 160 million people in Indonesia use social media descriptions with a percentage of smartphone use at 62%, computers at 16%, and tabs at 6% [3]. Facebook is one of the social media platforms where people can share statuses, moments, and photos and sell something there, this platform stores a lot of users' personal information and can be a risk if security does not work. One interesting aspect of Facebook is the use of apps and third-party interactions. This means that each Facebook page now acts as a web page, blog, instant messaging, email system, and third-party applications that enable real-time functionality [4]. On the Facebook site, there is a statistic on the number of Facebook users which states that active Facebook users have reached more than 300 million users. Even in Indonesia, the number of Facebook users is increasing every year and the percentage of Facebook users is 2997.3% in just 1 year [5].

The rapid development of social media and providing benefits to society is one of the positive impacts of technological advances, but the use of social media is also detrimental to society. Information security threats are negative impacts arising from the development of social media, various cases of information security threats often occur on social media accounts [6]. This can happen due to the negligence of social media users, and service providers, or the deliberate negligence of information criminals. Therefore, people who use social media must have been exposed to the risks of social media itself [7]. Crime attacks on social media have many types of attacks, one of which is phishing, these attacks are scams against social media users because the perpetrator will trick the victim into obtaining and digging someone's data through emails, or uploads on social media [8]. Phishing is in many ways an evolutionary threat, phishing can become bigger and worst if hackers get personal information [9].

Hackers usually do phishing to steal the victim's personal information, or simply try their expertise in Web Security (white hackers), phishing can be done on a Facebook account by duplicating a login site, and using the link to make the victim believe and logging in via a link. The existence of a phishing tool can make it easier for perpetrators to phish the victim's account because this phishing tool can be downloaded for free so it allows the perpetrator to carry out fraud attacks freely without any obstacles.

Thus, to examine the problems that have been previously presented, the phishing tools in this study consist of three types, namely Social Engineering Toolkit (Setoolkit), SocialFish, and HiddenEye. These three tools have uses to carry out various kinds of attacks, but the focus of this research is on the use of these tools in carrying out phishing attacks on target accounts. Testing on all three phishing tools will be carried out on the Kali Linux operating system and the social media account to be tested is a Facebook account. With this test, the public knows the performance of phishing tools in manipulating social media accounts to obtain the personal data of the account owner concerned. The tools will be tested and then analyzed for a comparison of the performance contained in each tool.

## II. METHODOLOGY

### A. *Research Methods*

In this test, the researcher uses a qualitative comparative method, which means that the researcher will compare the results of this test. Qualitative research includes research that explains research using analysis and is descriptive, while comparative is comparative research.

Qualitative methods are generally defined as "a collection of methods and rules followed in science or discipline"[10]. Qualitative methods have been widely used in research in Indonesian journals. Meanwhile, comparative research is research that focuses on the research subject group, then continues by focusing on the study of variables in the comparative group [11]. Comparative research is ex post facto, which means data is collected after all the events to be studied have occurred.

### B. *Research Design*

In these tests, the results are created based on each tool's data and test objectives. The tools used by the Tester are the Social Engineering Toolkit (Setoolkit), SocialFish, and HiddenEye which have been tested on Kali Linux OS. The purpose of this test is to compare and analyze the differences between each tool. Testers install each tool first on a computer with Kali Linux OS and perform phishing tests on each tool. The comparison of each tool is accurate, the features, and the effectiveness of each tool, and get a valid conclusion.


Fig. 1. Testing the Design Flowchart

## III. RESULTS AND DISCUSSION

### A. *Setoolkit (Social Engineering Toolkit)*

Social Engineering Toolkit or called Setoolkit is a tool that uses phishing methods and serves as manipulation of login pages on a site, the main purpose of this tool is to get information about the target account, especially in the password section [12]. Setoolkit can also be mentioned as an integrated set of tools specifically designed to carry out follow-up attacks on the human element [13]. These tools are used to find information on emails or find social media account passwords. By using the phishing method, the hacker must make the target click on a fake link that has been created by the hacker, usually, the hacker will approach the target by sending a link along with an attractive image that aims to make the target fall into a trap. The way Setoolkit works starts with hackers sending fake links through social media like email and so on. After the target clicks on the link that the hacker has sent, the target will be directed to a fake site, and on the site, there is a login form with the same appearance accompanied by convincing words so that the target will not realize that the site is just a clone. All information in the form of a username or Email and password that the target enters into the form such as login will be automatically saved to the hacker's PC. Therefore, the hacker will know the password and username of the target account so that the hacker can immediately change the password of the target account and get the hang of it.

#### 1) *Testing*

- In the first step, open the Kali Linux command terminal.
- Then write the command 'sudo su' and the password that was pre-created in Kali Linux in the terminal.
- Next, open the Social Engineering Toolkit that has been installed with the 'setoolkit' command.
- After that, choose option number 1 Social Engineering Attack.

- Then, select option number 2 Website Attack Vector.
- Select option number 3 of the Credential Harvester Attack Method.
- Select option number 2 in the Site Cloner section.
- Then create a new terminal file to get the IP address, with the same command as the previous terminal, write 'sudo su' and enter the password.
- Write the command 'ifconfig' to get the IP address then copy the selection.
- Paste the IP address selection on the first terminal.
- Enter the Facebook URL www.facebook.com.
- Copy the IP address and paste it into a browser on Kali Linux OS, for example, Firefox.
- The Facebook login menu will appear, it is a manipulated version of the phishing IP address.
- Targets enter their email and password, for example, using an email sandi1@gmail.com and password: password.
- After entering the email and password then click the login button and the tools will read the previous input data and appear in the terminal as a result of phishing (email and password).

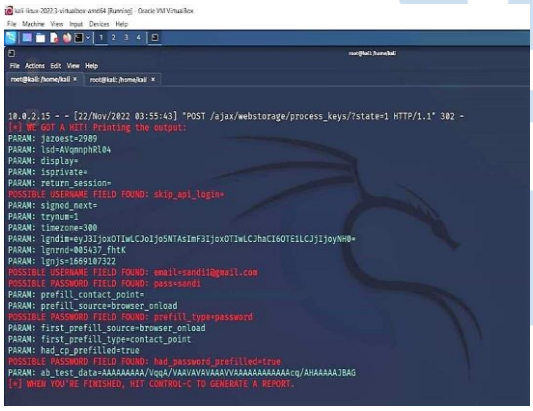


Fig. 2. Testing on Setoolkit


Fig. 3. Setoolkit Phishing Results

The testing steps are said to be successful by gathering real evidence. A little review in the seventh step because the hacker will phish the Facebook account so that the option in no. 2, namely the cloner site, functions so that it can enter the official website of the Facebook account, namely www.facebook.com.

*2) Result*

Setoolkit can phish a Facebook account at any Linux time using a pre-customized Ip address according to the one in the terminal. But in experiments, there are errors caused by signal interference. However, if you want to phish in Linux times with a device such as a different laptop, the network or localhost must be the same as this can be the cause of the error.

## *B. SocialFish*

In the world of social engineering, SocialFish is a tool that many hackers use to carry out phishing, especially email fraud. SocialFish is one of the Kali Linux tools that offer to phish on social media such as Facebook, Twitter, and Instagram as a feature [14]. With SocialFish it is necessary to get the target to click on the link to cheat the email. SocialFhis is one of the third-party tools outside of Linux, in contrast to the Social Engineering Toolkit or Setoolkit. SocialFhis must be installed first. In addition, the third version of the Python programming language needs to be installed for the tool to be used smoothly. In running SocialFish it is necessary to create a net to capture the target to be targeted. It is intended for users of the tool to create URLs http://0.0.0.0:5000 the address of this URL is used to trap the intended target. Most tools in general have provided the URL address of phishing attacks, for example, Facebook, Instagram, Twitter, and others. SocialFish is more flexible, so in using SocialFish users only copy links that will be the target of phishing, then the links will be duplicated into fake links which later the duplicate links previously had a http://0.0.0.0:5000 URL address. So, internet users who open fake links enter their login data so that it will appear in the SocialFish tool.

*1) Testing*

Testing the security of www.facebook.com pages from phishing attacks using one of the tools, called SocialFish v3, this tool is a third-party tool from the Linux operating system times. Here's how to test the security of www.facebook.com pages using the SocialFish v3 tool.

- SocialFish is a third-party phishing tool so it is necessary to download then install it to the directory, using the git clone command https://github.com/UndeadSec/SocialFish then wait until the installation process is complete.
- Run the pip install -r requirements.txt command to install the SocialFish requirements.
- Create a username and password that will be used to log in to the SocialFish web tool. With

the python3 command SocialFish.py username password.

- Next, go to the http://0.0.0.0:5000/neptune link then enter the username and password that have been created earlier.
- Then in the cloning and redirection column fill in the Facebook social media web link. Then refresh the http://0.0.0.0:5000 link then the link will become a fake link from Facebook.
- The features provided by the SocialFish tool don't have many advantages. The use of these tools is done completely manually without providing significant features. Users of this tool are required to enter the Facebook website address into the cloning and redirection field which will then be duplicated into a fake site that resembles a real URL.
- Next, the test is carried out ten times to obtain an accurate result. The first test was carried out to find out the accuracy of the "space" key on the keyboard because it was known that the "space" key can be entered into the password. A second test is performed to find out the standard rule on email, whether users can enter email addresses randomly. A third test was conducted to find out the standard rules on passwords because Facebook passwords are known to have at least eight characters. The fourth test is carried out in violation of the standard rules for both (email and password). The fifth test was carried out without entering an email address and password at all. And the sixth to tenth tests repeat the previous test from the first test to the fifth test to get the correct data accurate.

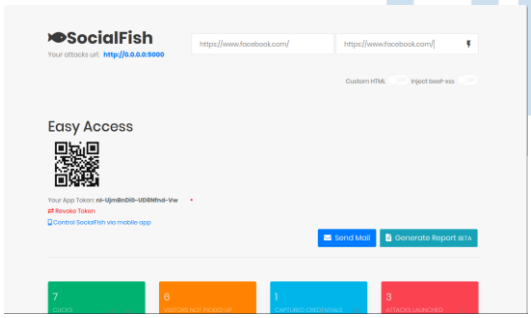


Fig. 4. Testing on SocialFish

*2) Result*

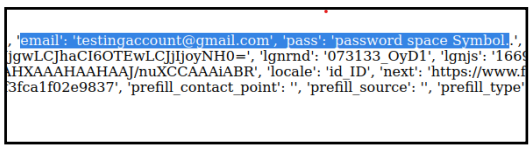


Fig. 5. SocialFish Phishing Results

Setoolkit can phish a Facebook account at any Based on the results of ten tests, the first accuracy test found that the username testingaccount@gmail.com and with a password named Symb0l password space ... not fully stated to be accurate. This is because the password cannot record "space" on the keyboard, has only one space spacing, and cannot be more than one space. In the second test, the result was that the standard and non-standard formats in E-mail and passwords could be recorded well. Then the third test had the result that passwords that were less than eight characters could be recorded in this tool. Please note that the results obtained are not only recorded results, but this tool does not have some kind of feature to recognize the standard format in E-mail and password of Facebook users. Although the results of some tests say that the results are indeed all well-recorded. However, there is one problem that is said to be less accurate, that is, the "space" button is not recorded properly as is usually done by users, that is, it can create a password with just the "space" button.

*C. HiddenEye*

HiddenEye is a highly effective Social Engineering tool that can be used to collect user credentials, and miscellaneous information [15]. This tool can be a great asset in an Enterprise-level Penetration Test or any other engagement. HiddenEye is included as a third-party tool that provides the most menu features in carrying out phishing attacks. Unlike the previous tools, HiddenEye has provided a phishing menu feature so that users only use these tools. These phishing tools are used by testers to phish the target's Facebook account, but there are other features in it such as Instagram or Twitter phishing. This feature is only for certain reasons and should not be used for illegal activities. HiddenEye does not exist by default in Kali Linux, so to use this tool the user needs to download it twice, namely on Kali Linux devices and operating systems. To use this tool, users only need to select the phishing menu provided by HiddenEye. For example, a tool user wants to use the menu feature on the Facebook page site, and in using the Facebook phishing feature, a URL will be given which is used as a medium in launching phishing attacks on the target to be targeted. Just like phishing tools in general, to use the HiddenEye tool the user simply copies or clones the original site. For example, on https://www.facebook.com site copy the URL https//127.0.0.1:1028. If the target enters a copied or fake URL, login data such as the Email and Password of the target will be sent to the HiddenEye tools so that the login data can be seen by the user of the HiddenEye tool.

*1) Testing*

The security testing on the www.facebok.com website using the HiddenEye tool is:

- The first thing that needs to be done is to download the HiddenEye tools using the git clone https://github.com/Morsmalleo/HiddenEye command and then wait for the installation process to complete.

- To enter the directory by performing the cd HiddenEye command. Next, download the Requirements file with the pip3 install -r requirements.txt command and wait for the installation process to complete.
- Run the tool from HiddenEye by using the python3 HiddenEye.py -h command.
- HiddenEye is designed to perform phishing attacks completely with a variety of features. There are several menu features options for carrying out phishing attacks, such as Facebook, Instagram, Netflix, Google, and others. And this test only focuses on the Facebook website.
- Inside the Facebook website attack menu feature. There are sub-features such as standard phishing methods, phishing sub-features using polling methods, phishing sub-features with methods of providing fake security information messages to users, and phishing sub-features through fake trust messages. But basically, these tools use the same methods to capture user information, but the feature aims to create different attack methods so that phishing attacks are much more effective and have minimal failures.
- Then, this test uses a local server with the URL address already provided by HiddenEye with the URL address http//127.0.0.1:1028/. URLs are used as a medium to launch phishing attacks or in other words URLs that match the original site so that the URLs look similar.
- Furthermore, at this stage using ten tests can produce maximum tool accuracy data. The first test was carried out to find out the accuracy of the "SPACE" key on the keyboard because it became known that the "SPACE" key can be entered in the password. The second test is performed to find out the standard rules in email and whether it is possible to select any Email for the user to enter. A third test was conducted to find out the standard password rules because it was known that the Facebook password has at least eight characters. The Fourth Test is performed by violating the standard rules of both (Email and Password). The fifth test was done by not entering the Email and Password at all. And the sixth to tenth tests repeat the previous test from test one to the fifth test to get the correct accurate data.

*2) Result*

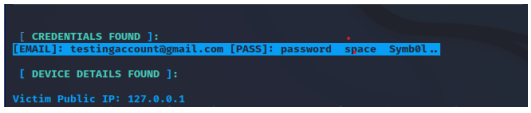


Fig. 6. HiddenEye Phishing Results

The results of the ten tests above are known that the first test by entering the username testingaccount@gmail.com and password: password space Symb0l... Get the test results that this tool has good accuracy to capture passwords from Facebook users because the "space" button is recorded on the password in the first test. Then the second test got the result that not all standard formats in E-mail can be recorded. It is accurate, but this tool cannot tell if the email has a standard format or just the text is recorded. In the third test, Facebook's rule was known that passwords must have at least eight characters to be declared valid. However, this tool only captures the user's password even if there is only one character in the password. And the fifth test got the result that, if the user does not enter the data into the tool, the tool will automatically record even though they did not enter the E-mail and Password. In other words, this accuracy test is said to be accurate, since every recorded data result entered by the user will be recorded as a whole. Although this tool cannot figure out how to format the standard email and password of Facebook users.

*D. Comparative Analysis*

Here is a comparison of the tests performed on the three phishing tools.

TABLE I. COMPARISON BY FEATURE

| No. | Tools | Feature |
|---|---|---|
| 1. | Setoolkit | The available features are easy to implement and the error problems are minimal, when phishing Facebook users there are no problems. The instructions for using the Setoolkit tools are clear and simple |
| 2. | SocialFish | SocialFish does not have any superior features that are effective for carrying out phishing attacks. The features provided by SocialFish are limited to generating fake URLs only and no direct attack features are provided |
| 3. | HiddenEye | HiddenEye features a social media menu to carry out phishing attacks directly. This feature is useful for carrying out phishing attacks without having to enter a URL which is then spoofed |

TABLE II. COMPARISON BASED-ON ACCURACY

| No. | Tools | Feature |
|---|---|---|
| 1. | Setoolkit | Setoolkit cannot record "spaces" on the keyboard, this is inaccurate if the target enters a password that has spaces |
| 2. | SocialFish | SocialFish is inaccurate in phishing because the "space" button is not recorded correctly, it is only recorded once and is not recorded more than twice the space in the character |
| 3. | HiddenEye | The accuracy of HiddenEye tool is very accurate, login data such as Emails and passwords are well recorded even if they use "spaces". Because "space" can be used as a password in an account. Therefore the HiddenEye tool is accurate and effective in phishing |

TABLE III. COMPARISON BASED-ON EASE OF INSTALLATION

| No. | Tools | Feature |
|---|---|---|
| 1. | Setoolkit | Setoolkit is available by default in Kali Linux operating systems. Users do not need to download the installation package from third parties and users can use it directly and easily |
| 2. | SocialFish | SocialFish needs to be downloaded from a third party. The user must copy the original URL to the fake URL as a phishing site, then the fake URL can record the user's login data |
| 3. | HiddenEye | HiddenEye requires downloading installation packages from third parties. The menu feature has been provided by HiddenEye. Users only use the available menu and do not need to create a new URL to launch a phishing attack |

## IV. CONCLUSIONS

Based on the results of a comparison of Setoolkit, HiddenEye, and SocialFish tools conducted ten tests, it can be concluded that:

1. The three tools tested are phishing tools that can be used on the Kali Linux operating system and have the same function, which can phish emails along with the password of the intended account. Despite the differences in how it works, the output remains the same.
2. In phishing, the three tools have their advantages and disadvantages which can be seen from the comparison between the features, accuracy, and convenience offered by each of these tools, such as the features provided by the HiddenEye tools have various features compared to the Setoolkit and SocialFish tools. The accuracy rate of the HiddenEye and Setoolkit tools is higher than that of the SocialFish tool. And the Setoolkit tool is easier to use compared to the other two tools because the tool users do not need to download the Setoolkit tool from third parties.
3. With this test, Facebook social media users know how phishing is done by hackers working on the target Facebook account so that users can be more vigilant and always maintain account security which can be done by downloading the information security guard application.

## ACKNOWLEDGMENT

The researcher would like to thank Mr. Gede Arna Jude Saskara and Mr. I Made Edy Listartha as teaching lecturers at the Ganesha University of Education who have supported and helped in guiding researchers to make articles well.

## REFERENCES

[1] E. Dwi and S. Watie, "Communication and Social Media," *Journal of The Messenger*, vol. 3, no. 2, pp. 69–74, March 2016.

[2] V. Taprial and P. Kanwar, "Classification of Social Media," *Understanding Social Media*, pp. 30, 2012.

[3] P. Study Ilmu Komunikasi, "TREN PENGGUNAAN MEDIA SOSIAL SELAMA PANDEMI DI INDONESIA," 2020.

[4] S. Leitch and M. Warren, "Security Issues Challenging Facebook Security Issues Challenging Facebook Security Issues Challenging Facebook," 2009, doi: 10.4225/75/57b4188730df5.

[5] M. Hanafi, "PENGARUH PENGGUNAAN MEDIA SOSIAL FACEBOOK TERHADAP MOTIVASI BELAJAR MAHASISWA FISIP UNIVERSITAS RIAU," 2016.

[6] M. Betty Yel and M. K. M Nasution, "SECURITY OF PERSONAL DATA INFORMATION ON SOCIAL MEDIA," *JIK)*, VOL. 6, NO. 1, 2022.

[7] "Social Media Security: Leveraging Social Networking While Mitigating Risk – Michael Cross – Google Books."

[8] N. Vadila and A. R. Pratama, "Analisis Kesadaran Keamanan Terhadap Ancaman Phishing," *Automata*, vol. 2, pp 14-17, 2021.

[9] Z. Ramzan, "Phishing Attacks and Countermeasures," *Handbook of Information and Communication Security*, pp. 433-448, 2010.

[10] G. R. Somantri, "Memahami Metode Kualitatif," *Makara Human Behavior Studies in Asia*, vol. 9, no. 2, pp. 57 – 65, Dec. 2005, doi: 10.7454/mssh.v9i2.122.

[11] A. B. Sesuai, Dengan, ASI ttry, and R. Nur Sasongko, "PENELITIAN KOMPARATIF," 2009.

[12] S. Wahyuni, I. M. Raazi, and D. I. Dwitawati, "Analisis Teknik Penyerangan Phishing Pada Social Engineering Terhadap Keamanan Informasi di Media Sosial Profesional Menggunakan Kombinasi Black Eye dan Setoolkit," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 1, 2022.

[13] N. Pavković and L. Perkov, "Social Engineering Toolkit — A systematic approach to social engineering," 2011 Proceedings of the 34th International Convention MIPRO, 2011, pp. 1485-1489.

[14] N. Helminen, K. Tero, and K. Sampo, "Description Information and communication technology," 2021.

[15] "John J Hacking." https://johnjhacking.com/blog/hiddeneye/ (accessed Nov. 15, 2022).

[16] UndeadSec, "UndeadSec SocialFish," June 11, 2022. https://github.com/UndeadSec/SocialFish (accessed November 15, 2022).

[17] Morshmalleo, "Morshmalleo HiddenEye," Sept. 06, 2022. https://github.com/Morsmalleo/HiddenEye (accessed November 19, 2022).

ISSN 2085-4552

# Web-based Writing Learning Application of Basic Hanacaraka Using Convolutional Neural Network Method

Dewi Candani Sulaiman[1], Teady Matius Surya Mulyana[2]
[1,2] Informatics, Faculty of Technology and Design, Universitas Bunda Mulia, Jakarta, Indonesia
[1]dewicandani8@gmail.com, [2]tmulyana@bundamulia.ac.id

Accepted 09 December 2022
Approved 29 December 2022

***Abstract*— The Javanese script, known as Hanacaraka, or Carakan, is one of the traditional Indonesian scripts developed and used on the island of Java. The government's efforts to preserve the use of Javanese language and script by making Javanese a compulsory subject of local content at the education level in Central Java and East Java. In the basic competence of writing, the Javanese script has a complicated shape so that students have difficulty writing and recognizing Javanese script writing. Through this research a web-based basic Javanese writing learning application was designed that can recognize handwriting digitally which aims to help learn basic hanacaraka writing for beginners, especially students at the basic education level in Central Java and East Java. Handwriting Recognition is a system that can recognize handwritten characters and convert them into text that can be read and understood by machines or computers. The handwriting recognition process in this study uses the Convolutional Neural Network (CNN) algorithm which has the capability and ability to recognize patterns in images. Based on the tests that have been carried out between the two architectural models that have been made, the performance of the CNN model that will be used from various experiments has an accuracy of 98.29% and a loss of 0.0746 on the training data. As well as producing an average accuracy value of 99.52%, an average error rate of 0.48%, an overall accuracy of 95.03% and an overall error rate of 4.97%.**

***Index Terms*— convolutional neural network; deep learning; hanacaraka; handwriting recognition**.

## I. INTRODUCTION

The Javanese script, known as Hanacaraka, or Carakan, is one of the traditional Indonesian scripts developed and used on the island of Java, which is a derivative of the Brahmi script originating from India. Javanese script is used in the field of literature, and in the daily life of the Javanese people from the 17th century until now [1]. But as a result of the times, the application of Javanese script in everyday life began to decline or be limited. The Government's efforts to maintain and preserve the use of the Javanese language and script by making Javanese a compulsory subject of local content at the elementary to senior high school levels in Central Java and East Java. In the basic competencies of writing and reading Javanese script, it is a difficult part because it has complicated shapes and complex elements, so that students have difficulty remembering the forms or patterns of Javanese script writing.

In learning a language whose writing is in the form of characters or symbols, regular writing practice is needed. Practicing writing can improve long-term memory, and the ability to recognize things because when writing motor skills work which connects to the nerves of the brain [2]. In this era of digitization, so that writing practice activities such as basic Javanese script are still being carried out so learning applications are needed that can recognize handwriting digitally.

Applications for recognizing digital handwriting can be built with the help of Artificial Intelligence (AI) which can analyze patterns based on data known as deep learning. Deep learning architecture that can recognize patterns in an image is the Convolutional Neural Network (CNN) [3]. Convolutional Neural Network is a deep learning algorithm that can recognize a pattern in an image such as numbers, characters, and in the case of computer vision and natural language processing [4] [5]. Where CNN has good ability and capability in performing image recognition. The recognition system used is online handwriting recognition system.

Based on previous research conducted by Dewa C, Fadhilah A, Afiahayati A with the title "Convolutional Neural Networks for Handwritten Javanese Character Recognition" using the CNN and Multi-Layer Perceptron (MLP) architectures to compare the performance of Javanese script handwriting recognition. Conclusion from this research, the accuracy of handwriting recognition using CNN is higher than using MLP. Although in terms of the time required for training, CNN requires a longer time than MLP [6].

So in this research a comparison will be made between two CNN models. This research was conducted with the aim of knowing the accuracy and

error rate of the convolutional neural network in recognizing hanacaraka handwriting.

## II. METHODOLOGY

### A. Javanese Script (Hanacaraka)

The Javanese script, known as Hanacaraka, or Carakan, is one of the traditional Indonesian scripts that is developing and actively used on the island of Java, especially in Central and East Java [7], and is a derivative of the Brahmi script that originated in India [1]. The existence of the Javanese script cannot be separated from ancient legends. The legend comes from a character named Ajisaka who wrote poetry for his two servants [1]. From this poem, a basic Javanese script was created which has 20 kinds of characters as shown in Figure (1) [6].

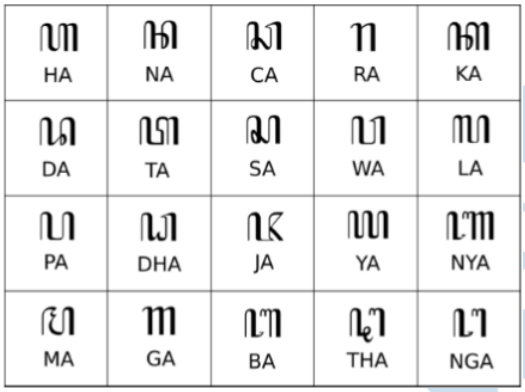


Fig. 1. Basic Javanese Script (Hanacaraka)

### B. Handwriting Recognition System

Handwriting Recognition is a system that can recognize handwriting and convert it into text that can be read and understood by computers. There are two types of handwriting recognition, Offline handwriting recognition which is recognition handwriting derived from an image such as a scanned human handwriting, and Online handwriting recognition which is recognition handwriting derived from digital writing written using a touchpad or touch screen [8] [9]. The Handwriting Recognition process consists of preprocessing, feature extraction, and classification [10].

### C. Convolutional Neural Network

Convolutional Neural Network (CNN) is a deep learning architecture that works based on multi-layer perceptron (MLP) [10]. The CNN architecture consists of a special layer for extracting the input image with a network model connected to the logistic regression classifier [6]. Weights, parameters, and bias play a role in the transformation from the original image to the vector to find out about the nature of the image during the training period. So that CNN can properly recognize the patterns contained in the image, the processed image will go through several stages. The process of CNN is divided into 3 parts, namely as follows [11]:

#### 1) Convolutional Layer

The convolution stage is a filtering stage in the image using a kernel in the form of a matrix which will be multiplied by each selected pixel value. The mathematical representation of the convolution layer is contained in Equation (1) [12].

$$z_{i,j,k}^{1} = w_k^1 * x_{ij}^1 + b_k^1 \quad (1)$$

where $z_{i,j,k}^{l}$ is the result of convolution, where $w_k^l$ is the weight, $x_{ij}^l$ is the coordinate point value (i, j) in the lth layer, and $b_k^l$ is the bias value of the kernel in l layer. After the convolution operation process, the convolution results will be processed using the activation function. The activation functions used are sigmoid, ReLU, tanh [13] [14].

#### 2) Pooling Layer

The pooling layer is a layer that aims to process data to be smaller by reducing dimensions or reducing pixels in convolved images in order to speed up data processing in the next layer. There are 2 types of pooling layers, namely average pooling and max pooling. Between the two pooling, what is often used is max pooling [15].

#### 3) Fully-Connected Layer

Fully-Connected Layer is the last layer on CNN, which is a neural network that is fully connected with the final result of the pooling layer being converted into a vector [15]. Then the output from the fully-connected layer will pass through the activation function, namely the softmax activation function [8]. The representation of the softmax activation function is contained in Equation (2):

$$s(x_n) = \frac{e^{X_n}}{\sum_{k=1}^{K} e^{x_k}} \quad (2)$$

where $s(x_n)$ is the probability value of each class that belongs to, $e^{x_n}$ is the exponential value of x from the nth class, and $\sum_{k=1}^{K} e^{x_k}$ is the sum of all the exponentials of fully-connected output layer [6].

### D. System Development Plan

In the development of a system required planning or design stages of the application to be made. Here are the steps needed:

#### 1) Data Collection

The data needed for this system is a basic hanacaraka dataset or collection of images that will be used for the training and testing process in the form of open-source datasets obtained from Kaggle and additional personal handwriting from researchers, where the dataset is in the form of basic hanacaraka handwritten images consisting of the 20 typeface characters. Each character consists of 150 images. As

much as 80 percent of the dataset will be used for training and as much as 20 percent will be used for testing.

*2) CNN Model Design*

In architectural design, there are 2 types of network architecture from the CNN model that will be used for this training process. Table (1) represents model A, with the simplest architecture consisting of only two convolution layers. Table (2) represents model B, with a higher number of convolution layers.

TABLE I. CNN ARCHITECTURE DETAILS OF MODEL A

| Layer | Size | Output Dimension | Parameter |
|---|---|---|---|
| Input | (100, 100, 1) | - | |
| Convolution 2D + ReLU | 32 (3 x 3) *filters*; *padding* "same" | (None, 100, 100, 32) | 320 |
| Pooling | (2 x 2) *max pool* | (None, 50, 50, 32) | |
| Convolution 2D + ReLU | 48 (3 x 3) *filters*; *padding* "same" | (None, 50, 50, 48) | 13872 |
| Pooling | (2 x 2) *max pool* | (None, 25, 25, 48) | |
| Flatten | | (None, 25 * 25 * 48) | |
| Dense + ReLU | 256 *perceptrons* | (None, 256) | 7680256 |
| Dropout (rate=0.5) | | (None, 256) | |
| Dense + Softmax | 21 *perceptrons* | (None, 21) | 5397 |
| Total Trainable Parameters | | | 7699845 |

TABLE II. COMPARISON BASED-ON ACCURACY

| Layer | Size | Output Dimension | Parameter |
|---|---|---|---|
| Input | (100, 100, 1) | - | |
| Convolution 2D + ReLU | 16 (3 x 3) *filters*; *padding* "same" | (None, 100, 100, 16) | 160 |
| Pooling | (2 x 2) *max pool* | (None, 50, 50, 16) | |
| Convolution 2D + ReLU | 32 (3 x 3) *filters*; *padding* "same" | (None, 50, 50, 32) | 4640 |
| Pooling | (2 x 2) *max pool* | (None, 25, 25, 32) | |
| Convolution 2D + ReLU | 32 (3 x 3) *filters*; *padding* "same" | (None, 25, 25, 32) | 9248 |
| Pooling | (2 x 2) *max pool* | (None, 12, 12, 32) | |
| Convolution 2D + ReLU | 64 (3 x 3) *filters*; *padding* "same" | (None, 12, 12, 64) | 18496 |
| Pooling | (2 x 2) *max pool* | (None, 6, 6, 64) | |
| Convolution 2D + ReLU | 64 (3 x 3) *filters*; *padding* "same" | (None, 6, 6, 64) | 36928 |
| Pooling | (2 x 2) *max pool* | (None, 3, 3, 64) | |
| Flatten | | (None, 576) | |
| Dropout (rate=0.5) | | (None, 576) | |
| Dense + ReLU | 128 *perceptrons* | (None, 128) | 73856 |
| Dense + Softmax | 21 *perceptrons* | (None, 21) | 2709 |
| Total Trainable Parameters | | | 146037 |

*3) System Planning*

The main feature of this application is the character analyzer feature, and the quiz feature. Here is a detailed description of the main features of this application:

- *Character Analyser*

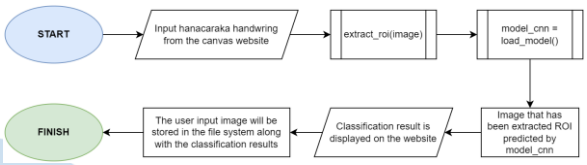


Fig. 2. Flowchart of handwriting image classification on canvas

Based on Figure (2) the handwriting written by the user on the canvas will be the input image into the system. After that, the image will go through a region of interest (ROI) search process from digital writing. The extracted image based on the ROI coordinates will be extracted for its features using the CNN model. Before making predictions, the weights of the CNN model will be entered into the system first. The output from CNN is in the form of predictions or classification results which are Javanese character class classifications from the image being analyzed. The digital writing image written by the user will be stored in the file system in the folder that matches the predicted character.

- *Handwriting Quiz*

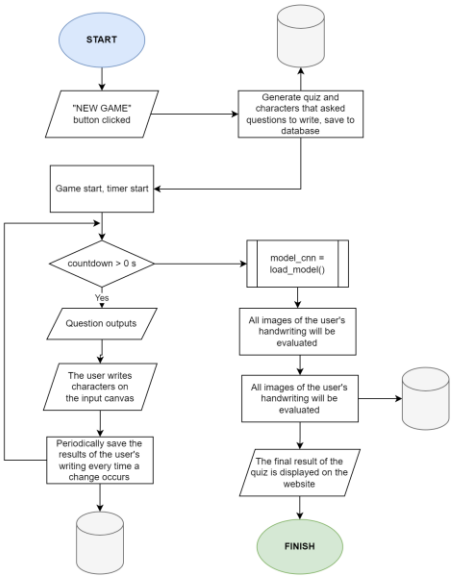


Fig. 3. Flowchart handwriting quiz

In Figure (3) the game will start when the user presses the "Start Game" button. After that the questions will be randomly generated, and the time system will run. Information about quiz questions, and the number of countdowns will be stored in the database. Users can fill in answers and change answers as long as the countdown count is still greater than 0. Any changes to answers that occur will be stored periodically by the system and update the information in the database. If the user has completed all the questions, then the user can submit answers. Quiz will auto-submit if the allotted time has expired. After that each answer will be evaluated using the CNN model, then the system will count the number of answers that were written correctly and the results will be stored in a database which will be displayed as game history.

- *Guess Word Quiz*

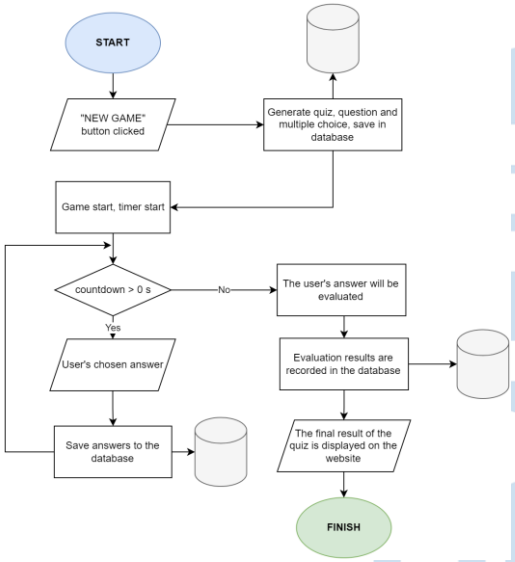


Fig. 4. Flowchart guess word quiz

In Figure (4) the game will start when the user presses the "Start Game" button. After that the questions and choices will be randomly generated, and the time system will run. Information about quiz questions, answer choices and countdown numbers will be stored in the database. Users can fill in answers and change answers as long as the countdown count is still greater than 0. All answers selected by the user will be directly stored in the database. If the user has completed the quiz, then he can submit an answer. If the given countdown time has ended, the system will auto submit. After that, all answers from users will be evaluated. The system will count the number of questions answered correctly, and the evaluation results will be stored in the database. User quiz results will be displayed on the history page on the website.

*4) ERD*

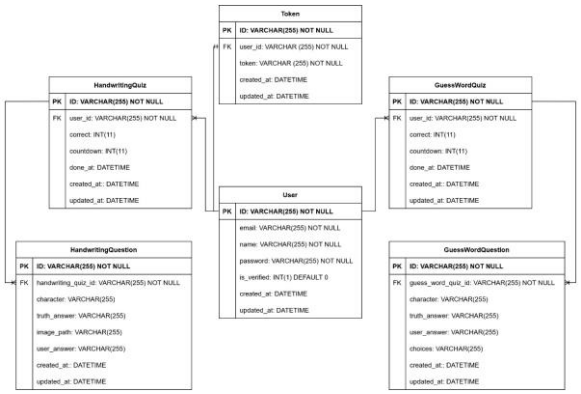


Fig. 5. Entity Relationship Diagram of the system

The designed application uses a database to store user data. The database used is the mySQL database. The following is a table used in the application where each table has an id attribute which is the primary key, the created_at and updated_at attributes are standard attributes for each entity to record when the data row was created and last modified:

- *User Table*
  The user table consists of the attributes name, email, password, and is_verified which is a sign to know whether the account has been verified.

- *Token Table*
  The token table consists of the user_id attribute which is a foreign key and has a relationship with the user table, and the token attribute which contains the token generated by the system for user account verification.

- *HandwritingQuiz Table*
  The handwriting quiz table consists of the user_id attribute which is a foreign key and has a relationship with the user table, the correct attribute to store the number of correct answers to quiz questions, the countdown attribute to store the remaining time to complete the quiz, and the done_at attribute to record when the quiz was completed by the user.

- *HandwritingQuestion Table*
  The HandwritingQuestion table consists of the handwriting_quiz_id attribute which is a foreign key and has a relationship with the HandwritingQuiz table, the character attribute which stores the Javanese characters in question, the truth_answer attribute to store answers to questions, the image_path attribute to store the storage location of each handwritten image from the user, and the user_answer attribute to store user answer data. The HandwritingQuiz table has a one-to-many relationship with the HandwritingQuestion table.

- *GuessWordQuiz Table*
  The guesswordquiz table consists of the user_id attribute which is a foreign key and has a

relationship with the user table, the correct attribute to store the number of correct answers to quiz questions, the countdown attribute to store the remaining time to complete the quiz, and the done_at attribute to record when the quiz was completed by the user.

- *GuessWordQuestion Table*
  The GuessWordQuestion table consists of the guess_word_quiz_id attribute which is a foreign key and has a relationship with the GuessWordQuiz table, the character attribute which stores the Javanese characters in question, the truth_answer attribute to store answers to questions, the user_answer attribute to store user answer data, and the choices attribute to store choices. answers to quiz questions. The GuessWordQuiz table has a one-to-many relationship with the GuessWordQuestion table.

## III. RESULTS AND DISCUSSION

This section covers the interface implementation of the application, and the test results of the system that has been developed

### A. *Interface Implementation*

The interface of this application consists of a login page, sign-up page, verification page, home page (character analyzer), handwriting history page, handwriting quiz page, guess word history page, guess word quiz page. The following is the result of the display of the main features of this application:

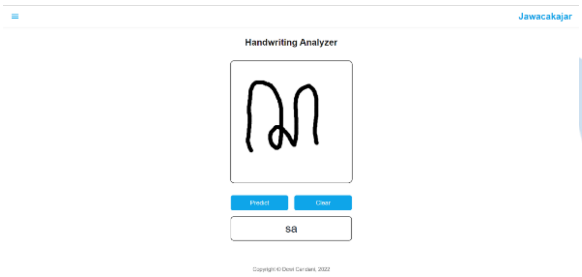


Fig. 6. Home Page Display (character analyzer)

On the home page as shown in Figure (6) there is a canvas to write the basic hanacaraka that will be predicted, the predict button to predict the posts written by the user, and the predicted results will appear in the column below the predict and clear buttons, as well as the clear button to delete the writing on canvas and in the predicted output column.


Fig. 7. Handwriting Quiz view

On the handwriting quiz page as shown in Figure (7) there are 5 questions which if clicked will display the questions as well as the canvas containing the user's answers (canvas write hanacaraka), the canvas where the user's answers are written (canvas write here), the clear button to delete the answers on the canvas write here, and there is also a submit answer button that will appear when all the questions have been answered.

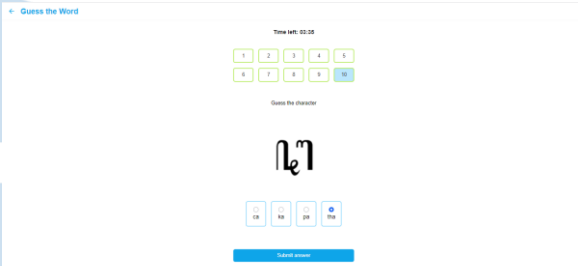


Fig. 8. Guess Word Quiz View

On the guess word quiz page as shown in Figure (8) there are 10 questions which if clicked will display the questions and their choices. There is also a submit answer button that will appear when all the questions have been answered.

### B. *Method Test Results*

Tests were carried out to determine the performance of the developed CNN model. Where there are several factors that will affect how the performance of the CNN model in classifying. These factors are the number of layers of the network, the type of pooling layer used, and so on. There are also other factors such as hyperparameters such as learning rate and epoch.

In this test it will be seen how accuracy and loss are based on the results of the training process. Testing the CNN model will be carried out using the data that has been collected which consists of 20 basic hanacaraka character classes and 1 unknown class (characters that are not basic hanacaraka). Based on the CNN model that has been designed, the results are obtained in the form of accuracy values, and losses from the training and testing processes as shown in the table in the form of Figure (9).

| No. | Model | LR | Epoch | Accuracy | Loss | Testing Accuracy | Testing Loss | Training Time |
|---|---|---|---|---|---|---|---|---|
| 1 | A | 1.00E-03 | 30 | 0.9578 | 0.1257 | 0.9441 | 0.1603 | 384 seconds |
| **2** | **B** | **1.00E-03** | **30** | **0.9815** | **0.0496** | **0.9940** | **0.0570** | **341 seconds** |
| 3 | A | 2.00E-03 | 30 | 0.9525 | 0.1478 | 0.9521 | 0.1707 | 409 seconds |
| 4 | B | 2.00E-03 | 30 | 0.9688 | 0.1010 | 0.9840 | 0.0755 | 361 seconds |
| 5 | A | 1.00E-03 | 50 | 0.9696 | 0.0866 | 0.9641 | 0.1692 | 745 seconds |
| 6 | B | 1.00E-03 | 50 | 0.9886 | 0.0361 | 0.9780 | 0.0900 | 914 seconds |

Fig. 9. Accuracy and Loss Value in the Training and Testing Process

From several training processes, it was found that the loss value was relatively low compared to other training trials, namely in the 2nd and 6th training, which is model B. From these results, model B shows relatively better performance based on the loss value and its accuracy if compared to model A.

The most important factors that affect the learning capacity of a deep learning model are the number of layers and the number of perceptrons that a model has. A model that has a sufficient number of layers and perceptrons will produce good performance. The more layers that can represent data, the more feature information will be extracted. This is the reason model B has better performance than model A.

Of the two candidate models trained, namely the 2nd and 6th training trial model, the training results from the 2nd model will be selected to carry out the hanacaraka classification process on the system of the application to be developed. Figure (10) is a graph of accuracy and loss from the training results of the 2nd experimental model and Figure (11) is a graph of accuracy and loss from the training results of the 6th experimental model.

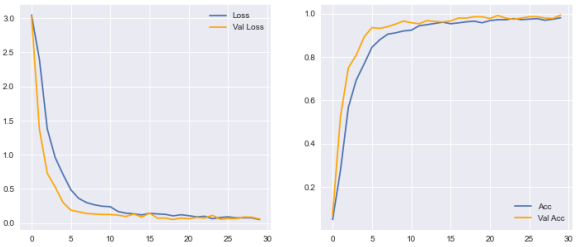


Fig. 10. Graph of loss (left) and accuracy (right) from the results of the 2nd training

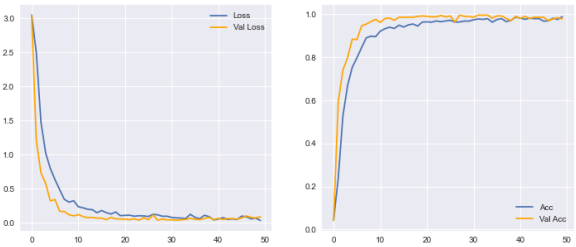


Fig. 11. Graph of loss (left) and accuracy (right) from the results of the 6th training

If you look at the graph in Figure (11), it can be seen that the accuracy and loss values during the training epoch have begun to stabilize and stagnate when entering the 30th epoch, and there are no longer significant changes in the accuracy and loss values. Therefore, if the training process is continued, it has a high potential to cause the model to "memorize" the given training dataset which in the end if it is continued it can cause overfitting. A significant decrease in the loss value can be seen in the epoch range from the 1st to the 10th epoch, and then the decreasing gradient starts to slope from the 10th to the 20th epoch. Due to the reasons above, for this study it was concluded that the CNN model from the results of the 2nd training experiment will be used in the Javanese character recognition system which will be implemented in learning applications.

Calculation of the accuracy and error rate of the CNN model in handwriting classification will use a confusion matrix for multiclass classification cases. To find out the accuracy and error rate of a model, you can do it by comparing the results of the predicted class with the actual class. All details regarding the true positive, true negative, false positive, false negative, accuracy and error rate of each class of character classification results from the 2nd training result model can be seen in Table (3), and the 6th training in Table (4).

TABLE III. ACCURACY AND ERROR RATE OF EACH CHARACTER CLASS CLASSIFICATION RESULTS OF THE 2ND TRAINING MODEL

| Lable | TP | TN | FP | FN | Akurasi | Error |
|---|---|---|---|---|---|---|
| ba | 24 | 478 | 2 | 0 | 0.996031746 | 0.003968254 |
| ca | 24 | 480 | 0 | 0 | 1 | 0 |
| da | 24 | 478 | 2 | 0 | 0.996031746 | 0.003968254 |
| dha | 24 | 480 | 0 | 0 | 1 | 0 |
| ga | 24 | 479 | 1 | 0 | 0.998015873 | 0.001984127 |
| ha | 19 | 479 | 1 | 5 | 0.988095238 | 0.011904762 |
| ja | 22 | 480 | 0 | 2 | 0.996031746 | 0.003968254 |
| ka | 24 | 476 | 4 | 0 | 0.992063492 | 0.007936508 |
| la | 24 | 477 | 3 | 0 | 0.994047619 | 0.005952381 |
| ma | 24 | 480 | 0 | 0 | 1 | 0 |
| na | 19 | 480 | 0 | 5 | 0.990079365 | 0.009920635 |
| nga | 24 | 480 | 0 | 0 | 1 | 0 |
| nya | 24 | 480 | 0 | 0 | 1 | 0 |
| pa | 24 | 474 | 6 | 0 | 0.988095238 | 0.011904762 |
| ra | 21 | 480 | 0 | 3 | 0.994047619 | 0.005952381 |
| sa | 24 | 480 | 0 | 0 | 1 | 0 |
| ta | 24 | 478 | 2 | 0 | 0.996031746 | 0.003968254 |
| tha | 24 | 480 | 0 | 0 | 1 | 0 |
| unknown | 22 | 477 | 3 | 2 | 0.990079365 | 0.009920635 |
| wa | 16 | 480 | 0 | 8 | 0.984126984 | 0.015873016 |
| ya | 24 | 479 | 1 | 0 | 0.998015873 | 0.001984127 |

TABLE IV. ACCURACY AND ERROR RATE OF EACH CHARACTER CLASS CLASSIFICATION RESULTS OF THE 6TH TRAINING MODEL

| Label | TP | TN | FP | FN | Akurasi | Error |
|---|---|---|---|---|---|---|
| ba | 24 | 480 | 0 | 0 | 1 | 0 |
| ca | 24 | 480 | 0 | 0 | 1 | 0 |
| da | 24 | 480 | 0 | 0 | 1 | 0 |
| dha | 24 | 479 | 1 | 0 | 0.998015873 | 0.001984127 |
| ga | 23 | 480 | 0 | 1 | 0.998015873 | 0.001984127 |
| ha | 19 | 480 | 0 | 5 | 0.990079365 | 0.009920635 |
| ja | 24 | 480 | 0 | 0 | 1 | 0 |
| ka | 23 | 479 | 1 | 1 | 0.996031746 | 0.003968254 |

| | | | | | | |
|---|---|---|---|---|---|---|
| la | 20 | 479 | 1 | 4 | 0.990079365 | 0.009920635 |
| ma | 23 | 480 | 0 | 1 | 0.998015873 | 0.001984127 |
| na | 21 | 479 | 1 | 3 | 0.992063492 | 0.007936508 |
| nga | 22 | 480 | 0 | 2 | 0.996031746 | 0.003968254 |
| nya | 23 | 480 | 0 | 1 | 0.998015873 | 0.001984127 |
| pa | 24 | 469 | 11 | 0 | 0.978174603 | 0.021825397 |
| ra | 23 | 479 | 1 | 1 | 0.996031746 | 0.003968254 |
| sa | 23 | 479 | 1 | 1 | 0.996031746 | 0.003968254 |
| ta | 23 | 480 | 0 | 1 | 0.998015873 | 0.001984127 |
| tha | 24 | 480 | 0 | 0 | 1 | 0 |
| unknown | 22 | 467 | 13 | 2 | 0.970238095 | 0.029761905 |
| wa | 19 | 480 | 0 | 5 | 0.990079365 | 0.009920635 |
| ya | 21 | 479 | 1 | 3 | 0.992063492 | 0.007936508 |

From the calculation of accuracy, error rate, true positive, true negative, false positive, and false negative values of each class resulting from the 2nd and 6th training model classification, it can then be calculated average accuracy, average error rate, overall accuracy, and overall error rate as a metric from the confusion matrix which will be used in this test to compare the performance of the two models. The results of calculating the metrics from the confusion matrix used can be seen in Table (5).

TABLE V. CONFUSION MATRIX METRICS FROM TESTING CLASSIFICATION MODEL RESULTS OF THE 2ND AND 6TH TRAINING

| Test No | Average Accuracy | Average Error Rate | Overall Accuracy | Overall Error Rate |
|---|---|---|---|---|
| **2** | **0.9952** | **0.0048** | **0.9503** | **0.0497** |
| 6 | 0.9941 | 0.0059 | 0.9385 | 0.0615 |

Based on Table (5), it can be seen that the classification performance of the 2nd training model has a higher overall accuracy rate, and a lower overall error rate when compared to the 6th training model. It is from these results that support the reason why the 2nd training result model was selected and implemented as a model for introducing basic Javanese script in the application being developed.

## IV. CONCLUSIONS

Based on the tests that have been carried out, researchers can draw several conclusions, namely:

1. The use of deep learning methods is quite effective in solving complex classification cases such as the use of convolutional neural networks to develop basic hanacaraka character recognition systems.
2. The performance of the CNN model used in the application is obtained after various experiments with an accuracy of 98.29%, and a loss of 0.0746 on the training data. In data testing obtained an accuracy of 97.82%, and a loss of 0.0924.
3. Tests carried out using the multiclass confusion matrix produced an average accuracy value of 99.52%, an average error rate of 0.48%, an overall accuracy of 95.03% and an overall error rate of 4.97%.

## REFERENCES


[1] A. Arismadhani, U. L. Yuhana, and I. Kuswardayan, "Aplikasi Belajar Menulis Aksara Jawa Menggunakan Android," *Jurnal Teknik Pomits*, vol. 2, no. 1, 2013.

[2] A. M. Silva and R. Limongi, "Writing to Learn Increases Long-term Memory Consolidation: A Mental-chronometry and Computational-modeling Study of 'Epistemic Writing,'" *J Writ Res*, vol. 11, no. vol. 11 issue 1, pp. 211–243, Jun. 2019, doi: 10.17239/jowr-2019.11.01.07.

[3] M. R. Alwanda, R. P. K. Ramadhan, and D. Alamsyah, "Implementasi Metode Convolutional Neural Network Menggunakan Arsitektur LeNet-5 untuk Pengenalan Doodle," *Jurnal Algoritme*, vol. 1, no. 1, pp. 45–56, Oct. 2020, doi: 10.35957/algoritme.v1i1.434.

[4] Y. Sun, B. Xue, M. Zhang, and G. G. Yen, "Evolving Deep Convolutional Neural Networks for Image Classification," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 2, 2020, doi: 10.1109/TEVC.2019.2916183.

[5] K. Wróbel, M. Karwatowski, M. Wielgosz, M. Pietroń, and K. Wiatr, "Compression of Convolutional Neural Network for Natural Language Processing," *Computer Science*, vol. 21, no. 1, 2020, doi: 10.7494/csci.2020.21.1.3375.

[6] C. K. Dewa, A. L. Fadhilah, and A. Afiahayati, "Convolutional Neural Networks for Handwritten Javanese Character Recognition," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 12, no. 1, 2018, doi: 10.22146/ijccs.31144.

[7] T. Matius Surya Mulyana, D. Widyaningrum, P. Studi Informatika, F. Teknologi dan Desain, U. Bunda Mulia Jl Raya Lodan, and J. Utara, "OCR HURUF JAWA DENGAN FITUR KODE RANTAI DAN LEVENSHTEIN DISTANCE."

[8] R. Plamondon and S. N. Srihari, "Online and off-line handwriting recognition: a comprehensive survey," *IEEE Trans Pattern Anal Mach Intell*, vol. 22, no. 1, pp. 63–84, 2000, doi: 10.1109/34.824821.

[9] P. C. Vashist, A. Pandey, and A. Tripathi, "A Comparative Study of Handwriting Recognition Techniques," in *2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM)*, IEEE, Jan. 2020, pp. 456–461. doi: 10.1109/ICCAKM46823.2020.9051464.

[10] I. Khandokar, M. Hasan, F. Ernawan, S. Islam, and M. N. Kabir, "Handwritten character recognition using convolutional neural network," in *Journal of Physics: Conference Series*, 2021. doi: 10.1088/1742-6596/1918/4/042152.

[11] Z. J. Wang *et al.*, "CNN Explainer: Learning Convolutional Neural Networks with Interactive Visualization," *IEEE Trans Vis Comput Graph*, vol. 27, no. 2, 2021, doi: 10.1109/TVCG.2020.3030418.

[12] J. Briliantio, N. Santosa, G. Ardian, and L. Hakim, "Penerapan Convolutional Neural Network untuk Handwriting Recognition pada Aplikasi Belajar Aritmatika Dasar Berbasis Web," *Jurnal Teknik Informatika Unika St. Thomas*, vol. 05, no. 02, 2020.

[13] M. Z. Alom *et al.*, "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches," Mar. 2018.

[14] Y. Tian, "Artificial Intelligence Image Recognition Method Based on Convolutional Neural Network Algorithm," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.3006097.

[15] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network-A Deep Learning Approach," in *Procedia Computer Science*, 2018. doi: 10.1016/j.procs.2018.05.069.

ISSN 2085-4552

# Gamification To Study Machine Learning Using Octalysis Framework

Monika Evelin Johan[1], Samuel Ady Sanjaya[2], Andreas Yogi Brata[3], Andrew Brilyant[4]
[1,2,3.4] Department of Information System, Universitas Multimedia Nusantara, Tangerang, Indonesia
[1]monika.evelin@umn.ac.id, [2]samuel.ady@umn.ac.id, [3]andreas.brata@student.umn.ac.id,
[4]andrew.brillyant@student.umn.ac.id

Accepted 16 March 2023
Approved 31 May 2023

***Abstract*— Machine learning is an area of study that enables computers to learn and improve their performance without explicit programming. Its primary application is assisting humans in processing vast amounts of data. Machine learning is a crucial area of study, and to assist students in learning about it, a gamification application has been developed using the Android platform. The aim of this research is to not only provide knowledge to students but also make the application interesting by using the Octalysis Framework. The application's level of acceptance will be tested on students at Multimedia Nusantara University who are taking machine learning courses. The Unified Theory of Acceptance and Use of Technology Model (UTAUT) method will be used to evaluate the application, which received an average score of 86.77%. This score indicates that students strongly agreed that the gamification approach using the Octalysis Framework and Android platform for machine learning was well-received.**

***Index Terms*— android application; gamification; machine learning; octalysis framework**.

## I. INTRODUCTION

Machine learning is a field of study that enables computers to learn without being explicitly programmed, where the computer can then work according to what it has learned. [1]. Machine learning is can be used to process vast quantities of data, often in diverse formats, that are beyond human ability to process quickly. Because of that, machine learning is needed to help humans in processing data. Machine learning is commonly used to teach machines how to process data efficiently so that large amounts of data can be processed and become valuable information. Machine learning has several different algorithms for solving problems with data [2]. Machine learning can also be used to make forecasts with the available data. With the various benefits of machine learning that are useful today, machine learning is a subject that should be passed down to the current generation, as it enables them to make good use of future data and extract valuable insights from it.

Gamification is a way to hand down knowledge using an element from the game, such as challenge and feedback to help the learning process. Gamification is a method for using in-game mechanisms for non-game environments [3]. Gamification takes several in-game elements such as badges, rankings, rankings, etc., and uses them to transform non-game environments into game environments. One of the uses of this gamification is to assist in the field of learning.

The gamification method has been proven to be able to meet its target in channeling machine learning knowledge. A study showed that there were workshops teaching machine learning to junior high school students using the gamification method which was carried out in Thailand [4]. In this study, it was explained that junior high school students were given an explanation about machine learning and were given the challenge to classify sweet mangoes by looking at the physical shape of the mangoes, such as the color of the mangoes, the hardness of the fruit, and so on. The students then entered the mango data into an application called Rapid Miner, which is a UI-based application for machine learning for classification. In the end the students succeeded in classifying the mangoes, and based on observations, the students looked happy and the model built by the students proved to be successful.

Research [5] concludes that some of the studies that have been reviewed have weaknesses, such as the absence of an appropriate framework to use. Researchers are too focused on making basic programming studies so they don't pay attention to the fun factor to encourage users to complete games. So, this research will use the Octalysis Framework in designing machine learning gamification applications. Octalysis framework is an approach to creating a gamification system that was introduced by Yu-Kai Chou [6]. The Octalysis framework assumes that in gamification there must be encouragement so that users want to interact and interact in the game. The driving factors are divided into eight in the Octalysis Framework, including epic meaning, development and accomplishment, empowerment of creativity and feedback, ownership and possession, social influence and relatedness, scarcity and impatience, unpredictability and curiosity, and loss and avoidance [7]. By using the Octalysis Framework approach, users get encouragement so that users want to interact in the

game. Gamification with Octalysis Framework has also been proven to transmit knowledge, such as the Java programming language [8], and machine learning. The gamification technique built using the Octalysis Framework method is not only aimed at young people, but has also been tried on elder people [9]. By focusing on the 8 core drivers of the Octalysis Framework, the application of gamification techniques in education has the opportunity to create experiential independent learning, for all ages, and can be expanded exponentially [10].

This research using RAD (Rapid Application Development) method to develop the mobile gamification application. The RAD system development method was chosen because it allows researchers to be able to change the design of the application during the application development process [11]. In addition, the application development process can be revised repeatedly on the prototype that has been made, so that the developer does not need a long time in the prototyping process.

With the gamification method using the Octalysis Framework that has the potential to help the teaching process of machine learning knowledge, this research will design a mobile gamification application using Octalysis Framework for Android platform. Then at the end of this research, to evaluate the resulting application this research will use the UTAUT model, which has proven in evaluating a new information technology [12]. This application is expected to help students who are studying machine learning to gain knowledge and can also increase student interest in learning machine learning.

## II. METHODOLOGY

### A. *Object of Research*

The object of this research is to design and build a mobile machine learning gamification application for students at Multimedia Nusantara University, department of Information Systems study program who are studying machine learning. The mobile machine learning gamification application that has been created will later be tested for user acceptance using the UTAUT (Unified Theory of Acceptance and Use of Technology) model to be able to measure the acceptance rate of Multimedia Nusantara University student's batch 2019 Information Systems study program who are studying machine learning. The applications that have been built from this research will be distributed to students, then those who are needed and available to try the application will be provided with the UTAUT model. In this study there were as many as 13 students who were available to be respondents of this research and had the necessary requirements, namely having an Android phone. Then the results of the entire research will be analyzed and will be a suggestion for future research.

### B. *Research Method*

RAD (Rapid Applications Development) is a method for system development. RAD itself makes it possible to build an application quickly and is suitable for use if you don't have a large number of teams. In the RAD method, developers can create an application by prioritizing feedback from users, so that the prototype that has been made can be revised repeatedly so as to get optimal application results. Figure 1 explaining about the steps available on RAD [13]. In the RAD method, the application development process can be revised repeatedly on the prototype that has been made, so that the developer does not need a long time in the prototyping process.[11].

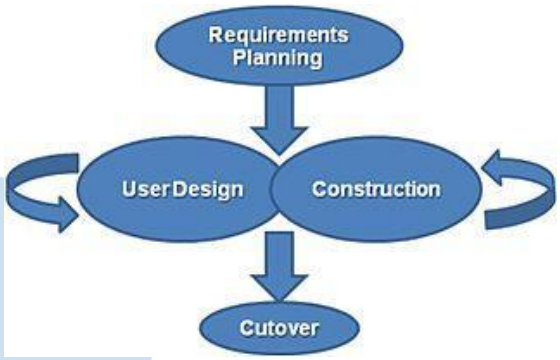


Fig. 1. Steps Available on RAD

#### 1) *Requirements Planning*

In this research, the goal is to help students who are studying about machine learning in a more engaging way. Therefore, in this research, we built a mobile machine learning gamification using the Octalysis Framework. Observations are made by analyzing material about machine learning that needs to be learned, then studying techniques for making gamification according to the Octalysis Framework.

#### 2) *User Design*

In this step, researchers plan the application to be built according to the Octalysis Framework. Therefore, the researcher ensures that all eight cores contained in the Octalysis Framework are contained in machine learning gamification applications. The following is a user design of each core in the Octalysis Framework which will be implemented on the construction site.

- Epic Meaning and Calling

  This core is about driving people to believe that they are chosen to do something. This core will be implemented when the user plays for the first time, there will be words indicating that the user is the chosen one to be able to finish this game.

- Development and Accomplishment

  This core is about driving people to always make development to accomplish some goals. This core will be implemented when the user clicks

the right answer, and the user will get additional levels and scores.

- Empowerment of Creativity and Feedback

  This core focuses on empowering the user's creativity and giving the user's feedback right after they use their creativity. This core will be implemented in giving users creativity to create their own in-game name, and the name will be displayed on the leaderboard.

- Ownership and Possession

  This core is about giving user the feel that they own something and triggering the user to improve it. This core will be implemented with the user's score and level. Every user in this game has their own level and score and every user wants to improve their level and score.

- Social Influence and Relatedness

  This core is about all social aspects that can motivate people. This core will be implemented with a leaderboard that can display the rank of each player that plays the game.

- Scarcity and Impatience

  People tend to want something that is rare and not immediately attainable. This core will be implemented with each stage in this game and cannot be passed without completing the previous stage.

- Unpredictability and Curiosity

  This core is about driving people to be curious about something because of the unpredictability. This core will be implemented with different questions at each stage.

- Loss and Avoidance

  This core is about driving people to avoid them from losing something. This score will be implemented when the user clicks the wrong button or input the wrong answer, then the user's score will be reduced by 25 points.

*3) Construction*

In this step, researcher will build the application using C# and the PHP programming language. The implementation of this research resulted in a machine learning gamification application.

*4) Cutover*

The cutover phase is carried out to measure the level of student acceptance in machine learning gamification applications that have been designed by distributing machine learning gamification applications on the Android platform. Researchers distributed Google Drive links to students who were taking machine learning courses, then distributed questionnaires in the form of google forms to fill out. The questions in the questionnaire distributed are based on the unified theory of acceptance and use of technology (UTAUT) model which can measure the level of acceptance of new technologies in students [14] [15].

TABLE I. ACCEPTANCE LEVEL CATEGORY INTERVAL

| Percentage Interval | Acceptance Level Category |
|---|---|
| 0% - 20% | Strongly Disagree |
| 20.01% - 40% | Disagree |
| 40.01% - 60% | Uncertain |
| 60.01% - 80% | Agree |
| 80.01% - 100% | Strongly Agree |

## III. RESULTS AND DISCUSSION

### *A. Octalysis Framework Implementation*

The following is a user design plan for each core in the Octalysis Framework that will be implemented in the construction phase.

*1) Epic Meaning and Calling*

When the user first enters the system, there are words to indicate that the user is the chosen person who can complete all the existing stages.

*2) Development and Accomplishment*

If the user has completed the existing stage, then the user will get an additional level and also an additional score.

*3) Empowerment of Creativity and Feedback*

Users get the freedom to write the desired name so that it can appear on the leaderboard.

*4) Ownership and Possession*

In the first step, open the Kali Linux command terminal.

*5) Social Influence and Relatedness*

There is a leaderboard that displays the rank of each user so that it can encourage users to compete between friends to get the most scores.

*6) Scarcity and Impatience*

Every stage in this game cannot be skipped. The user must complete the previous stage in order to advance to the next stage.

*7) Unpredictability and Curiosity*

There are several different questions at each existing stage that increase user interest in advancing to the next stage.

*8) Loss and Avoidance*

If the user presses or enters the wrong answer, then the user's score will be reduced by 25 points.

The following is a flowchart to help understand the application to be built.

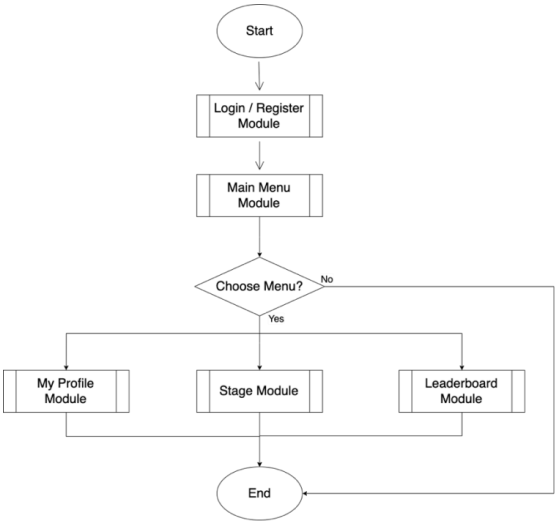


Fig. 2. Main Flowchart

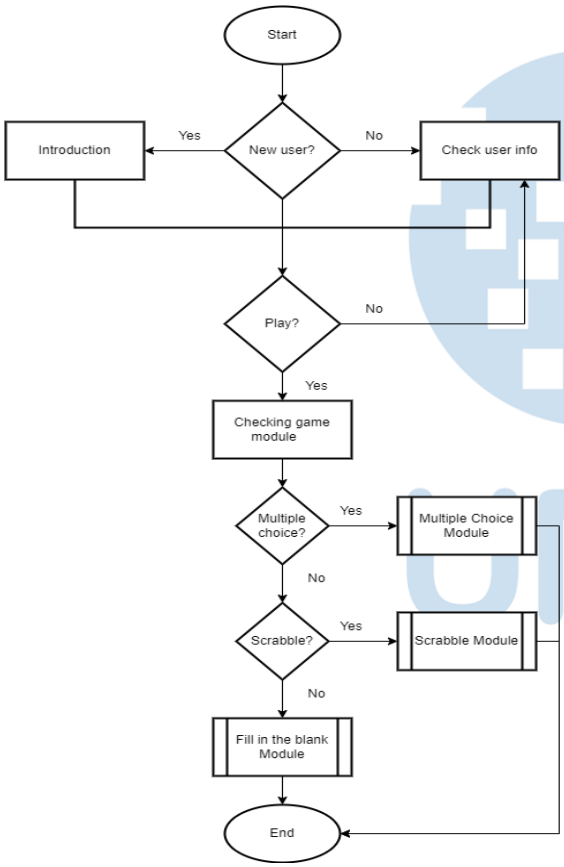


Fig. 3. Stage Module Flowchart

Figure 3 above is the flow of the stage module. If the user presses the play button on the main menu page, the user will be taken to the stage module. For users who have just registered their account, an introduction will appear which contains words that make the user feel special or selected. This is in accordance with the Octalysis Framework factor, epic meaning and calling. Then if the user has played this game before, the level information will be checked and adjusted to the stage button that can be accessed. Then when the user presses the stage button, the system will see the game module at that stage, whether it is a multiple choice, scrabble, or fill the blank type stage.

### *B. Construction*

The application was built using C# and PHP programming language, and uses MySQL to store user data so the leaderboard can display all the player's rank. To use this application, a new user has to register their account so they can log in into the game.

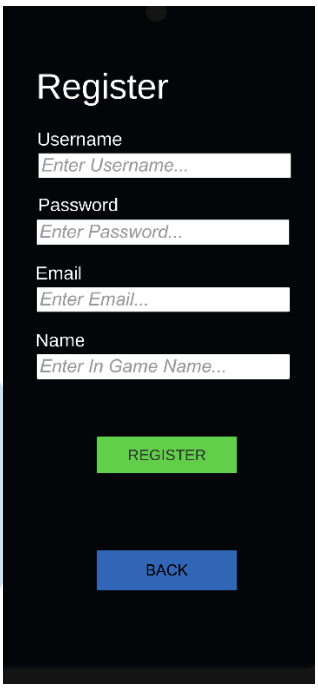


Fig. 4. Register Page

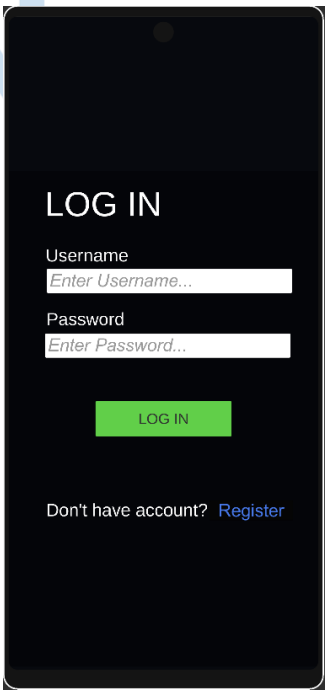


Fig. 5. Login Page

After the user has successfully logged in, the user will see the main menu module as shown in the Figure 6.


Fig. 6. Main Menu Page

Figure 6 is the page that users see when they have successfully logged into the game. On the main menu page, users can choose 4 menus, namely Play, Leaderboard, My Profile, and Quit. If the user presses the Play button, then the user can play the game. Then if the user presses the Leaderboard button, the user will go to the Leaderboard page which contains the ranking of each user who plays this game. If the user presses the My Profile button, the user will go to a page that displays information about the user's account. Finally, if the user presses the Quit button, the user will exit my Machine Learning application.

On the Leaderboard page, users can see the rank that the user has or see the rank that is owned by the user's friends or other users. The rank that appears on the leaderboard adjusts to the number of scores owned by the user. This score is one of the elements used in the Octalysis Framework, ownership and possession which means the user has something, in this case the user has a score that will be obtained when the user answers the question correctly. Then there are also elements of social influence and relatedness that are applied in the form of a leaderboard, which can challenge users to compete with other users.


Fig. 7. Leaderboard Page

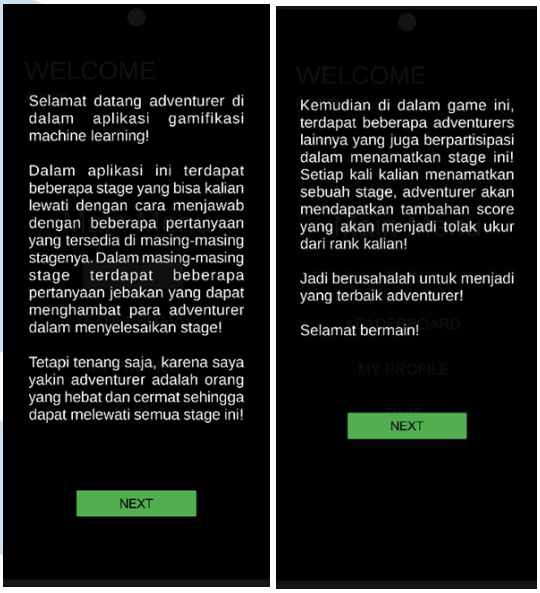


Fig. 8. Introduction Page

Figure 8 is the user introduction page. Users who play my Machine Learning application for the first time will see this page when they tap the Play button. This page displays making the user believe that the user is the chosen person who can complete all the stages in this game. This is an example of the application of the Octalysis Framework, epic meaning and calling, which makes players believe there is something greater than themselves or the player is someone who is chosen.

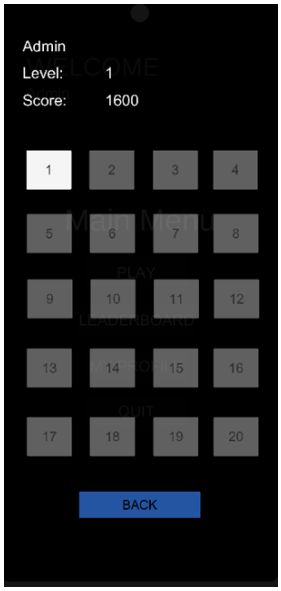


Fig. 9. 20 Stages of the Game

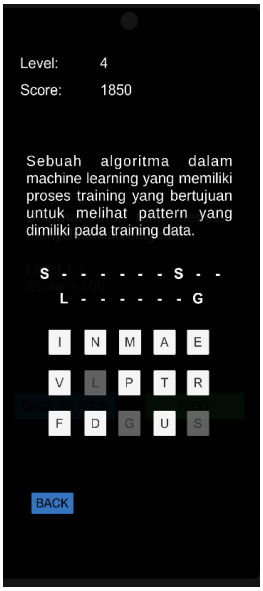


Fig. 11. Scrabble Page

After tap Next button, user can see the stage list that is available at this game. This game has 20 stages, and users have to clear each stage in order to proceed to the next stage. In each stage, there will be a different machine learning question. From the 20 stages, there are 3 types of game. The first is multiple choice, the second is scrabble, and the last is fill the blank. An example of multiple choice shown in Figure 10, scrabble stage in Figure 11, and Fill in the Blank Stage in Figure 12.

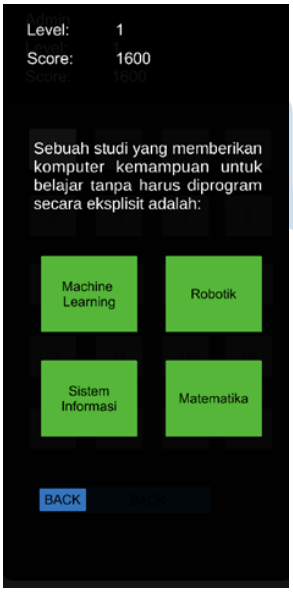


Fig. 10. Multiple Choice Stage

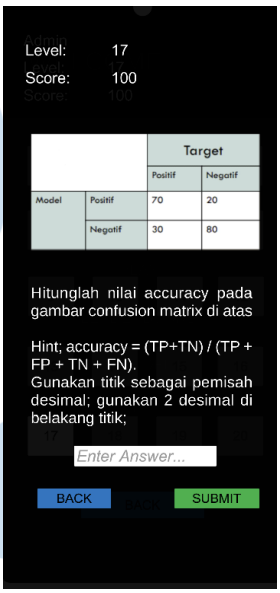


Fig. 12. Fill in the Blank Stage

The stage with the fill the blank type is one of the implementations of the Octalysis Framework at the point of unpredictability and curiosity, which makes this application unpredictable because there is an unexpected type of stage that suddenly appears in the final stage.

For each question, if the user successfully answers the question, a successful answer page will appear as shown in Figure 13 and get 100 points. But if the user presses the wrong button, then the user will lose the 25 score points they have. This is in accordance with one of the points in the Octalysis Framework, namely loss and avoidance. With the reduction of the user's score points, the user will try to answer questions in a stage correctly and will avoid the risk of losing the score points they have.

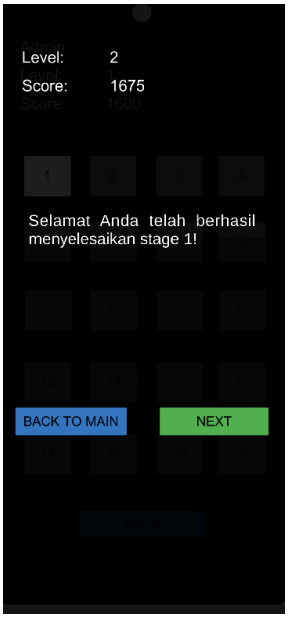


Fig. 13. Interface if User Answer Correctly

### C. *Cutover*

The cutover stage is carried out to measure the level of student acceptance of the machine learning gamification application that has been designed by sharing the machine learning gamification application on the Android platform. Measurements were made by distributing questionnaires to students who were testing the application. The questions on the questionnaire distributed were made based on the unified theory of acceptance and use of technology (UTAUT) model which can measure the level of acceptance of new technology in students. Then from all the list of questions, will use a Likert scale to calculate the level of acceptance of student information. Each of the questionnaire questions was answered using a scale of one to five, where one means strongly disagrees and five indicates strongly agree. The results of the questionnaire that has been calculated using the UTAUT method are fully discussed in the next chapter.

## IV. TESTING AND IMPLEMENTATION

The completed application will be tested using the UTAUT model. The UTAUT model has five indicators to measure the level of acceptance of new information technology. The indicators consist of performance expectancy, effort expectancy, social influence, facilitating conditions, and behavioral intentions. The data obtained for this test was obtained from a questionnaire distributed to the sixth semester students who are studying machine learning. In this study, there were 13 students who were willing to become respondents and had the requirement to have an Android phone. Then to be able to measure indicators in the UTAUT model, this study uses a Likert [16] scale. Table II show the results of each indicator in the UTAUT model.

TABLE II. ACCEPTANCE LEVEL CATEGORY INTERVAL

| No. | UTAUT Model Indicator | Average of Acceptance | Acceptance Rate Category |
|---|---|---|---|
| 1 | Effort Expectancy | 89.23% | Strongly Agree |
| 2 | Performance Expectancy | 88.46% | Strongly Agree |
| 3 | Social Influence | 84.61% | Strongly Agree |
| 4 | Facilitating Condition | 87.69% | Strongly Agree |
| 5 | Behavioral Intention | 83.84% | Strongly Agree |
| Average total acceptance percentage | | 86.77% | Strongly Agree |

Table II is the calculation result of each UTAUT model indicator. It can be seen that the average percentage of acceptance on the effort expectancy indicator is 88.46%. Then on the UTAUT performance expectancy indicator, there is an average acceptance percentage of 88.46%. Then on the UTAUT social influence model indicator, the average percentage of acceptance is 84.61%. Then on the UTAUT facilitating condition indicator, the average percentage of acceptance is 87.69%. And finally on the UTAUT behavioral intention indicator, the average percentage of acceptance is 83.84. Then from all the indicators of the UTAUT model, the average percentage of acceptance is 86.77%. Based on the Table 2, it can be concluded that the respondents strongly agree that this machine learning gamification application can be accepted by students who are studying machine learning. The lowest score is on behavioral intention, which is triggered by a display that is still simple so that it is less attractive to users to use the application. Therefore, changes to the appearance are made to increase user interest as shown in Figure 14 below.

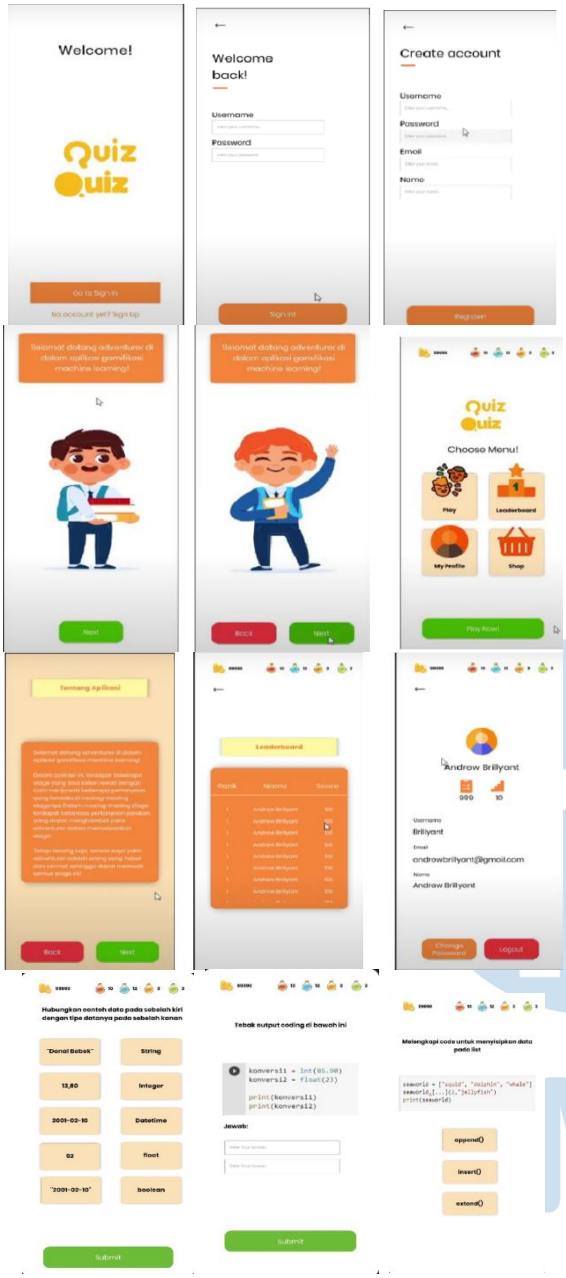


Fig. 14. Interface if User Answer Correctly

## V. CONCLUSION AND SUGGESTION

### A. *Conclusion*

In this study, the machine learning gamification application using the Octalysis Framework gamification method has been designed using the C# programming language and using the UNITY game engine platform. This gamification application has also been connected to the MySQL database with PHP to connect the front end with the C# programming language. Then to support the live leaderboard feature, this research uses hosting so that users can access this application online. This machine learning gamification application can already be installed and used by downloading the apk file that can be found on the researcher's google drive. Then the machine learning gamification application as a result of this research has been evaluated by 13 respondents using the UTAUT model to see the percentage of acceptance by students who are taking machine learning courses. From the distributed questionnaires, the average percentage figure for all UTAUT factors is 86.77%, which can be concluded that the respondents strongly agree that this machine learning gamification application can be accepted by students who are studying machine learning.

### B. *Suggestions*

In this study, there are several suggestions that can be given that can be used for the development of machine learning gamification applications in the future such as, added a friend list feature to the application and created a leaderboard that can be filtered based on friends, as well as several categories so that users get an incentive to get number one ranking, and it is also possible to monitor student progress, then added a database for questions for each stage so that it is possible to create new stages in machine learning gamification applications more practically, so that users can learn more about machine learning, and lastly adding a type of stage that allows the user to fill in more complicated coding, so that it will add variety to the questions contained in each stage.

## ACKNOWLEDGMENT

The authors would like to thanks Universitas Multimedia Nusantara for supporting this research.

## REFERENCES

[1] B. Mahesh, "Machine learning algorithms-a review," *International Journal of Science and Research (IJSR)*, vol. 9, no. 1, pp. 381-386, 2020.

[2] J. Alzubi, A. Nayyar and A. Kumar, "Machine learning from theory to algorithms: an overview," *Journal of physics: conference series*, vol. 1142, no. 1, p. 012012, 2018.

[3] E. Çeker and F. Özdaml, "What" Gamification" Is and What It's Not.," *European Journal of Contemporary Education*, vol. 6, no. 2, pp. 221-228, 2017.

[4] B. Sakulkueakulsuk, S. Witoon, P. Ngarmkajornwiwat, P. Pataranutaporn, W. Surareungchai, P. Pataranutaporn and P. Subsoontorn, "Kids making AI: Integrating Machine Learning, Gamification, and Social Context in STEM Education," *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, pp. 1005-1010, 2018.

[5] M. Shahid, A. Wajid, K. U. Haq, I. Saleem and A. H. Shujja, "A review of gamification for learning programming fundamental," in *2019 International Conference on Innovative Computing (ICIC)*, Lahore, Pakistan, 2019.

[6] Y. K. Chou, "yukaichou.com," [Online]. Available: https://yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/. [Accessed May 2022].

[7] F. Marisa, A. L. Maukar, S. S. S. Ahmad, R. D. Marcus and A. H. Widodo, "Evaluation of student core drives on e-learning during the covid-19 with octalysis gamification framework," *International Journal of Advanced Computer Science and Applications*, vol. 11, p. 11, 2020.

[8] L. Christopher and A. Waworuntu, "Java Programming Language Learning Application Based on Octalysis

Gamification Framework," *IJNMT (International Journal of New Media Technology)*, vol. 8, no. 1, p. 65069, 2021.

[9] C. Gellner, I. Buchem and J. Muller, "Application of the Octalysis Framework to Gamification Designs for the Elderly," in *Proceedings of the 15th European Conference on Games-Based Learning*, Berlin, Germany, 2021.

[10] S. Duarte and M. Cruz, "The gamification octalysis framework within the primary english teaching process: The quest for a transformative classroom," Repositório Científico do Instituto Politécnico do Porto, 2018.

[11] A. K. Nalendra, "Rapid Application Development (RAD) model method for creating an agricultural irrigation system based on internet of things," in *IOP Conference Series: Materials Science and Engineering*, 2021.

[12] R. D. Mahande, "UTAUT Model: Suatu Pendekatan Evaluasi Penerimaan E-Learning pada Program Pascasarjana.," *INA-Rxiv*, p. 784–788, 2017.

[13] Kissflow, "Rapid Application Development (RAD) Model : An Ultimate Guide For App Developers in 2022," 2021. [Online]. Available: https://kissflow.com/low-code/rad/rapid-application-development/. [Accessed 18 May 2022].

[14] S. S. Alghazi, S. Y. Wong, A. Kamsin, E. Yadegaridehkordi and L. Shuib, "Towards Sustainable Mobile Learning: A Brief Review of the Factors Influencing Acceptance of the Use of Mobile Phones as Learning Tools," *Sustainability*, vol. 12, no. 24, p. 1–19, 2020.

[15] "Beyond TAM and UTAUT: Future directions for HIT implementation research," *Journal of biomedical informatics*, vol. 100, 2019.

[16] Meiryani, "MEMAHAMI SKALA LIKERT DALAM PENELITIAN ILMIAH," 2021. [Online]. Available: https://accounting.binus.ac.id/2021/08/13/memahami-skala-likert-dalam-penelitian-ilmiah/.

ISSN 2085-4552

# Methods of Stealing Personal Data on Android Using a Remote Administration Tool with Social Engineering Techniques

Ridwan Satrio Hadikusuma[1], Lukas[2], Epril M Rizaludin[3]
[1,2]Electrical Engineering Master Departement, Universitas Katolik Indonesia Atma Jaya, Jakarta, Indonesia
[3]PT. Klik Digital Sinergi, Jakarta, Indonesia
[1]ridwan.202200090017@student.atmajaya.ac.id, [2]lukas@atmajaya.ac.id, [3]epril_mohamadrizaludin@klik-ss.com

Accepted 24 March 2023
Approved 28 June 2023

***Abstract*— IT security is a significant concern of the internet because almost all communication occurs today. The purpose of testing personal data theft with the social engineering method is to ensure that the system and network on the user's Android have security holes to be hacked if the user is not aware of social engineering that allows data theft through the remote administration tool (RAT) which is accidentally downloaded on the Android User. Installing a RAT by applying social engineering is the possible and proper way to steal Android user privacy data. This study outlines some basic concepts of data theft, from recent call data and personal data to controlling Android users' cameras and microphones remotely.**

***Index Terms*— RAT; security; social engineering**.

## I. INTRODUCTION

One of the most significant inventions in human history that have changed how we do things is the Internet. The Internet has changed how we communicate, do business, and interact with each other[1-2]. Today, people no longer have to wait to hear from someone because they can communicate easily through the Internet. The Internet has also enabled businesses and individuals to receive payments instantly and track their fleets and cargo [3]. Security and information technology are increasingly important to society and the ICT(Information and Communication Technology) industry in this modern era. Security experts have developed various high-performance security tools to ensure that information on the Internet remains safe and not vulnerable to attack [4]. Various techniques, such as Layered Design, Assurance or Proof of Correctness, Software Engineering Environment, and Penetration Testing, test a complete, integrated, and reliable software, hardware, and people operational computer base[5, 6].

One example is using open-source frameworks such as Metasploit for exploit creation and penetration testing, which comes with over 1,600 exploits and 495 payloads to attack computer networks and systems. No matter how strong an android's security system is, it can still be penetrated if the user is still easy to manipulate, especially using social engineering methods. Social engineering is a manipulation technique that exploits human error to access private information or valuable data [7-9]. In the world of cybercrime, this type of human hacking scam can lure unsuspecting users. The most common is manipulating Android users to install an application (possibly under the guise of an e-ticket, package delivery receipt numbers for tracking, to other deceptive applications), where these applications are remote administration tools used for personal data theft [10][11]. Remote addressing tools android hacking is a technology that allows someone to remotely access an Android device without having to be near the device. Hackers use remote addressing tools to exploit vulnerabilities or loopholes in the Android security system to access the device and gain access to sensitive data such as photos, text messages, or financial information [12, 13].

Hackers use remote addressing tools (RAT) for malicious purposes such as data theft, extortion, or other criminal activities. However, this technology can also be used positively, for example, to help Android device owners who have forgotten their passwords or pins or to monitor people who need help, such as children or the elderly [14]. It is important to remember that using remote addressing tools for unethical purposes can compromise the privacy and security of one's Android device. Therefore, keeping your Android device safe is essential by installing the latest security apps, updating the operating system regularly, and avoiding downloading apps from sources you do not trust. This article will discuss some of the most frequently used remote addressing tools in Android hacking and how to protect your Android device from these threats.

## II. METHODOLOGY

Before the authors conduct research, the authors conduct several literature studies from several related studies that are still relevant to the research to be conducted. one of them is research conducted by

Huang, Y., & Han, X. [15] in his research entilted "Security Analysis of Remote Administration Tools for Android Devices". In this study, the researchers analyzed the security of six popular remote administration tools for Android devices, and found that these tools are vulnerable to social engineering attacks. this is also in line with research conducted by Iliyasu, A. M., & Ahmad, M. O [16] in his research related to "A Comprehensive Study on Android Remote Administration Tools: Threats, Vulnerabilities and Countermeasures". In this study, the researchers conducted a comprehensive analysis of Android remote administration tools, and identified various threats and vulnerabilities associated with these tools. The researchers also proposed a set of countermeasures to mitigate these risks.

Different from what Prakash, S., & Jadhav, S [17] did in an article entitled "Social Engineering Attacks in Android Platform. 2018 International Conference on Intelligent Computing and Control Systems (ICICCS)". In this study, the researchers analyzed the various social engineering attacks that can be used to exploit vulnerabilities in Android devices, and proposed a set of countermeasures to prevent such attacks. there is another study entitled "A Study on Remote Administration Tools and Their Impact on Android Devices" conducted by Ravikumar, N., & Gokulnath, C [18]. In this study, the researchers analyzed the impact of remote administration tools on Android devices, and identified various security risks associated with these tools. The researchers also proposed a set of countermeasures to mitigate these risks. Finally, what is interesting for the author's research in conducting this research is in research entitled "A Review on Security Threats and Countermeasures for Android Remote Administration Tools" conducted by Singh, G., & Kapoor, S [19]. In this study, the researchers conducted a review of the security threats associated with Android remote administration tools, and proposed a set of countermeasures to prevent these threats. The researchers also discussed the importance of user education in preventing social engineering attacks.

## III. Research Method

The research method on remote administration tools Android hacking using social engineering can be done in several stages. First, researchers must select and identify the types of remote administration tools used on Android devices. Then, the researcher must identify the security vulnerabilities in each remote administration tool. After that, researchers must conduct trials of each remote administration tool by carrying out attacks by exploiting the security holes found. In this case, researchers will use social engineering techniques to trick Android device users so that they can install remote administration tools unknowingly. After successfully installing the remote administration tools on the Android device, researchers will conduct testing and analysis of the data successfully retrieved from the Android device. Furthermore, researchers will evaluate the results of the tests and analyses carried out and provide recommendations regarding actions that need to be taken to improve the security of Android devices from attacks using remote administration tools and social engineering. Briefly, the flow of this research is described in Figure 1 below.

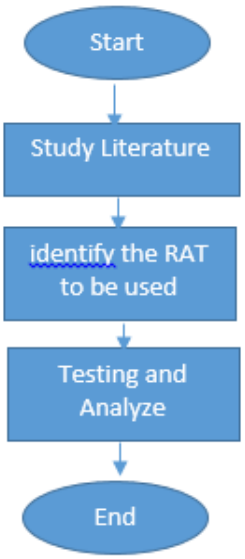


Fig. 1. Research Flowchart

### *A. Determination of Remote Administration Tools*

Installing remote administration tools on the Kali Linux operating system can be done by downloading the installation package according to the operating system and architecture used. After that, the installation package can be installed on Kali Linux using the appropriate installation commands. However, remember that using remote administration tools in hacking activities is illegal and can violate privacy and harm others. Therefore, remote administration tools must be used with good ethics and pay attention to security and privacy aspects. In addition, using Kali Linux must be done for good and legal purposes, such as conducting security tests on a system or network owned or with permission from the owner of the system or network.

Determining remote administration tools for Android hacking can be done by looking for references from various sources, such as underground forums, websites, or blogs about information security, or using a unique search engine such as Shodan. Then, it is necessary to evaluate the tools found in terms of functionality, capability, security, and the legality of their use. Data collection methods related to Android hacking remote administration tools can be done using scanning and enumeration techniques on the target system or network. This can be done using special tools like Nmap, hoping, or the Metasploit framework. In

addition, using tools such as Wireshark can also assist in collecting data related to network traffic that occurs when using remote administration tools on the target.

### *B. Social Engineering Techniques*

Social engineering methods of installing RAT applications can be carried out in various ways, such as creating fake messages or emails that look genuine, creating websites or pages that mimic the appearance of official sites, or using other tactics that trick the target into downloading and installing applications that contain malware. These techniques usually involve psychological manipulation of the target, such as making fraudulent offers or promises, intimidating or threatening, exploiting curiosity, and taking advantage of the target's trust or ignorance. It is important to remember that these actions are illegal and may harm others. Therefore, it is crucial for technology users always to be vigilant and careful when obtaining information or downloading applications from unknown sources. In addition, it is also essential to update the device security system and use the latest security software to avoid harmful malware attacks.

### *C. Methods of Data Collection and Prevention*

Collecting personal data using hacking remote administration tools (RAT) on Android devices can be done in various ways. One common way is sending malicious applications embedded with RATs to the target device. Once the application is installed on the target device, RATs can collect personal data such as text messages, phone calls, and browsing history. In addition, RATs can also be used to take control of the target device, such as activating the camera or microphone and recording user activity without their knowledge. Another method is to use phishing techniques, such as sending fake messages or emails that trick users into providing their personal information or clicking on links containing malicious applications embedded in RATs. In addition, users can also become victims of RATs attacks through unprotected Wi-Fi networks or applications that are vulnerable to attack, such as banking or e-commerce applications [7].

Therefore, it is essential for Android users always to be careful and avoid downloading apps from untrusted sources and using protected and trusted Wi-Fi networks. In addition, users are also advised to use security applications to monitor and protect their devices from RATs and other malware attacks. Android's fairly tight security system protects its users from remote administration tools (RAT) attacks and other malicious applications. One way to increase Android security is to activate security features provided by the operating system, such as a password or PIN, fingerprint sensor, or screen pattern lock settings [20]. In addition, users are also advised not to download applications from untrusted sources or use antivirus applications to identify and block malicious applications. Suppose you suspect that RATs or other malicious applications have infected your Android device. In that case, the first step is to remove the application from the device and perform a system scan using an antivirus. Also, it is recommended to continuously update the Android operating system to the latest version and avoid using unprotected or untrusted Wi-Fi networks. By paying attention to these security measures, Android users can minimize the risk of attacks by RATs and other malicious applications and protect their privacy and personal data.

TABLE I. ACCEPTANCE LEVEL CATEGORY INTERVAL

| Percentage Interval | Acceptance Level Category |
|---|---|
| 0% - 20% | Strongly Disagree |
| 20.01% - 40% | Disagree |
| 40.01% - 60% | Uncertain |
| 60.01% - 80% | Agree |
| 80.01% - 100% | Strongly Agree |

## IV. RESULTS AND DISCUSSION

### *A. Operating System Installation and Remote Administration Tools Identification Results*

Kali Linux is a distribution specifically designed for penetration testing, including in android hacking. Kali Linux has several advantages in its use that make it easier for practitioners to do hacking. One of the advantages of Kali Linux is that it is equipped with a variety of complete hacking tools. Hence, users no longer need to install additional tools manually. Apart from that, Kali Linux also has an intuitive user interface so that users can efficiently operate the system and the tools provided. Another advantage is modifying and customizing the tools according to user needs through manual configuration or built-in features such as meta-packages. Thus, Kali Linux is one of the right choices for android hacking practitioners in conducting security testing on the Android system. Here are the Kali Linux installation steps [11]:

1. Download the Kali Linux ISO file from the official Kali Linux website.
2. Prepare an empty USB flash drive with a minimum capacity of 4 GB.
3. Download and install the Rufus application to create a bootable USB. Open the Rufus application and select the USB flash drive to use.
4. In the "Boot selection" section, click the "SELECT" button and select the Kali Linux ISO file downloaded in step 1.
5. Ensure the USB flash drive partition is in "MBR" mode, and the file system is "FAT32".
6. Click the "START" button and wait until the bootable USB creation process is complete.

7. After the process, insert the USB flash drive into the computer where Kali Linux will be installed.
8. Set the BIOS settings to boot from the USB flash drive the first time. The method depends on the type and brand of your computer or laptop.
9. Select the "Graphical Install" option on the Kali Linux boot menu after successfully booting from the USB flash drive.
10. Follow the on-screen installation instructions, including selecting the language, time zone, and hard drive partition to use.
11. Select the root password setting and create a new user account.
12. Wait for the installation process to finish. Once done, Kali Linux is ready to use.

Installing Kali Linux requires basic knowledge of the operating system and BIOS settings. Therefore, ensure you understand the instructions and the associated risks before starting the installation process. Another alternative in Linux installation is to use a virtual machine (virtual server) to run the operating system.

After the operating system is installed, the next step is to determine the RAT that will be used. the author uses AhMyth as a RAT which will be used to commit personal data theft. AhMyth is a popular Android hacking tool and relatively easy to use on Kali Linux. Here are the steps to run AhMyth on Kali Linux [18]:

1. First, ensure that Kali Linux is installed and updated with the latest version.
2. Then, open the terminal on Kali Linux and run the command "git clone https://github.com/AhMyth/AhMyth-Android-RAT.git" to download the AhMyth source code from GitHub.
3. After successfully downloading, enter the AhMyth directory with the command "cd AhMyth-Android-RAT".
4. Next, run the command "sudo sh AhMyth.sh" to start the AhMyth installation process on Kali Linux.
5. Wait for the installation process to finish, and after that, open AhMyth by typing the command "sudo sh ahmyth".
6. After AhMyth opens, users can start creating an Android application that will be injected with the AhMyth payload. To do so, users can follow the instructions available on AhMyth.
7. Once done, AhMyth is ready to hack the target Android device.

However, remember that unauthorized use of AhMyth on devices not belonging to the user is illegal and can result in serious legal consequences. Therefore, AhMyth must be used ethically and comply with applicable regulations. As for AhMyth's appearance as shown in Figure 2 below.

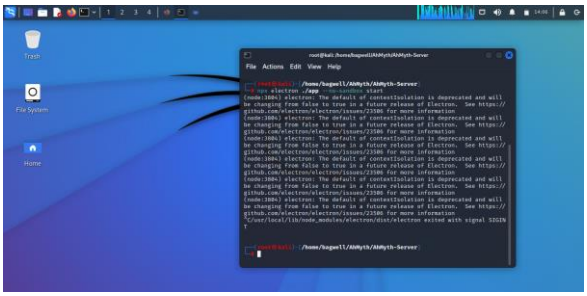

Fig. 2. Display the AhMyth RAT and run the server command

### *B. Results of Social Engineering*

After the RAT is deployed into an application, the next step is manipulating the target to install the RAT application on their Android device. Various ways can be done, such as sending e-tickets and package delivery receipts to government assistance programs, as shown in Figure 3 below.

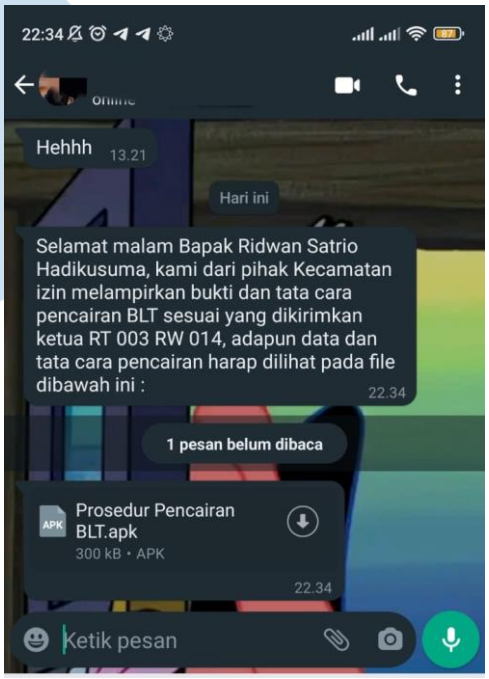


Fig. 3. Social Engineering Results

When doing social engineering to get someone to install Remote Administration Tools (RAT) on an Android device, it takes some persuasive skills to convince the victim. First, an aggressor needs to know about the victim's interests and wants. Furthermore, the attacker can create a scenario to interest the victim and make him interested in installing the RAT application. One trick often used is to promise an attractive offer, such as a free application or a premium service at a low cost. Attackers can also use phishing techniques by sending fake emails or text messages that look genuine and offer an attractive application or service. In addition, attackers can also take advantage of the

victim's fear or worry about the security of Android devices.

Attackers can promise better device security by installing a RAT application when in fact, the application is dangerous malware. To ensure success in getting someone to install the RAT application, the attacker needs to master effective and creative persuasive techniques and constantly update and improve these techniques so that they can always trick the victim. However, it is essential to remember that such actions are illegal and can cause harmful effects on others and, therefore, should not be carried out.

### *C. Results of Personal Data Theft*

After the RAT is installed on the target device, the author can find the country, device type, and IP used, as shown in Figure 4. Since the first installation, all user data on the Android device can be accessed and fully controlled by the author (also in this research, the target is part of the author and already with permission).

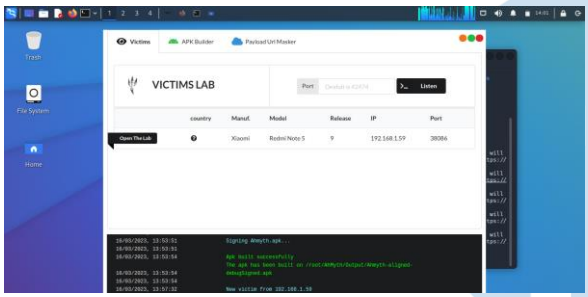


Fig. 4. Device information from the attacker side

The first data result is that the author steals or takes contact data stored on the target android device (see figure 5). the contact data can be used for various things, such as being sold to online gambling sites for marketing needs, fraud, and many other digital crimes.

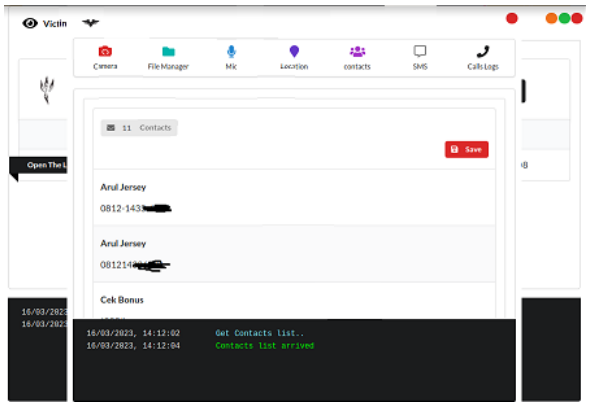


Fig. 5. Contact information is stored on the target device

The author also managed to hack the camera, microphone and real-time location of the target android device, which can be controlled 24 hours a day, as shown in Figure 6 and Figure 7.

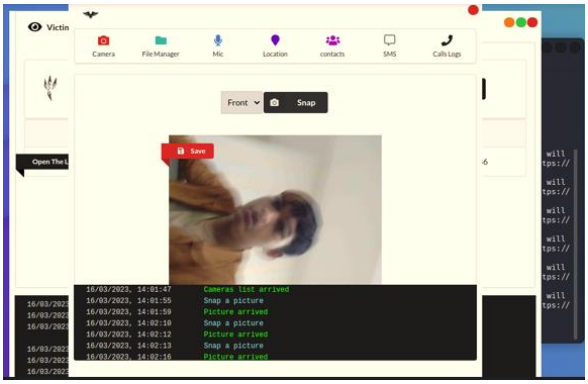


Fig. 6. The author view currently controlling the target android device's camera

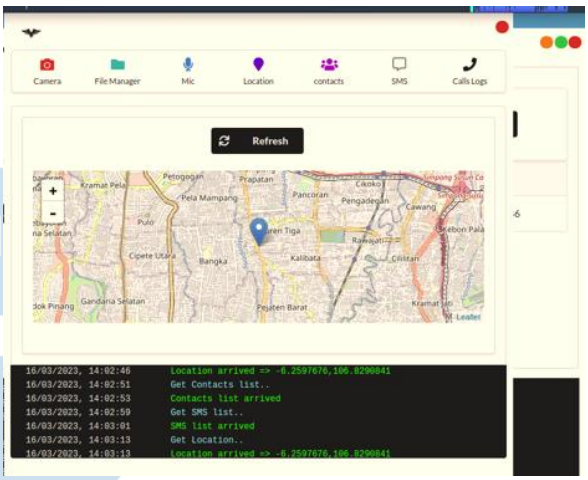


Fig. 7. Author display in monitoring the real-time location of the target device

Finally, the author can access all storage files from the target android device as a whole, starting from photo and external storage files to the android system files themselves (see fig. 8). Of course, this crucial data can be used for various digital crimes, the most dangerous of which is hacking an M-Banking account installed on a user's device.

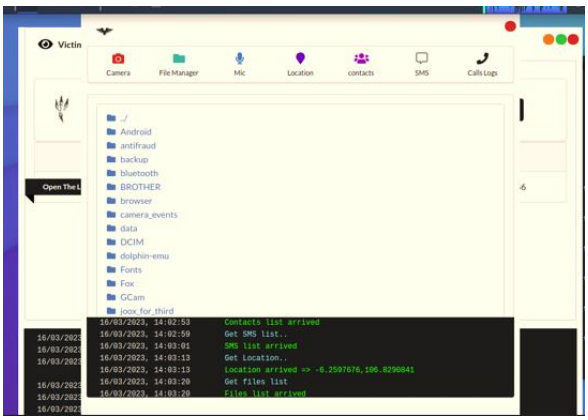


Fig. 8. Monitoring of target android device folder from author side

## V. CONCLUSION

Based on the research that has been done, the use of social engineering techniques in installing Remote Administration Tools on Android devices is very effective. In this study, researchers convinced respondents to install applications containing Remote Administration Tools by making convincing fake messages or phone calls. In addition, the research results also show that security on the Android system is still very vulnerable to Remote Administration Tools attacks that can take over the device remotely and collect users' data without their knowledge.

Therefore, it is necessary to take better precautions and safeguards on Android devices to prevent malicious Remote Administration Tools attacks. Some steps to avoid these attacks include downloading apps only from trusted sources, keeping your Android device updated with the latest security patches, and installing reliable antivirus software. Additionally, awareness and education are also needed for Android users to recognize and prevent Remote Administration Tools attacks that use social engineering techniques.

## REFERENCES


[1] A. R. Maulana and D. P. Wardhana, "Remote Administration Tool (RAT) Implementation using AhMyth and Social Engineering Techniques," 2020 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM), Surakarta, Indonesia, 2020, pp. 1-5, doi: 10.1109/CENIM51083.2020.9317252.

[2] W. R. Pratama and A. F. T. Riyadi, "Uji Penetrasi Remote Administration Tool pada Android dengan Teknik Social Engineering," Jurnal Teknologi Informasi dan Komunikasi, vol. 8, no. 2, pp. 56-62, 2022.

[3] R. H. Putra, M. N. Huda and T. A. Wisesa, "Android Hacking Using AhMyth RAT with Social Engineering Techniques," 2020 4th International Conference on Informatics and Computing (ICIC), Jakarta, Indonesia, 2020, pp. 1-5, doi: 10.1109/IAC50653.2020.9259369.

[4] S. Pradana, D. D. Setiawan and N. E. Darmawan, "Remote Administration Tool (RAT) Implementation using AhMyth RAT and Social Engineering Techniques," 2021 International Conference on Advanced Informatics: Concept, Theory and Application (ICAICTA), Malang, Indonesia, 2021, pp. 1-6, doi: 10.1109/ICAICTA51487.2021.9488906.

[5] S. R. S. Maharjan, S. Maharjan and S. Adhikari, "Social Engineering Techniques and the Use of Remote Access Trojans (RATs) in Android Devices," 2019 4th International Conference on Computing, Communication and Security (ICCCS), Rome, Italy, 2019, pp. 1-6, doi: 10.1109/CCCS.2019.8887181.

[6] R. Agarwal, S. Saha, S. Chaki, "Security Issues and Threats in Android Platforms: A Survey", International Journal of Computer Applications, vol. 52, no. 6, pp. 28-35, 2012.

[7] B. Al-Duwairi, "A Study on Security Issues of Mobile Devices and Applications", International Journal of Advanced Computer Science and Applications, vol. 8, no. 9, pp. 239-244, 2017.

[8] S. Arora, A. Singh, "Mobile Malware Detection: A Review", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 7, no. 5, pp. 76-79, 2017.

[9] A. Babar, M. Masood, M. Farooq, "An Analysis of Mobile Malware: A Comprehensive Study", International Journal of Computer Science and Network Security, vol. 15, no. 3, pp. 103-112, 2015.

[10] S. Bhattacharya, S. Kumar, "A Study on Android Malware Detection and Prevention Techniques", International Journal of Computer Applications, vol. 168, no. 5, pp. 8-12, 2017.

[11] K. Chandrakar, S. Bhoi, "Mobile Security: Issues, Challenges and Future Directions", International Journal of Advanced Research in Computer Science, vol. 8, no. 3, pp. 118-123, 2017.

[12] A. Dhiman, S. Sharma, "Security Issues and Solutions for Mobile Devices", International Journal of Computer Applications, vol. 104, no. 12, pp. 21-26, 2014.

[13] N. M. N. Lestari and A. W. Nugroho, "Penerapan Teknik Social Engineering dalam Remote Administration Tools pada Android," Jurnal Keamanan Informasi, vol. 5, no. 1, pp. 12-21, 2021.

[14] D. F. Maulana, R. A. Hidayat, and R. D. Saputra, "Analisis Penggunaan Remote Administration Tool dengan Teknik Social Engineering pada Android," Jurnal Ilmiah Informatika, vol. 10, no. 2, pp. 123-130, 2019.

[15] Huang, Y., & Han, X. (2018). Security Analysis of Remote Administration Tools for Android Devices. 2018 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA).

[16] Iliyasu, A. M., & Ahmad, M. O. (2021). A Comprehensive Study on Android Remote Administration Tools: Threats, Vulnerabilities and Countermeasures. Journal of Information Security, 12(2), 79-97.

[17] Prakash, S., & Jadhav, S. (2018). Social Engineering Attacks in Android Platform. 2018 International Conference on Intelligent Computing and Control Systems (ICICCS).

[18] Ravikumar, N., & Gokulnath, C. (2016). A Study on Remote Administration Tools and Their Impact on Android Devices. International Journal of Advanced Research in Computer Science and Software Engineering, 6(8), 262-268.

[19] Singh, G., & Kapoor, S. (2018). A Review on Security Threats and Countermeasures for Android Remote Administration Tools. International Journal of Advanced Engineering Research and Science, 5(11), 122-126.

[20] N. Hidayatullah, I. A. Akbar, and R. Kurniawan, "Social Engineering-Based Attack on Android Mobile Device," in 2019 International Conference on Information Management and Technology (ICIMTech), 2019, pp. 82-87.

ISSN 2085-4552

# Development of A Mobile-Based Student Grade Processing Application Using the Waterfall Method

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

Accepted 03 April 2023
Approved 07 July 2023

***Abstract*— One of the public elementary schools in Medan is SDN 064017. At SDN 064017, the procedure for processing student grade data has not been computerized. Gradebooks are used to track and process student grade information. There are still certain problems that arise with this method. SDN 064017 requires a lot of time and effort to process and find information on student scores. The aim of the research is to develop a mobile-based application that will simplify and speed up SDN 064017's processing and searching for student grade data. The system development method used is the waterfall method, which consists of the stages of needs analysis, system design, implementation, testing, deployment, and maintenance, while the system testing method uses the black box testing method. In the process of developing this mobile-based application, it uses the Kodular framework and the Airtable database. The results of the research conducted are an information system that can help SDN 064017 in processing and searching for student grade data to be more effective and efficient through mobile Android.**

***Index Terms*— Android; Grades; Information System, Students; Waterfall**.

## I. INTRODUCTION

Modern technology is developing rapidly, especially in the field of information technology, which includes mobile technology and the internet [1]. In addition, Android technology, one of the mobile technologies that is currently being developed, On the other hand, because Android is an open-source platform, developers are free to build new applications or modify existing ones [2]. Modernization in schools increases efficiency and makes processes simpler.

One of Medan City's public elementary schools is SDN 064017. The data processing of student scores at SDN 064017 is still manual or not yet digital. The process of recording and processing student grade data is carried out in a grade book. In this way, there are still some problems that occur. SDN 064017 requires a lot of time and effort to record and process student grade information. In addition, finding student-grade information becomes a time-consuming and ineffective task.

The waterfall method, which includes the stages of needs analysis, system design, development, testing, implementation, and maintenance, is used to manually change the student grade data processing system into an Android mobile-based application. Thus, SDN 064017 can process and search for student grade data through an Android mobile device with the help of an application that is being developed using the Kodular framework and the Airtable database.

In 2021, Muhammad Afifudin and Agung conducted similar research. This research examines the processing of grade data, which is still being carried out using a spreadsheet application. As a result, there is still the problem of data hoarding, which has an impact on the difficulty of obtaining information. By using the PHP programming language, the CodeIgniter framework, and the MySQL database, this research focuses on creating a web-based information system that can be used to process valuable data quickly [3].

Similar research was conducted by Dedi Irawan and I Putu Arya Aryanto in 2020. Processing student grades using Microsoft Excel, which is discussed in this study, is less effective and efficient, so this study aims to develop a web-based student grade data processing system using the PHP programming language [4].

In 2020, Handini Widyastuti and Rizat Sakmir conducted a similar study. This study looked at the difficulties associated with manually processing student grade data, which is still stored in leger books and report cards and requires a lot of time and work to complete. To overcome these problems, a study was conducted that aimed to create a web-based information system for processing student grades [5].

The research conducted differs from previous studies in several important respects. By using the Kodular framework and the Airtable database, the purpose of this research is to change the manual student grade data processing system into an Android mobile-based student grade data processing application.

## II. METHODOLOGY

### A. System Development Method

The method used in system development using the waterfall approach is preferred because the steps are completed sequentially and continuously. For projects where tool quality is important, the waterfall methodology is highly recommended but requires significant documentation [6]. The stages of the waterfall method can be seen in Figure 1.

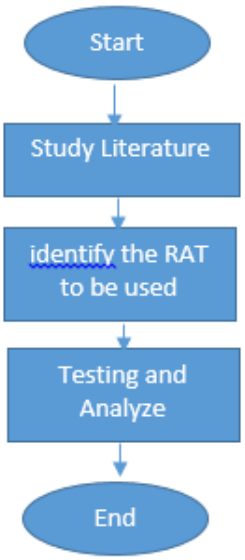


Fig. 1. Research Flowchart

The following is an explanation of each stage contained in the waterfall method [7].

1. Requirements

This research activity begins with analysis. The analysis carried out includes an analysis of future system requirements as well as problems that arise in the current student grade data processing system.

2. Design

The design of Android mobile applications that will be developed using UML (Unified Modeling Language) tools consisting of use case diagrams, sequence diagrams, activity diagrams, and class diagrams

3. Implementation

The Kodular Framework and Airtable database are used to create Android-based mobile applications, translating the design results into a language that computers can understand.

4. Verification

At this stage, a black box testing approach is used to test the applications that have been developed to evaluate the efficiency and precision of each process they contain.

5. Maintenance

Maintenance is carried out on applications that have been built to ensure the application functions properly and no errors occur when used by the user.

### B. Method of Collecting Data

This study uses data collection techniques or procedures in addition to system development methodology. Without knowledge of data collection techniques, it is certain that researchers will not be able to collect data that meets predetermined standards. Data collection techniques are the main techniques in a study, starting with obtaining data, which is the main goal of research [8].

The following methods were used to collect data for this study:

1. Observation

To record, process, and find data on student scores at SDN 064017, direct observations were made at the research location, looking for problems that might arise during the procedure.

2. Interview

Direct questions and responses from SDN 064017 teachers and staff members about the procedure for recording and processing grade data as well as any issues that came up during the process served as the basis for the data collection process.

3. Literature review

For the theories needed for research activities related to the creation of Android-based mobile applications, data collection was carried out by searching books, journals, the internet, and other sources.

## III. RESULTS AND DISCUSSION

### A. Problem Analysis

The first stage is to analyze the running system before starting the design stage. The analysis is centered on student assessment procedures at SDN 064017 Medan. Student grades are written in a grade book as part of the assessment process, which is still done manually or not digitally. The low level of protection for student grade data, lost or damaged student grade books, as well as the time, effort, and costs required to process and search for student grade data at SDN 064017 Medan, are some of the remaining problems.

### B. System Requirements Analysis

Based on the findings of the system analysis, a process of identifying several problems was carried out, especially those related to the time, cost, and effort required to process and search for student grade data. The research aims to develop an Android-based student grade data processing application to process and search for student grade data at SDN 064017 Medan. The Kodular framework and Airtable database, which can be accessed online via an Android phone, are used to develop the application.

### C. Design System

By using use case diagrams, sequence diagrams, activity diagrams, and class diagrams, researchers

establish the general architecture and design details of the applications that will be made at this stage.

1. Use Case Diagram

Use case diagrams show how a system should operate, while use cases show how actors and systems interact. A person or actor is a human or technological entity that communicates with a system to perform a certain task [9].

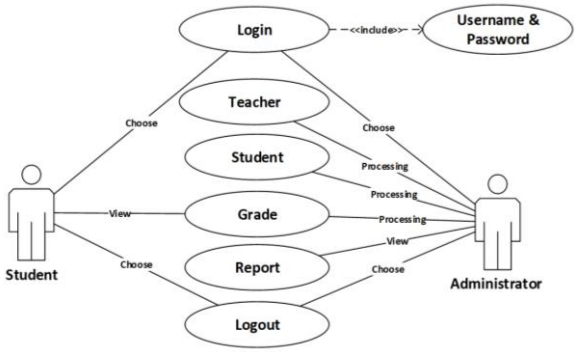


Fig. 2. Use Case Diagram

Figure 2 illustrates the use case diagram of the application to be built. Use case diagrams to describe what actors can do in the application to be built. Android-based student grade processing applications can be accessed by administrators and students. Administrators can log into the application by entering the specified username and password. Administrators can process teacher data and student data. Administrators can process student grade data processing and can view student grade data reports that have been processed. Students can see the grades that have been processed by the administrator, after logging in to the application by entering the specified username and password.

2. Sequence Diagram

Scenarios that are performed in response to an event to create a particular output are described using sequence diagrams. What causes the action, what internal processes and changes occur, and what results are all considered first [10].

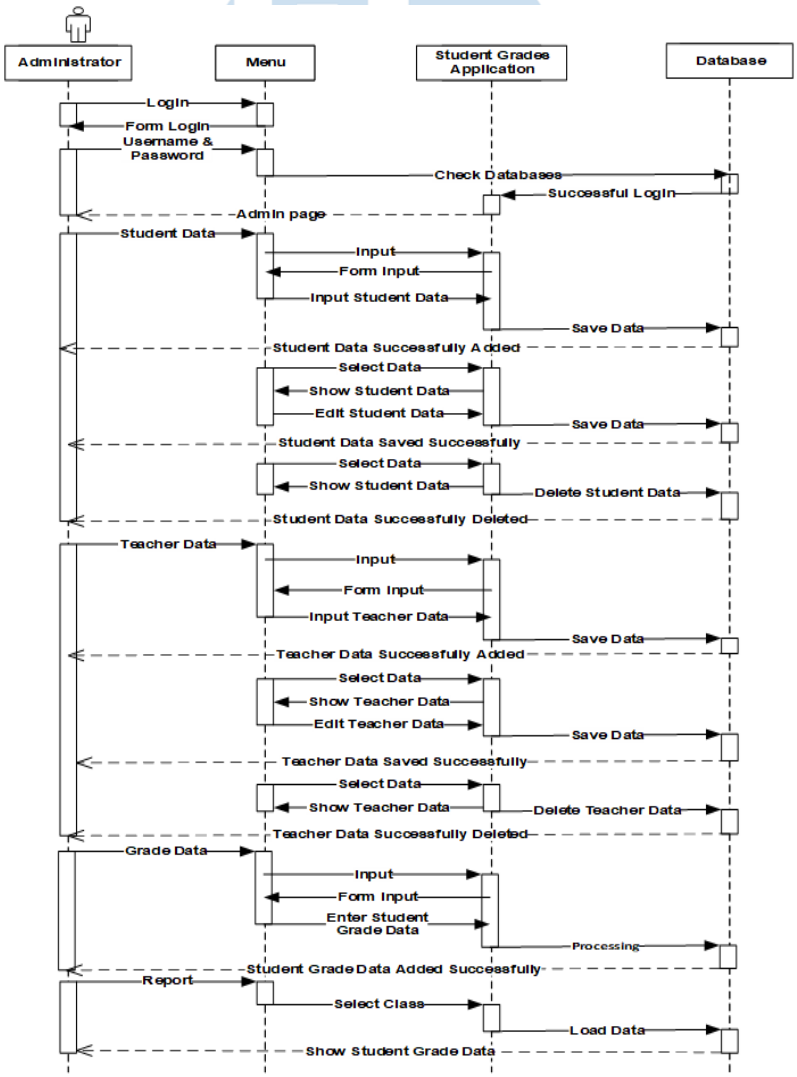


Fig. 3. Sequence Diagram

Figure 3 illustrates the sequence diagram of the application to be built. The sequence diagram of the Android-based student grade processing application that will be built begins when the administrator logs into the application. When the administrator enters the username and password used in the login process, the system will check the username and password data in the database. Then the administrator will enter the administrator page. When the administrator selects the student menu or teacher menu, the administrator can process student data or teacher data consisting of input, editing, and deleting. In the input student data or teacher data options, the administrator can fill in student data or teacher data on the form provided by the application and save the data into the database, in the edit option, the administrator can select student data or teacher data and make changes to the selected data and save the results of the changes in the database, in the delete option, the administrator can select student data or teacher data to be deleted and delete the data in the database. On the grade menu, the admin can process student grade data by selecting student name, teacher name, and subject, entering the number of student scores, and then storing student grade data in the database. On the report menu, administrators can see a summary of student grade data reports that have been processed.

3. Activity Diagram

The amount of activity flowing through the system, the start of each activity, the potential fork, and the conclusion of the program are all described using activity diagrams [11].

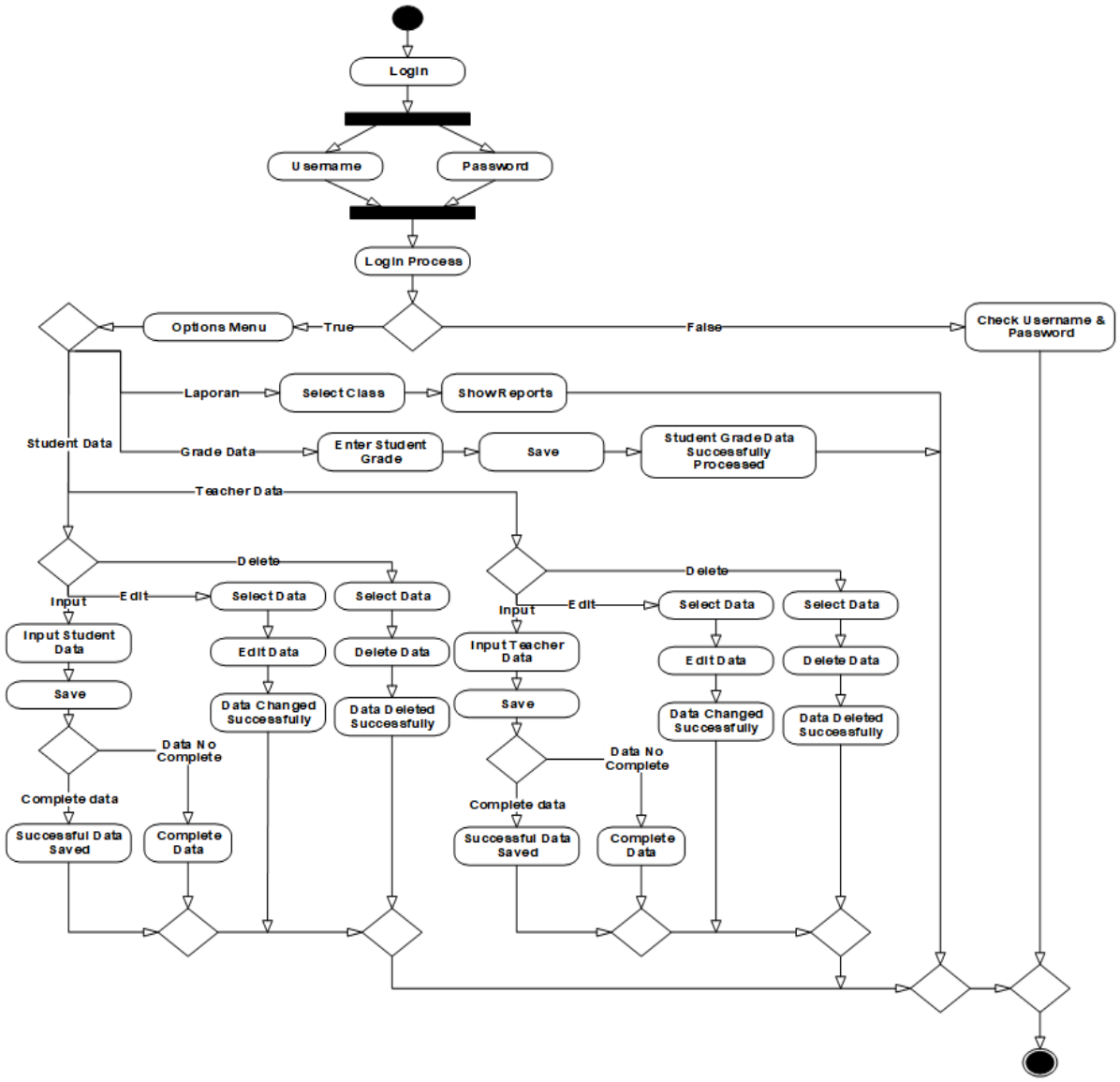


Fig. 4. Activity Diagram

Figure 4 illustrates an activity diagram. The activity diagram of the student grade processing application that will be built begins when the administrator logs in to the application by entering the specified username and password. Then the application verifies the username and password entered by the administrator, if the

username and password are correct, then the administrator can enter the options menu page and select an available menu. If the username and password are incorrect, the administrator cannot enter the application and must re-check the username and password entered. The application provides several menu options that can be selected by the administrator. When the administrator selects the student menu or teacher menu, the application will provide input options and edit and delete student or teacher data. In the input options, the application checks the completeness of the student data or teacher data entered by the admin, if the data is complete, then the data can be processed and stored in the database, but if the student data or teacher data is incomplete, then the data cannot be processed and the admin must complete the data first. In the edit and delete options, the application will save the results of changes to student or teacher data made by the admin. When the admin selects the grade menu, the admin chooses the class, student name, teacher name, and subject, then enters the student grades and saves the grade data that has been entered by the admin. On the report menu, the application will display the results of the grade data recapitulation that has been processed by the admin.

4. Class Diagram

The system structure is depicted in a class diagram in the form of classes that will be developed to build the system. The features of a class include attributes and methods or activities [12]. The class diagram is the foundation of object-oriented development and design because it is a specification that, when instantiated, creates an object [13].

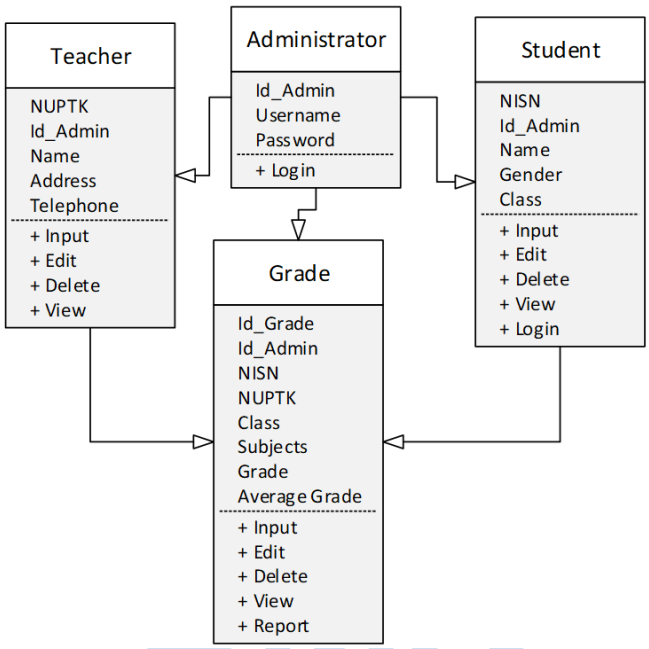


Fig. 5. Class Diagram

Figure 5 illustrates the relationship between each table in the student grade processing application at SDN 064017. The admin table structure consists of id_admin, username, and password, the student table structure consists of NISN, id_admin, name, gender, and class, the teacher table structure consists of NUPTK, id_admin, name, address, and telephone, the grade table structure consists of id_grade, id_admin, NISN, NUPTK, and class. subjects., grade and average grade. The class admin can process teacher data in the teacher class, student data in the student class, and grades in the grade class. Class teachers can enter grade data according to their respective subjects in class grades. Class students can see the grade data that has been processed by the teacher in the teacher class or admin in the admin class.

*D. System Implementation*

The process of translating the design based on test results into a specific programming language and deploying a system that functions in the actual environment is known as the system implementation step. The following is a display of the student grade processing application. The appearance of the application consists of displaying the login page, student data page, teacher data page, grade data page, and report page.

1. Login Page

The administrator must login on the login page using the username and password before he can access the student grades application. The login page can be shown in the Figure 6.


Fig. 6. Login Page

2. Student Data Page

Administrators can process student data consisting of input, edit, delete, and view student data on this page. The student data page can be shown in the Figure 7.


Fig. 7. Student Data Page

3. Teacher Data Page

Administrators can process teacher data consisting of input, edit, delete, and view student data on this page. The teacher data page can be shown in the Figure 8.


Fig. 8. Teacher Data Page

4. Grade Data Page

Administrators can process student grade data by selecting a class, student name, teacher name, and subject name, and entering student grades on this page. The grade data page can be shown in the Figure 9.


Fig. 9. Grade Data Page

5. Report Page

Administrators can view student grade reports that are processed by class on this page. The report page can be shown in the Figure 10.


Fig. 10. Report Page

*E. System Testing*

Application testing is carried out to check the functionality of the application being developed and look for deficiencies or errors. The black box method is used in testing this research. Black box testing is a type of testing based on application specifications [14] comprising the user interface or external look of the program, process performance, input, processing, and output generated by the application, and how to use application features in the user environment, followed by validation testing [15]. The results of blackbox testing can be seen in the Table I.

TABLE I. SYSTEM TEST TABLE

| Testing Module | Testing Procedures | Input | Output | Conclusion |
|---|---|---|---|---|
| Login Admin | - Open Application<br>- Entry the username "sirka", password "admin"<br>- Click Login | Username “sirka”, password “admin” | Admin can enter into the application select the available menu | Valid |
| Student Data | - Open Application<br>- Login<br>- Select the student menu<br>- Entry complete student data<br>- Click input | Complete student data | Student data successfully added | Valid |
| Teacher Data | - Open Application<br>- Login<br>- Select the teacher menu<br>- Entry complete teacher data<br>- Click input | Complete teacher data | Teacher data successfully added | Valid |
| Grade Data | - Open Application<br>- Login<br>- Select the grade menu<br>- Select class, student name, teacher name and subjects<br>- Entry student grade data<br>- Click input | Complete grade data | Data student grades successfully processed | Valid |
| Report | - Open Application<br>- Login<br>- Select the report<br>- Select class | Class Data | View student grade reports | Valid |

In addition, testing the application on an Android smartphone The application is tested in terms of appearance, process, and performance when installed on different Android smartphone devices. The classification of Android devices determined in this test consists of the operating system and Android smartphone brand used by the user, which consists of ten devices: two Android devices version 8, two Android devices version 9, two Android devices version 10, two Android devices version 11, and two Android devices version 12. The results of application implementation testing can be seen in Table II.

TABLE II. TABLE OF TESTING THE APPLICATION ON ANDROID DEVICES

| Devices | Android Version | Information System Display | Information System Process | Information System Performance |
|---|---|---|---|---|
| Infinix Hot 6 Pro X608 | Version 8 | Success | Success | Success |
| Xiaomi Redmi 6A | Version 8 | Success | Success | Success |
| Samsung Galaxy A10 | Version 9 | Success | Success | Success |
| Realme 3 Pro | Version 9 | Success | Success | Success |
| Xiaomi Redmi Note 8 | Version 10 | Success | Success | Success |
| Xiaomi Mi 8 Pro | Version 10 | Success | Success | Success |
| Xiaomi Redmi 10 | Version 11 | Success | Success | Success |
| Vivo V20 | Version 11 | Success | Success | Success |
| Redmi 10 Prime | Version 12 | Success | Success | Success |
| Vivo V60 Pro | Version 12 | Success | Success | Success |

After testing is done, the next step is to make a questionnaire to get evaluation results from the use of the application that has been built. The questionnaire was given to the teachers and students of SDN 064017. The questionnaire instrument was related to the ease of use of the application, the appearance of the application, and the performance of the application in helping to facilitate the process of processing student grades at SDN 064017. The results of the questionnaire can be seen in Figure 11, Figure 12 and Figure 13.

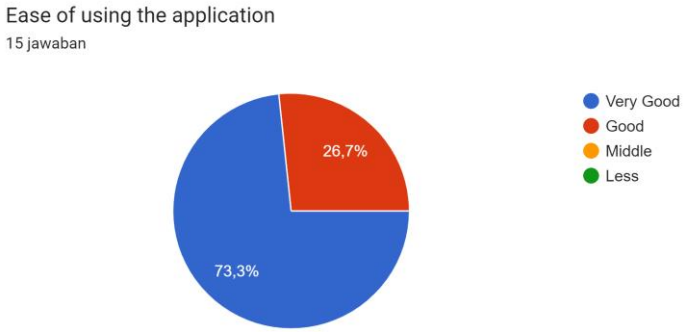


Fig. 11. Application Usage Survey Results

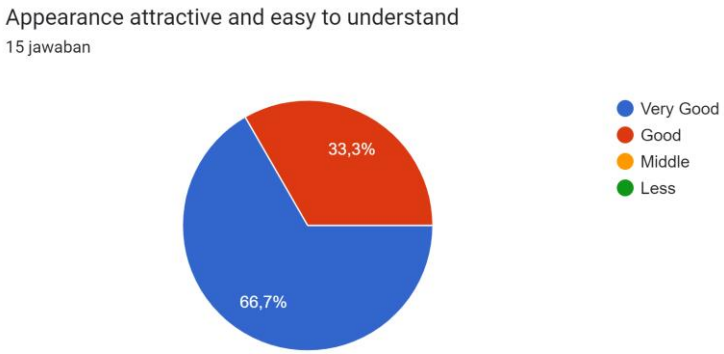


Fig. 12. Application Design Survey Results

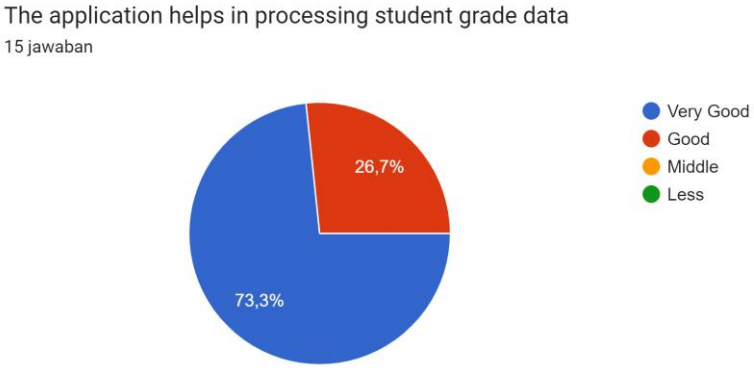


Fig. 13. Application Benefits Survey Results

## IV. Conclusion

Conclusions that can be drawn regarding the development of student grade data processing applications at SDN 064017 based on mobile Android that have been built after conducting this research are based on the stages contained in the waterfall method. A study of its utilization produces applications that function properly according to user needs. The student grade processing application that was built can replace the manual grade data processing system at SDN 064017 Medan City with an information technology-based grade data processing system using Android mobile devices. Based on the results of a questionnaire of the fifteen respondents who filled out the questionnaire form, more than 73.3% answered very well for the ease of use of the application, and 26.7% answered well for the ease of use of the application. 66.7% of respondents answered very well regarding the appearance of the application, and 33.3% of respondents answered well regarding the appearance of the application. 73.3% of respondents answered that the benefits of the application were very good in facilitating the processing of student grade data, and 26.7% of respondents answered that the benefits of the application were good in facilitating the processing of student grade data. So it can be concluded that 71.1% of respondents answered very well about the application that was built in terms of ease of use, appearance, and benefits to facilitate the processing of

student grade data. In addition, conclusions that can be made from the processing of student grade data before and after using the application can be seen in the Table III.

TABLE III. BEFORE AND AFTER USING THE APPLICATION

| Indicator | Before using the app | After using the app |
|---|---|---|
| Grade data processing time | The processing time for student grade data is longer because it is done manually by writing and counting one by one | Processing time for student grade data is faster because it is done automatically by the application |
| Search student grade information | The process of searching for student value data takes longer because it is done manually by searching data one by one and sequentially | The process of searching for student value data is faster because the application is equipped with a search feature so that it can perform searches quickly based on the keywords entered |
| Making student grade reports | Making student grade reports takes longer because it is done manually by creating and writing one by one from the processed student grades | The application provides a report feature, so student grade reports are available |
| Distribution of student grade reports | Student grade reports that have been made are then printed and distributed or delivered directly to those who need them, such as the school principal. | The application can be accessed by various types of users, so that those who need student grade data can directly obtain it through the application that has been accessed |

## REFERENCES


[1] J. R. Yuangga and R. Agustina, "Rancang Bangun Sistem Monitoring Akademis Siswa Berbasis Android Di Sd Negeri 1 Tlogosari Kecamatan Tirtoyudo Kabupaten Malang," *Rainstek J. Terap. Sains dan Teknol.*, vol. 3, no. 1, pp. 34–41, 2021, doi: 10.21067/jtst.v3i1.4160.

[2] M. Aman, "Aplikasi Informasi Nilai Siswa SMA Sukatani Dengan Metode Cload Test Berbasis Android," *J. PSIKOM*, vol. 7, no. 2, pp. 1–5, 2019.

[3] M. Afifudin and A. Riyantomo, "Sistem Informasi Pengolahan Data Nilai Siswa Berbasis Web (Studi Kasus Mi Darussalam Tlogoboyo)," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 3, no. 2, p. 125, 2021, doi: 10.36499/jinrpl.v3i2.4605.

[4] D. Irawan and I. P. A. Aryanto, "Pengolaha Data Nilai Siswa Pada Smp Negeri 7 Kota Metro Berbasis Web," *JIKI (Jurnal llmu Komput. lnformatika)*, vol. 1, no. 2, pp. 70–79, 2020, doi: 10.24127/jiki.v1i2.675.

[5] H. Widyastuti and R. Sakmir, "Perancangan Sistem Informasi Pengolahan Data Nilai Siswa Berbasis Web," *IMTechno J. Ind. Manag. Technol.*, vol. 1, no. 1, pp. 19–26, 2020, doi: 10.31294/imtechno.v1i1.65.

[6] M. Usnaini, V. Yasin, and A. Z. Sianipar, "Perancangan sistem informasi inventarisasi aset berbasis web menggunakan metode waterfall," *J. Manajamen Inform. Jayakarta*, vol. 1, no. 1, p. 36, 2021, doi: 10.52362/jmijayakarta.v1i1.415.

[7] S. Khadafi, A. Salim, R. Prabowo, and C. Anam, "Rancang Bangun Website UKM Reviora Tanggulangin Sidoarjo Menggunakan Metode Waterfall Sebagai Media Pemasaran Online," *Semin. Nas. Sains dan Teknol. Terap. VII - Inst. Teknol. Adhi Tama Surabaya* , pp. 705–710, 2019.

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

[9] Y. F. Apriyanda and M. B. Sanjaya, "Sistem Informasi Penyedia Info Bandara dan Wisata Berbasis Website dan Web Servis," *eProceedings …*, vol. 5, no. 3, pp. 2343–2358, 2019, [Online]. Available: https://openlibrarypublications.telkomuniversity.ac.id/index.php/appliedscience/article/view/11406.

[10] A. Arman, S. Sotar, and K. Ulya, "Sistem Informasi Pencatatan Hafalan Al-Qur'an Pada Smp Islam Terpadu Mutiara Kota Pariaman Berbasis Website," *Rang Tek. J.*, vol. 4, no. 2, pp. 316–324, 2021, doi: 10.31869/rtj.v4i2.2612.

[11] L. Sembiring and B. Sinuraya, "*Perancangan Sistem Informasi Persetujuan Bongkar / Muat Barang Khusus Berbasis Web Dengan Menggunakan Metode UML* (*Studi Kasus : Kantor Kesyahbandaran Utama Belawan*)," *CTIS*, vol. 3, no. 1, pp. 1–12, 2019.

[12] V. Kristianingrum and M. F. Y. Al-Fadillah, "Perancangan Website E-Commerce Penjualan Ikan Cupang," *JBMI (Jurnal Bisnis, Manajemen, dan Inform.*, vol. 18, no. 3, pp. 164–180, 2022, doi: 10.26487/jbmi.v18i3.19538.

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

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

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

ISSN 2085-4552

# Analysis of User-generated Content in Visitor Reviews of Tourist Attractions Using Semantic Similarity

Ni Made Satvika Iswari[1], I Gede Juliana Eka Putra[2]
[1,2]Informatics Study Program, Universitas Primakara, Denpasar, Indonesia
[1]iswari@primakara.ac.id, [2]gedejep@primakara.ac.id

Accepted 24 March 2023
Approved 28 June 2023

***Abstract*— The tourism industry plays an important role in the world economic sector because it makes a significant contribution to the global economy, creates jobs, and strengthens economic growth. Therefore, visitor satisfaction is very important in this industry. Currently tourists use information in online media to find tourist attractions that suit their needs and expectations. Reviews of tourist attractions are growing on the internet. Customers can post reviews, recommendations, or ratings of a tourist spot. Online reviews in the form of User-generated Content (UGC) can provide benefits for business managers to obtain feedback from customers and improve certain product attributes or service characteristics to increase business value and support marketing activities. In this research, a user-generated content analysis method has been produced in Visitor Reviews of Tourism Objects Using Semantic Similarity. The results of the sentiment analysis can be seen that positive, negative, and neutral sentiments have similar percentages for each category. This is because the designed algorithm cannot cut sentences for only one category that is assessed. The algorithm designed in this study has limitations, that is only able to analyze up to the semantic similarity stage but has not been able to cut which sentences are relevant to the category to be analyzed for sentiment.**

***Index Terms*— semantic similarity; sentiment analysis; tourist attraction; User-generated Content (UGC); visitor review**.

## I. INTRODUCTION

The tourism industry plays an important role in the world economic sector. In 2018, global tourism increased by 6% based on the United Nations World Tourism Organization (UNWTO) and is expected to increase by 3% to 4% in 2019. According to data from the World Tourism Organization (UNWTO), around 1.4 billion people travel international travel in 2018. The tourism industry makes a significant contribution to the global economy, creates jobs, and strengthens economic growth. Therefore, visitor satisfaction is very important in this industry[1].

However, the worldwide spread of Covid-19 in 2020 caused the tourism industry to face a catastrophic crisis as migration and travel became restricted. The ongoing pandemic has had an impact on tourist behavior. Micro, Small, and Medium Enterprise (MSME)s are most national tourism service providers. However, MSMEs lack the infrastructure and resources to provide products and services to more tourists. Estimated tourist arrivals can help MSMEs to significantly reduce the risk of not achieving sales targets, and even allow MSMEs to start building personalized tour packages based on the interests and needs of tourists [1].

Currently, tourists use information in online media to find tourist attractions that suit their needs and expectations. Reviews of tourist attractions are growing on the internet. Customers can post reviews, recommendations, or ratings of a tourist spot. Online review in the form of user-generated content (UGC) give business management advantages by enabling them to get client feedback and enhance features of products or services to boost brand value and help marketing initiatives [2].

Several related studies have provided significant findings about feedback in tourism. Most of these studies use data collection methodologies such as customer surveys, interviews, or Focus Group Discussion (FGD)s to analyze the factors influencing tourist behavior. Although this research provides many perspectives on tourist behavior, the Big Data Analytics approach can provide significant insights [3]. In this study, data collection was carried out from one of the UGC tourism data sources, namely visitor reviews on Google Maps. Data analysis was performed using the semantic similarity processing method. Analysis of the data will produce comprehensive information about tourist's feedback.

## II. LITERATURE REVIEW

### *A. Related Study*

Data is a vital asset to support business progress, including in the tourism sector. Information is necessary, but analytical outcomes are as important, particularly with big data technology. Big Data is used by businesses and organizations to assist in decision-

making. As a result, the analysis must be correct. The wrong choice will be made because of incorrect analysis, which will eventually be disastrous. Due to feature learning, Deep Learning algorithms—which are utilized in massive data analysis—can now make judgments that are comparable to those of humans. Big data analysis can improve intelligent computing by helping it to analyze unstructured real-world data[4].

Compared to conventional tourist survey data, a larger volume of tourist data is available online. The data allows decision makers to understand tourist behavior in a more granular way. Social media and social networking websites are an online platform that allows users to share experiences and opinions on a product or service [5]. Currently, tourists use information in online media to find tourist attractions that suit their needs and expectations. Given the large amount of unstructured content from social media, big data analytics approach like sentiment analysis, geographic visualization, and frequency analysis have been widely applied for User Generated Content (UGC) analysis in several research domains. UGC content offers the opinions and sentiment analysis of numerous individuals in diverse areas[6]. UGC is currently growing because it provides a valuable source of data for many parties to extract information that can be used as a competitive advantage[7].

Sentiment analysis-based neural networks are utilized to assist the information architecture of the tourism industry [1]. Big data analysis and content analysis are utilized in the restaurant industry to examine user-generated content (UGC) relevant to the meal experience for diners who have food allergies [8]. Supervised machine learning approach used on big data analytics cloud-based to support customer insight-based design innovation for the SME domain [9]. In this study, the UGC analysis was carried out using the natural language processing approach. The data analysis will produce comprehensive information about customer satisfaction.

### *B. Semantic Similarity*

In computer science, semantics is the mathematical justification for a valid string that is specified by a programming language. The rules regulating program, or syntax, are the exact reverse of this. It's simpler to conceive of semantics as meaning and syntax as structure. Computers can now read, compare, and extract meaning from human languages thanks to the field of computer science known as natural language processing. To create models from text and speech, natural language relies on the disciplines of linguistics and computer science [10].

The question of "how two words/phrases/documents are similar to each other?" is a fundamental one for research and applications in Natural Language Processing (NLP). To determine how closely two words, phrases, or documents resemble one another, use text similarity. This similarity could be lexical or in meaning.

Lexical similarity refers to the similarity of the word set, whereas semantic similarity refers to the similarity of the meaning. The cosine distance between the two embedded vectors in a sentence serves as an indicator of its semantic similarity. Although many people consider this computation to be difficult, building word or phrase embeddings is significantly more difficult. Text must first be transformed into a vector of features before the algorithm can choose an appropriate features representation, such as TF-IDF. Similarity research on text representation vectors is the last step.

There are numerous methods for determining text similarity, whether they take semantic relations into account or not. In addition to these methods:

1. Jaccard Similarity
2. Cosine Similarity
3. K-Means
4. Latent Semantic Indexing (LSI)
5. Etc.

The degree to which two texts' meanings are similar is gauged by their semantic similarity. Typically, this measurement is scored between 0 and 1. 0 denotes that they are utterly unrelated, and 1 indicates that they nearly share the same meaning.

The fields of linguistics and natural language processing are actively conducting research in text semantic similarity. Additionally, it participates in a variety of applications for informatics and natural language processing. In many natural language processing (NLP) applications, such as sentiment analysis, natural language comprehension, machine translation, question answering, chatbots, search engines, and information retrieval, we make use of semantic similarity.

Applications for informatics sciences exist in the biological and geo-informatics fields. Semantic similarity approaches are mostly used in biomedical informatics to develop biological ontologies, such as Genes Ontology. Ontologies for geographic features used in geo-informatics rely on topological and statistical measures of semantic similarity. The OSM Semantic Network, which calculates the semantic similarity of tags in OpenStreetMap, is one of the most well-known tools for this kind of application [11].

## III. METHODOLOGY

This research was conducted in 3 stages of activity, namely data collection, data analysis, visualization, and data interpretation, as shown in Figure 1. The first stage carried out was the collection of User-generated Content (UGC) data which would be processed in the

next stage. The data will be taken from the Google Maps website, which consists of visitor review data.

Because UGC analysis provides valuable information like opinions, evaluations, suggestions, experiences, or customer demands, it can be very helpful to organizations. The internet, social media, blog posts, tweets, product evaluations, and survey responses with accompanying images all provide open access to this information. UGC typically only exists in unstructured or semi-structured formats, making it impossible to process using conventional data mining methods.

Figure 2 shows visitor reviews on Google Maps which are available in several categories, such as scenery, massage, malls, traffic hours, tattoos, dusk, surf boards, etc. While Figures 3 and 4 show several user reviews based on word occurrence per category. The appearance of the word is the same word as the category or result of the translation. For example, for the category of massage, the results of a review that contains the word massage appear. Meanwhile for the scenery category, the results of a review that contain the word scenery or its translation in other languages appear .

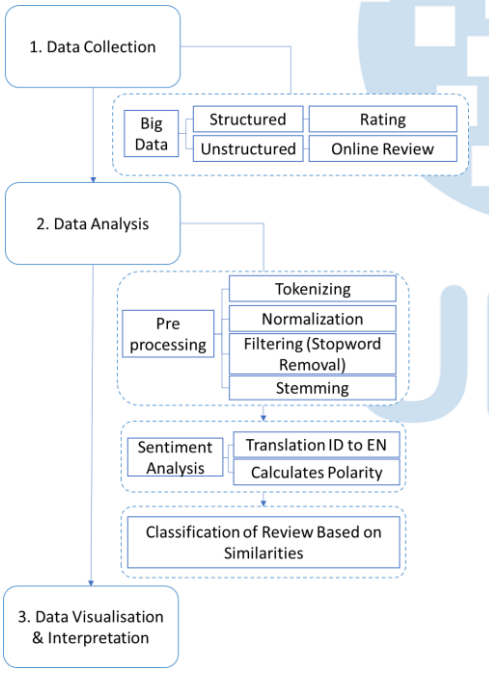


Fig. 1. Research Methodology

In this study, the Semantic Similarity approach was used to find review results that not only contained the exact same keywords as the categories but were also able to find review results that did not contain the exact same keywords but had the same semantics (meaning) as the categories. Semantics in computer science is the mathematical reasoning behind legal string specified by the programming language. It is the opposite of syntax, the rules that govern computer programs. It's easiest to think of syntax as structure and semantics as meaning.

Semantic analysis is a process of extracting meaning from text. Grammatical analysis and recognition of the relationships between certain words in certain contexts allow computers to understand and interpret phrases, paragraphs or even manuscripts. Using semantic analysis can help businesses in many ways, such as interpreting customer reviews more comprehensively. In terms of analyzing visitor reviews, semantic analysis can produce comprehensive information about tourist feedback. This analysis is needed for visitor review data that reviews a category for a tourist spot that does not explicitly state the name of the category[12].

The second stage of this study is data analysis, which includes pre-processing stages, sentiment analysis, and review classification based on similarity of context (similarities). Pre-processing is the first step taken in data mining techniques to convert raw data into cleaner data and ready to be used for further analysis. Some of the pre-processing stages carried out include [13]:

1. *Tokenizing* is the stage of splitting the string into tokens. Previously, the process of number removal and punctuation removal was first carried out.
2. *Normalization* is a process of cleaning words which are words in everyday language/*slank words* be the correct Indonesian word.
3. *Filtering* is a stop word removal process. At this stage, the Indonesian stop word was used which was obtained from the NLTK library (https://www.nltk.org/).
4. *Stemming* is the process of removing word affixes and changing them into basic words

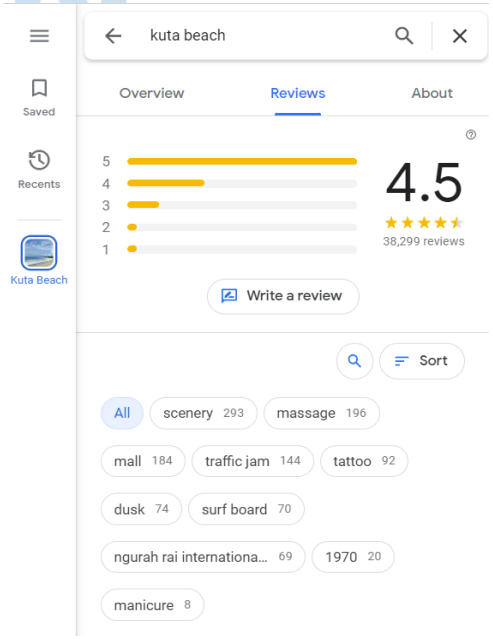


Fig. 2. Visitor Reviews on Google Maps Available in Several Categories

After going through the pre-processing stages, the data is ready to be processed in the semantic similarity analysis process. At this stage, 13 categories have been prepared, which are keywords that are reviewed quite a lot by visitors to tourist attractions. The categories consist of scenery, massage, mall, traffic jam, tattoo, dusk, surfboard, Ngurah Rai Int Airport, Manicure, Facilities, Location, Meals, and Sanitation. This semantic similarity analysis process will produce reviews related to the 13 categories mentioned above.

After going through the semantic similarity process, the next stage is the sentiment analysis process. The sentiment analysis process is carried out using the TextBlob Library (https://pypi.org/project/textblob/). TextBlob is a python library used to analyze data in the form of text. After going through this process, visitor review data is obtained based on the category and sentiment analysis results (positive, negative, or neutral).

The final stage is the interpretation and visualization of the data. Data that has been processed will be easier to understand if displayed in a visual form. Business managers may ultimately find this information useful in gathering client feedback and refining features of their products or services in order to boost their company's value and assist marketing initiatives [2].

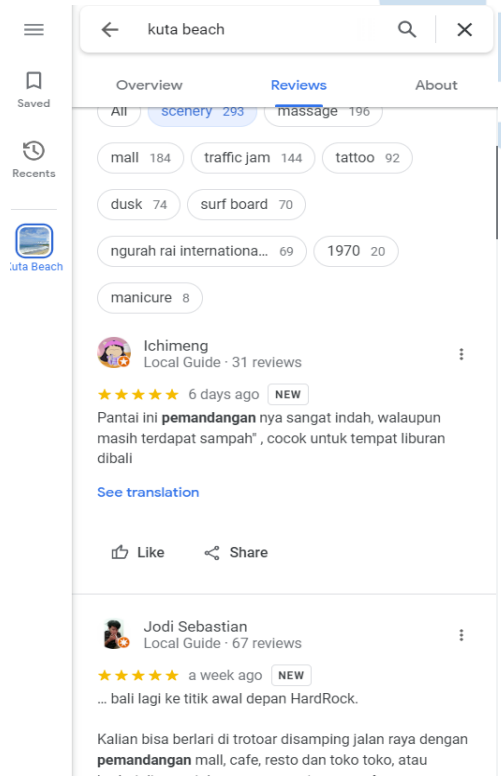


Fig. 3. Classification of Reviews Based on the Appearance of the Word “Massage” Category

## IV. Result

### A. Tools Used

To carry out the semantic similarity process, the SBERT and Cosine Similarity algorithms are used. Sentence-BERT (SBERT) leverages the advanced performance of BERT, with a different architecture. This allows things like cosine similarity to be found more quickly. For example, searching for sentence similarities for 65 hours on BERT would take 5 seconds with SBERT [10].

To implement the SBERT algorithm, we use SentenceTransformers (https://www.sbert.net/) in Python. SentenceTransformers is a Python framework for sophisticated embedding of sentences, text, and images. This framework can be used to calculate sentence/text embedding for more than 100 languages. This embedding can then be compared, e.g., with cosine-similarity to find sentences with similar meanings. This can be useful for semantic textual similarity, semantic search, or paraphrase mining.

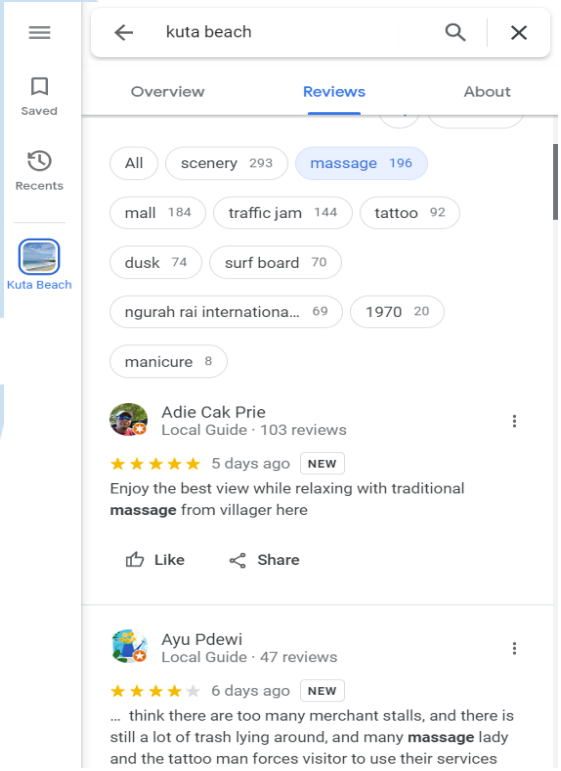


Fig. 4. Classification of Reviews Based on the Appearance of the Word "Massage" Category

Installing Kali Linux requires basic knowledge of The sentiment analysis process is carried out using the TextBlob Library (https://pypi.org/project/textblob/). TextBlob is a python library used to analyze data in the form of text. After going through this process, visitor review data is obtained based on the category and sentiment analysis results (positive, negative, or neutral).

### B. Dataset

The dataset used in this study was taken from visitor review data at Kuta Beach, Bali, Indonesia which can be publicly accessed via the URL https://maps.google.com/. Visitor review data totals the top 120 reviews sorted by the most recent review. The original review data consists of various languages, such as English, Indonesian, etc. which are then translated into English before being processed. The review data contains impressions from visitors who have visited Kuta Beach, Bali.

### C. Experiments

In the Semantic Similarity Analysis experiment that was carried out, the Cosine Similarity calculation was carried out between words in the specified category and sentences that were already available in the corpus. The cosine similarity results range from minus numbers to positive numbers. The higher the result of cosine similarity, the more similar in meaning the two sentences being compared. In this experiment, the cosine similarity limit for sentences that are considered similar is above 0. Only those sentences will be processed in the sentiment analysis process.

The sentiment analysis process carried out with the TextBlob Library is carried out by calculating polarity. After that, the polarity calculation will return the levels of polarity. Polarity lies between (-1, +1) with -1 meaning negative sentiments and +1 meaning positive sentiments. Polarity is reversed by negation words[14].

TABLE I. PERCENTAGE OF SENTIMENT ANALYSIS PER CATEGORY

| No | Category | Sentiment Positive | Sentiment Negative | Sentiment Neutral |
|---|---|---|---|---|
| 1 | Scenery | 0,908333333 | 0,075 | 0,016666667 |
| 2 | Massage | 0,903508772 | 0,078947368 | 0,01754386 |
| 3 | Mall | 0,907563025 | 0,075630252 | 0,016806723 |
| 4 | Traffic Jam | 0,907563025 | 0,075630252 | 0,016806723 |
| 5 | Tattoo | 0,901785714 | 0,080357143 | 0,017857143 |
| 6 | Dusk | 0,906779661 | 0,076271186 | 0,016949153 |
| 7 | Surfboard | 0,908333333 | 0,075 | 0,016666667 |
| 8 | Ngurah Rai Int Airport | 0,929824561 | 0,052631579 | 0,01754386 |
| 9 | Manicure | 0,91011236 | 0,06741573 | 0,02247191 |
| 10 | Facilities | 0,908333333 | 0,075 | 0,016666667 |
| 11 | Location | 0,908333333 | 0,075 | 0,016666667 |
| 12 | Meals | 0,910714286 | 0,071428571 | 0,017857143 |
| 13 | Sanitation | 0,9 | 0,081818182 | 0,018181818 |

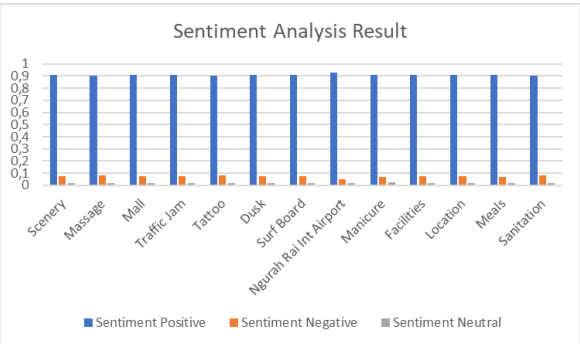


Fig. 5. Sentiment Analysis Result

The results of the sentiment analysis can be seen that positive, negative, and neutral sentiments have similar percentages for each category. This is because the designed algorithm cannot cut sentences for only one category that is assessed. One review being analyzed may contain more than one category, as shown in Figure 5. In this example, the review given by a visitor is a collection of paragraphs. Each paragraph consists of several sentences, where each sentence is the result of a review for one or more types of categories, including scenery, expense, massage, etc.

The algorithm designed in this study is only able to analyze up to the semantic similarity stage but has not been able to cut which sentences are relevant to the category to be analyzed for sentiment. In other words, sentiment analysis will be assessed for one review. So, if there is a review that contains several categories at once, then the overall sentiment analysis will be taken. This is a limitation of this research and will be carried out as future works.

To be able to perform a more comprehensive analysis, it is necessary to improve the designed algorithm. The algorithm should be able to cut sentence by sentence on visitor reviews first. The sentence can then be classified into one or more categories. Sentences that have been classified per new category can then be determined by the sentiment analysis results.

## V. CONCLUSION

In this research, a user-generated content analysis method has been produced in Visitor Reviews of Tourism Objects Using Semantic Similarity. The algorithm designed in this study has limitations, that is only able to analyze up to the semantic similarity stage but has not been able to cut which sentences are relevant to the category to be analyzed for sentiment. In other words, sentiment analysis will be assessed for one review. So, if there is a review that contains several categories at once, then the overall sentiment analysis will be taken. This is a limitation of this research and will be carried out as future works.

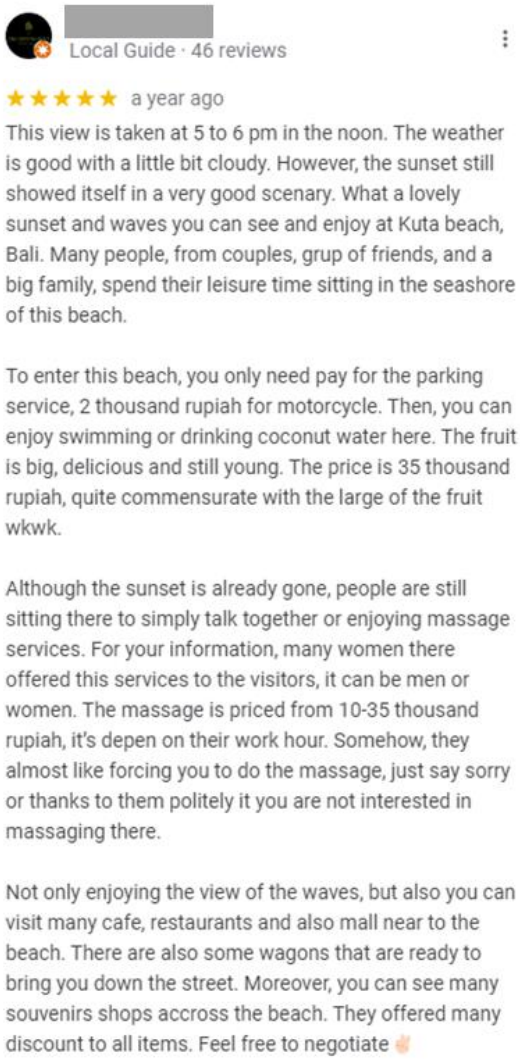


Fig. 6. Example of a review consisting of several categories

As future works, it is necessary to improve the algorithm designed in this study. The algorithm needs to be designed to be able to cut reviews consisting of several sentences, to then be classified based on the review category. Thus, the resulting sentiment analysis can be more detailed, namely on a sentence-by-sentence basis and not generalized for an overall review.

## REFERENCES

[1] C. Science and J. M. Moguerza, "Murga J . et al . ( 2020 ) A Sentiment Analysis Software Framework for the Support of Business Information Architecture in the Tourist Sector . In : Hameurlain A . et al . ( eds ) Transactions on Large-Scale Data- and Knowledge-Centered Systems XLV . Lect," vol. 12390, 2020.

[2] F. Kitsios, M. Kamariotou, P. Karanikolas, and E. Grigoroudis, "Digital marketing platforms and customer satisfaction: Identifying ewom using big data and text mining," *Applied Sciences (Switzerland)*, vol. 11, no. 17, 2021, doi: 10.3390/app11178032.

[3] J. Q. Dong and C. H. Yang, "Business value of big data analytics: A systems-theoretic approach and empirical test," *Information and Management*, vol. 57, no. 1, p. 103124, 2020, doi: 10.1016/j.im.2018.11.001.

[4] N. Balakrishnan, D. Pelusi, and S. Ganesan, "Special issue on ' Big Data Analytics and Deep Learning for E - Business Outcomes ,'" *Information Systems and e-Business Management*, vol. 18, no. 3, pp. 281–282, 2020, doi: 10.1007/s10257-020-00486-0.

[5] W. Lu and S. Stepchenkova, "User-Generated Content as a Research Mode in Tourism and Hospitality Applications: Topics, Methods, and Software," *Journal of Hospitality Marketing and Management*, vol. 24, no. 2, pp. 119–154, Feb. 2015, doi: 10.1080/19368623.2014.907758.

[6] N. Afriliana, N. M. S. Iswari, and Suryasari, "Sentiment Analysis of User-Generated Content: A Bibliometric Analysis," *Journal of System and Management Sciences*, vol. 12, no. 6, pp. 583–598, 2022, doi: 10.33168/JSMS.2022.0634.

[7] N. M. S. Iswari, N. Afriliana, and Suryasari, "User-Generated Content Extraction: A Bibliometric Analysis of the Research Literature (2007–2022)," *Journal of Hunan University Natural Sciences*, vol. 49, no. 11, pp. 120–126, Nov. 2022, doi: 10.55463/issn.1674-2974.49.11.14.

[8] H. Wen, E. Park, C. W. Tao, B. Chae, X. Li, and J. Kwon, "Exploring user-generated content related to dining experiences of consumers with food allergies," *Int J Hosp Manag*, vol. 85, no. August, 2020, doi: 10.1016/j.ijhm.2019.102357.

[9] Y. Liu, A. Soroka, L. Han, J. Jian, and M. Tang, "Cloud-based big data analytics for customer insight-driven design innovation in SMEs," *Int J Inf Manage*, vol. 51, 2020, doi: 10.1016/j.ijinfomgt.2019.11.002.

[10] D. Kaplan, "What Does Semantic Mean In Computer Science." https://enjoymachinelearning.com/blog/finding-semantic-similarity-between-sentences-in-python/ (accessed Apr. 17, 2023).

[11] Baeldung, "Semantic Similarity of Two Phrases." https://www.baeldung.com/cs/semantic-similarity-of-two-phrases (accessed Apr. 17, 2023).

[12] adminlp2m, "Analisis Semantik - Definisi, Cara Kerja dan Contohnya," 2023. https://lp2m.uma.ac.id/2023/02/11/analisis-semantik-definisi-cara-kerja-dan-contohnya/#:~:text=Analisis%20Semantik%2C%20diungkapkan%2C%20adalah%20proses,paragraf%2C%20atau%20bahkan%20seluruh%20manuskrip. (accessed May 04, 2023).

[13] RN Is Here, "Analisa Sentimen Ulasan di Tokopedia - 1," 2021. https://www.kaggle.com/code/raniaprillyap/analisa-sentimen-ulasan-di-tokopedia-1/data (accessed Oct. 12, 2022).

[14] A. A. Chaudhri, S. S. Saranya, and S. Dubey, "Implementation Paper on Analyzing COVID-19 Vaccines on Twitter Dataset Using Tweepy and Text Blob," 2021. [Online]. Available: http://annalsofrscb.ro.

ISSN 2085-4552

# E-Commerce System for Media Group Cooperative

Mikha Melvin Agustian[1], Ririn Ikana Desanti[2], Wella[3]
[1,2,3] Information Systems Study Program, Universitas Multimedia Nusantara, Tangerang, Indonesia
[1]mikha.agustian@student.umn.ac.id, [2]ririn.desanti@umn.ac.id, [3]wella@umn.ac.id

Accepted 09 June 2023
Approved 28 June 2023

***Abstract*— E-Commerce is a way to shop online, where it is easier to process transactions between sellers and buyers without having to meet face-to-face. During the covid-10 pandemic, face-to-face activities should be reduced to prevent the transmission of covid-19 virus. Based on these conditions, it is necessary to create an e-commerce system for the media group employee cooperative so that buying and selling transactions can continue. The SDLC waterfall method was used to identify suitable and appropriate features. The tools used were Visual Studio code and Heidi SQL. The results of this research are an e-commerce system made using PHP, JavaScript, Bootstrap, and CSS for the front end and back end of the website. The proposed system can accommodate buying and selling transactions online and reduce the possibility of data loss in cooperatives.**

***Index Terms*— cooperatives; e-commerce system; waterfall methodology; web-based system**.

## I. INTRODUCTION

Because technological developments are closely related to knowledge, technological advances in this era are inevitable [1] technology provides many conveniences for humans. The development of communication technology in this world is growing rapidly, with the ability to interact and communicate with individuals and groups using the internet [2]. If people want to get information, they just need to use the internet. This makes the activity more practical. The impact of technological advances is that the exchange of information will be easier and faster, easier to work with, more effective and efficient.

With the development of technology and communication, many services have sprung up [2]. One of them is the use of digital cooperatives. The development of technology and communication in cooperatives at this time has penetrated online such as the existence of an e-commerce website. E-commerce users are now a requirement for organizations or companies, so that the company can develop/compete globally [3]. E-commerce will change all marketing activities and at the same time cut operational costs for trading activities (trading).

Media Group has an employee cooperative called KOKARMINDO, which has 1,500 members. Members of this cooperative are permanent employees within the scope of the media group. By having this number of members, this cooperative is required to improve its performance to be more optimal in serving its members. The conditions of the COVID-19 pandemic have caused a division of work schedules, namely work from home (WFH) and work from office (WFO). The division of work schedules causes the cooperative to not be busy, product sales decrease, and it is difficult to disseminate information about cooperative programs or the latest news that is happening in the cooperative.

Therefore, we need an e-commerce website that aims to provide services such as selling goods in cooperatives, disseminating information about motorcycle sales, discount promos for holidays, etc. Cooperative members can easily get information about ongoing programs at the cooperative.

## II. LITERATURE REVIEW

### A. Theory

Cooperatives are one of the economic boosters for the community, become a driving force for the national economy and become a supporter in the economy of the State of Indonesia [4]. According to Rosa, "Cooperatives are associations of people who voluntarily unite themselves in improving their economic welfare by making a business entity managed democratically [5]. From the various meanings of cooperatives, we can conclude that cooperatives are institutions or business entities that have a member in improving their economy as a form of team that is voluntary based on kinship.

SDLC is a model of SDLC (Software Development Life Cycle). for development systems that have been replaced by other means, attempts to overcome one of the shortcomings inherent in traditional SDLC [6]. Development (SDLC) is a methodology that can be used in the description of the process in designing an information system [2], to carry out system development in a structured way.

### B. Previous Study

This research is based on several previous studies. The first research entitled analysis and design of web-based e-commerce systems in students cooperatives [7].

The research resulted in an e-commerce system for student cooperatives that provides various products for student needs. Just like the previous research, this research also builds an e-commerce system for cooperative companies. The second research entitled web-based e-commerce design and business transaction notification system in klister networking system [8]. This research is used as a reference for website design that provides authority to administrators as website managers. The third research entitled designing e-commerce application to increase UMKM revenue [9]. This research is used as a reference for designing a website using waterfall methodology.

## III. Research Methodology

### A. Research Object

The discussion this time will discuss the cooperative which will be the object of this research in the media group employee cooperative, this cooperative was established to help improve the welfare of employees under the auspices of the Media Group. The establishment of Kokarmindo is a Media Group employee cooperative to help provide various needs, such as merchandise, housing, loans, and so on. Currently, Kokarmindo is one of the largest cooperatives in West Jakarta with 1,458 members and 9 business units under the Media Group.

The process of selling various needs is quite long due to the information process. By being one of the largest cooperatives in West Jakarta, it is possible that this cooperative has shortcomings, this can be seen in the sales process and has difficulty in recapitulating data and disseminating information about current programs. cooperatives are held for members; the process is still done manually so that sometimes it slows down performance. This problem can hinder cooperatives in achieving their vision and goals.

### B. Research Methods

The method that can be applied is the waterfall method. The waterfall method can be applied because this method is first carried out to create a flow of the system running process that will be applied to the problems contained in the Media Group Employee Cooperative, namely the process of buying and selling transactions and informing members about programs or information that is currently happening in the cooperative. The system development is done sequentially using the waterfall method [12]. In the early stages of the waterfall method, interviews were conducted with the Media Group Employee Cooperative then given a solution, namely implementing an information system in the form of e-commerce that can assist in making sales. Then a system is made in the form of news on the e-commerce website that is needed and implemented in the Media Group Employee Cooperative.

### C. Data Collection Techniques

Data collection is done by means of in-dept interview. In-dept interview is the process of obtaining information to obtain research objectives by conducting questions and answers while meeting the interviewer face to face with the interviewee [10].

With data obtained through interviews with cooperative management and related parties, namely the chairman of the cooperative and 2 cooperative staff, questions were asked during the interview about how the conditions in the cooperative currently, what problems are being faced, how to deal with the problem, how often and what solutions will be taken. [11] From the interview, you will get the variables needed in making a website and implementing it.

## IV. Analysis and Discussion

### A. Analysis

At this stage an interview was conducted with the chairman of the cooperative, they are Mr. Andreas and Mr. Zay from the media group employee cooperatives. The interview process aims to get information about the data and also an overview of the ongoing business processes in the cooperative so that the functional system that will be needed when designing the system can be known.

Based on the results of interviews conducted with the chairman of the media group employee cooperative on January 20, 2022, with a duration of approximately 45 minutes, from the results of the interview it is known that the current state of business processes, documents and several problems occur in cooperatives, especially in sales monitoring problems. The number of members of the media group employee cooperative is approximately 1,500 members.

The main problem faced is that the sales system at this cooperative is still done manually, users have to make purchases via WhatsApp, especially during this pandemic, employees find it difficult to get information about goods sold by cooperatives. The next problem is the recording of sales and expenditures of goods that are still recorded and stored in the form of a Microsoft Excel file, then also when checking the stock of goods is still done manually, so there is a risk of lost data, and the data is not in sync with the amount of stock available. As for the solution to the first problem, it is proposed to create a website that can conduct online sales transactions. For the second problem, at the back end there will be a feature to view the display of sales of goods and be able to make sales reports.

### B. System Modeling

The system modeling process is carried out using UML diagrams consisting of use case diagrams, activity diagrams and class diagrams. Use case diagram is a diagram that describes an overview of the business process [13]. This system has 3 actors, namely visitors,

users (cooperative members), and admin (cooperative administrators). In fig 1, visitors only have access to see the products being sold as well as cooperative news. Users have access to view news, view products, make product purchases, and view purchase transaction history. The admin will have access to product updates, manage news, manage sales transactions, print sales reports.

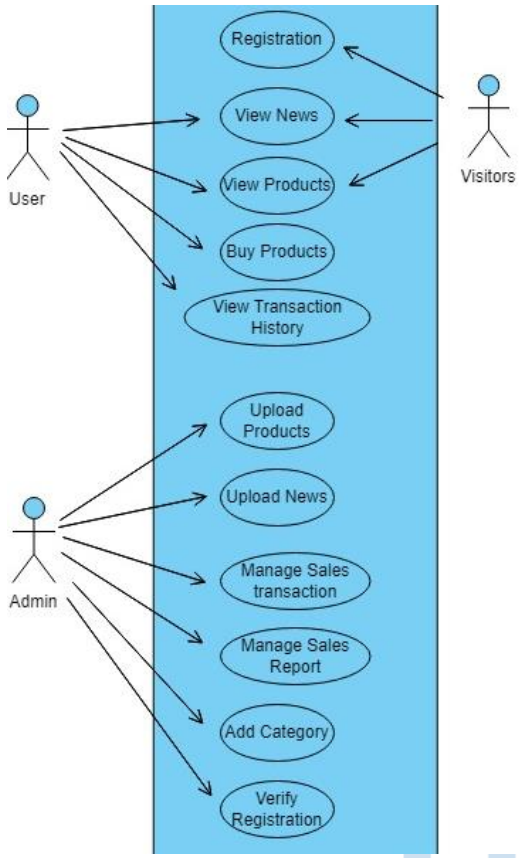


Fig. 1. Use Case Diagram of Proposed Systems

After modeling the main functions of the system, the business processes of each function will be modeled in the form of an activity diagram.

After creating a use case diagram, an activity diagram is created to show the main tasks and how they relate to one another in a process [14].

Fig 2 below shows the business process when visitors want to register.

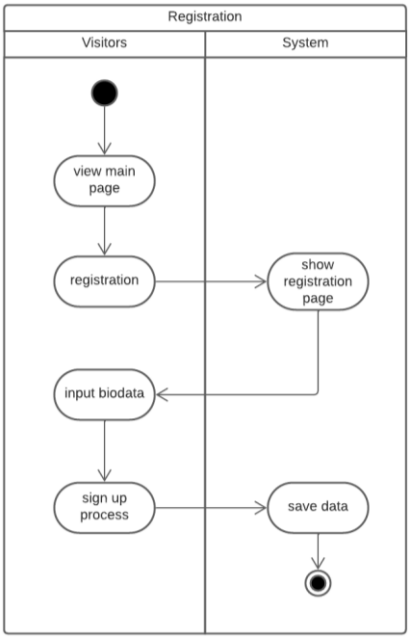


Fig. 2. Activity Diagram Register

Fig 3 below esplains the business process to view list of products sold in the cooperative.

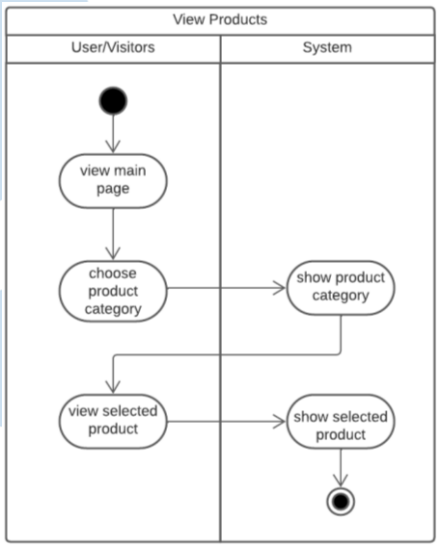


Fig. 3. Activity View Product

Fig 4 below explains the business process for the product purchase process.

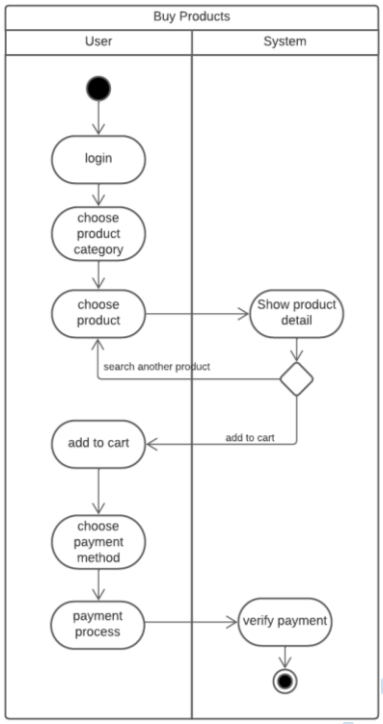


Fig. 4. Activity Diagram Buy Products

Fig 5 below shows the activity details of view transaction history.

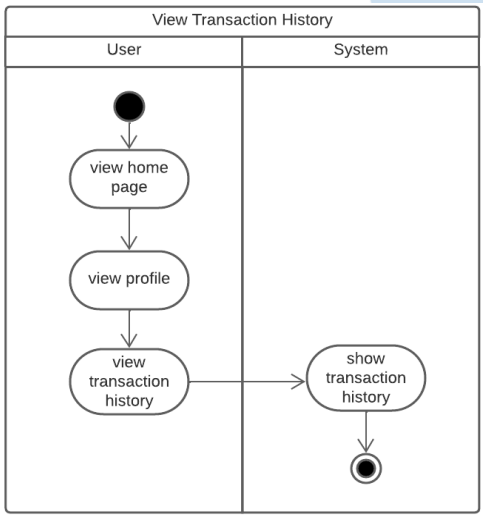


Fig. 5. Activity Diagram View Transaction History

Fig 6 below shows the business process for managing sales transaction.

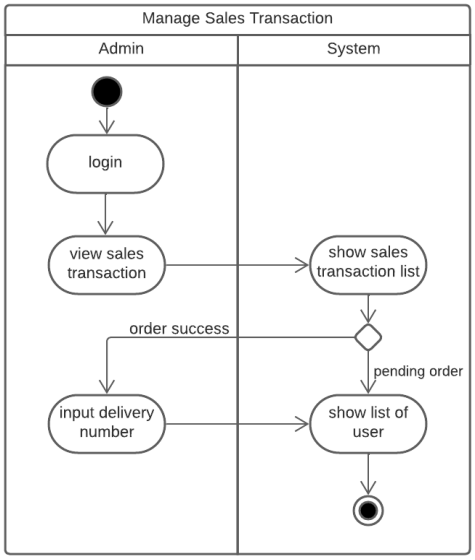


Fig. 6. Activity Diagram Manage Sales Transaction

Fig 7 below shows the activity details of how to upload new products into the system.

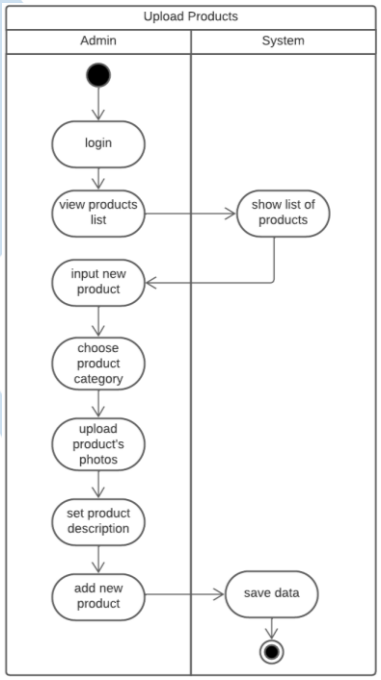


Fig. 7. Activity Diagram Upload Products

Fig 8 below shows the activity details of how to create a new product category.

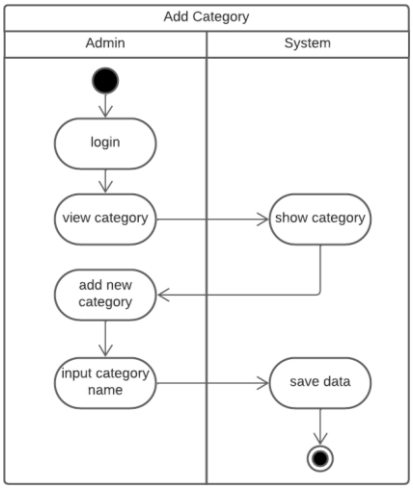


Fig. 8. Activity Diagram Add Category

## C. *System Design*

Interface design is the first step in the system design process. At this point, a thorough description of the system's functions and features is provided [15].

Fig 9 below is a dashboard display. On this page, users will be able to see the products being sold and able to view information or news about the media group employee cooperative.

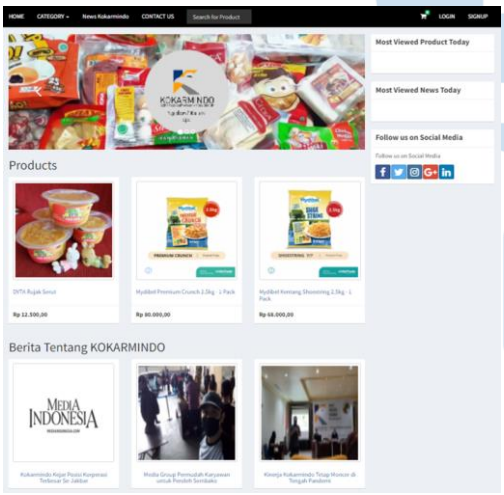


Fig. 9. The homepage of E-Commerce System

Fig 10 below is a product display based on the category that we have selected on the navigation bar. Here contains product information for sale starting from product names, product images, and product prices.

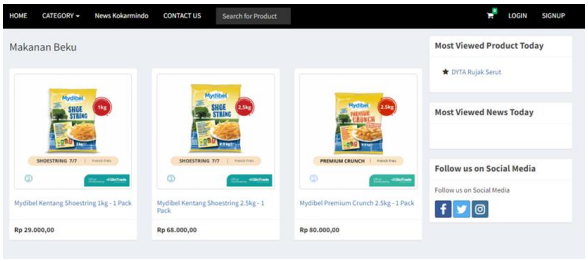


Fig. 10. Display Product by Category

Fig 11 below is a display of detailed products being sold. The information displayed starts from the product name, product image, product category, product stock, product description and product price.

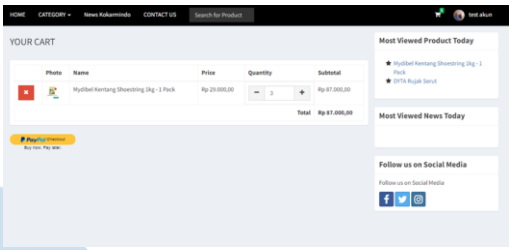


Fig. 11. Display Detail Product

Fig 12 below is a display of cart that contains products to be purchased by an employee.

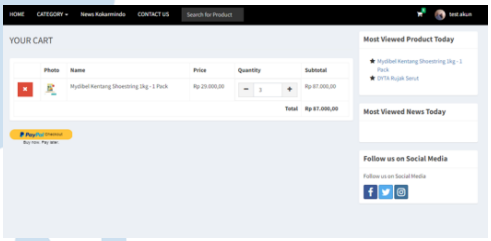


Fig. 12. Display Cart

Fig 13 below is a payment page. Users can choose the desired payment method using Midtrans or Paypal.

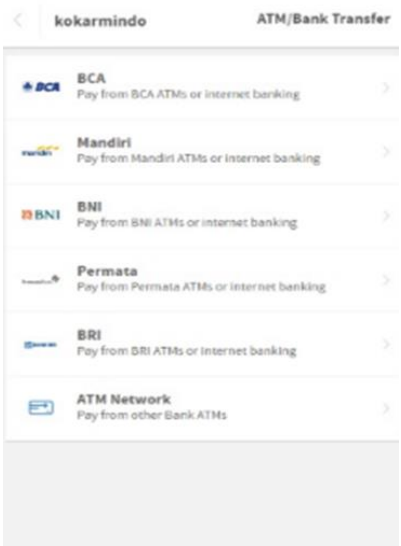


Fig. 13. Display Payment Midtrans

Fig 14 below is a profile page and purchase transaction where users can check their purchase transactions, and also edit their biodata.

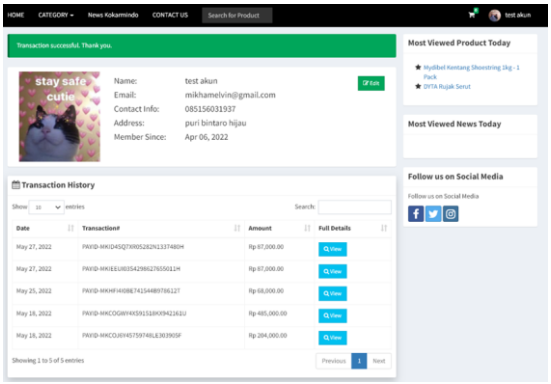


Fig. 14. Display Transaction History

There are several differences between the cooperative before and after using the system. Starting with selling products online and making it easier to monitor transactions that occur in real-time. Table 1 details the differences.

TABLE I. COMPARISON OF OLD SYSTEM AND PROPOSED SYSTEMS

| Category | Old System | Proposed System |
|---|---|---|
| Monitoring | Monitoring related to sales transactions is still done manually by checking the Ms. Excel file of sales transactions | Monitoring can be done at any time because it can be easily accessed through the website |
| Sales | Sales transactions are carried out using chat media such as WhatsApp. | Sales transactions are done online through the website. |
| Recording | Transaction recording is still done using Microsoft Excel and it takes a long time to document the report. | Sales recording is done by the system to make it easier for administrators to see real-time product sales summaries. |

## V. CONCLUSION

This E-Commerce website can improve the performance of cooperatives in making sales and reduce loss of sales report data. The system can expand the marketing of products sold by cooperatives and can be accessed quickly, therefore may increase sales and orders for cooperative products.

To make the system even better, there are suggestions to the next research, such as:

1. The system needs to be connected to a third party as a delivery service provider.
2. There is a need for push notifications via email if there are new products uploaded to the system or if there are products that are on sale.

## ACKNOWLEDGMENT

I would like to send my gratitude to Universitas Multimedia Nusantara especially to Information System Study Program for allowing me to conduct this research.

## REFERENCES

[1] K. Kasmi and A. N. Candra, "Penerapan E-Commerce Berbasis Business To Consumers Untuk Meningkatan Penjualan Produk Makanan Ringan Khas Pringsewu," J. Aktual, vol. 15, no. 2, p. 109, 2017.

[2] A. Tirtana, A. Zulkarnain, B. K. Kristanto, S. Suhendra, and M. A. Hamzah, "Rancang Bangun Aplikasi E-Commerce Untuk Meningkatkan Pendapatan UMKM," J. Ilm. Teknol. Inf. Asia, vol. 14, no. 2, p. 101, 2020.

[3] Alwendi, "Penerapan E-Commerce Dalam Meningkatkan," Manaj. Bisnis, vol. 17, no. 3, pp. 317–325, 2020, [Online]. Available: http://journal.undiknas.ac.id/index.php/magister-manajemen/.

[4] Y. D. Pasca, "Pengaruh Jumlah Anggota, Simpanan, Pinjaman Dan Modal Kerja Terhadap Shu Pada Koperasi Kpri Mitra Kabupaten Majalengka," Eqien J. Ekon. dan Bisnis, vol. 8, no. 2, pp. 322–332, 2021.

[5] R. Fitriana and Novitasari, "Sistem Dan Prosedur Pelaksanaan Simpan Pinjam Pada Koperasi Karyawan Rsu Bina Sehat Rosa," l Akunt. Fak. Ekon. UNIBBA, vol. 10, no. 3, pp. 58–71, 2019.

[6] Y. S. Dwanoko, "Implementasi Software Development Life Cycle ( SDLC ) Dalam Penerapan Pembangunan Aplikasi Perangkat," J. Teknol. Inf., vol. 7, no. 2, pp. 83–94, 2016.

[7] R.I.A. Insan, B. Hariadi, T. Amelia, "analisis dan perancangan sistem e-commerce berbasis web pada koperasi mahasiswa", JSIKA, Vol. 9, No. 3, 2020.

[8] S. Tanone and Radius, "perancangan e-commerce berbasis web dan sistem notifikasi transaksi bisnis pada sistem jejaring klister", Jurnal Teknologi Informasi, Vol. 10, No. 2, 2013.

[9] A. Tirtana, A. Zurkarnain, B.K. Kristanto, Suhendra, M.A. Hamzah, "Rancang Bangun Aplikasi E-Commerce untuk Meningkatkan Pendapatan UMKM", Jurnal Ilmiah Teknologi Informasi ASIA, Vol. 14, No. 2, 2020.

[10] J. W. Satzinger, R. B. Jackson, and S. D. Burd, 6th, Systems Analysis and Design in a changing world. 2012.

[11] T. Hamed, "How to Conduct an Effective Interview; A Guide to Interview Design in Research Study Authors", International Journal of Academic Research in Management (IJARM), 11 (1), pp.39-51. ffhal-03741838, 2022.

[12] R. Sherman. Business Intelligence Guidebook from Data Integration to Analytics. Elsevier. 2015.

[13] A. Dennins, B.H. Wixom, R.M. Roth. System Analysis and Design, 7th Edition. Wiley. 2018.

[14] A. Dennis, B.H. Wixom, D. Tegarden. System Analysis & Design An Object Oriented Approach with UML, 5th Edition. Wiley. 2015.

[15] J. Tidwell, C. Brewer, A. Valencia. Designing Interfaces: Pattern for Effective Interaction Design, 3rd Edition. O'Reilly Media. 2020.

ISSN 2085-4552

# Comparative Study of Robot Framework and Cucumber as BDD Automated Testing Tools

Khaerunnisa[1], Nungki Selviandro[2], Rosa Reska Riskiana[3]
[1,2,3]Faculty of Informatics, Telkom University, Bandung, Indonesia
[1]khaerunnisa@students.telkomuniversity.ac.id, [2]nselviandro@telkomuniversity.ac.id,
[3]rosareskaa@telkomuniversity.ac.id

Accepted 20 June 2023
Approved 11 July 2023

***Abstract*— Automation testing is much more efficient and accurate, the script is easy to document, and update compared to manual testing. Testing a website may necessitate time and effort to learn the tools to be used. Cucumber and Robot Framework are well-known open-source frameworks, according to Stack Overflow and GitHub. Cucumber and Robot Frameworks are known on an international scale, especially Robot Frameworks, which are often used by large companies. Each uses the Java and Python languages, which both support BDD. The comparative efforts of the two tools aim to help testers compare and determine automated testing tools in the BisOps Logee Port Web Admin case study based on the effectiveness and efficiency of the tools and create specific test cases as test documentation so that testers do not need to spend time analyzing both. Because this research involves evaluation and comparison, several criteria were chosen to support the evaluation process, namely functionality, reliability, usability, performance efficiency, and portability. The results of this study show that both tools can be recommended for novice QA's who want to learn the basics of automation by implementing BDD. Meanwhile, for QA's who have done automation before and want to do more in-depth configuration and reporting, it is recommended to use Robot Framework because the syntax is short, has lots of keywords that make it easier for testers, and can make the testing system shorter but more specific.**

***Index Terms*— automated-testing; behavior-driven development; comparative-evaluation; cucumber; Robot-framework**.

## I. INTRODUCTION

Software product testing is one of the phases of the Software Development Life Cycle (SDLC), which aims to find errors in the source code that might cause bugs in software functionality. Testing can improve the quality of software. There are two types of testing, namely manual testing, and automation testing. Automation testing is far more efficient and accurate, and the script can be easily documented and updated compared to manual testing [1]. When automating testing on web-based applications, the tester may need to invest time and effort in learning the tools that will be used later [2].

The method used in this research is BDD. Behavior-Driven Development (BDD) is an agile software development methodology that assists teams in creating high-quality, fast-moving software [3]. BDD was first introduced by Dan North in the early 2000s as an easier way to teach and practice test-driven development (TDD) [4]. The BDD method was chosen because its main benefit is to facilitate organized communication within teams, meaning that product owners, developers, and testers will have a better-shared understanding of how the system works. Requirements written by the customer in a given-when-then format can be immediately used as a starting point for acceptance tests. This means that it is easier for non-developers to participate in the creation of acceptance tests [5]. Many tools that can be used for automation testing have been developed. There are at least 59 tools that can be used for automated testing. The first tool, Cucumber, was chosen because it is a well-known open-source testing tool that supports BDD. Cucumber has 3200 stars on GitHub and 634 forks. Robot Framework was chosen as the second tool because it is popular and supports BDD. Robot Framework has 5600 stars on GitHub and 1600 forks [5]. The first testing tool, Selenium Cucumber, was chosen because it is an internationally renowned framework and open source. The second test tool is the keyword-driven open-source framework (Robot Framework), which was also chosen because it is comprehensive and is used by many large international companies [6] and the robot framework is very easy to use in writing test scripts. The Robot Framework has a very modular architecture [7]. The two test tools are the most well-known testing tools according to the GitHub platform and support BDD, which will be used in this research.

In Juuso Jokio's research, namely "Test automation tools: Robot Framework vs. Selenium-Cucumber" [6], which focuses on automatic testing to test the functionality of the e-mail service, namely the login feature that is implemented in several browsers, there is no specific test case used by researchers. Therefore, in this study, the researcher uses the same framework as the previous study but with more specific test cases, not only for the login feature but also for other features, so that the researcher can see and analyze the differences

between the two betters. Researchers tried comparative efforts on Cucumber and robot frameworks, which were carried out to assist testers in comparing and determining automated testing tool frameworks in the BisOps Logee Port Web Admin case study based on the effectiveness and efficiency of testing tools so that testers do not need to spend time trying and analyzing both, and testers making specific test cases as test documentation, which will help the testers in terms of trying and analyzing both. Because this research involves evaluation and comparison, several criteria were chosen to support the evaluation process, namely functionality, reliability, usability, performance efficiency, and portability [1].

The case study in this research refers to the Web Admin BisOps Logee Port, which is a web-based application for Internal Admin Operations that is useful for managing NPCT-1, NLE, and KOJA master data and is located in Indonesia. NPCT-1 is a web-based one-stop service platform for handling import and export containers and ordering fleets to and from Container Terminals; NLE is a web-based application for a logistics ecosystem that aligns the flow of international goods and documents traffic from the arrival of the means of transportation until the goods arrive at warehouses; KOJA is a web-based one-stop service platform for handling import and export containers and ordering fleets to and from Container Terminals.

This research effort was carried out to help the Logee port QA team (developers) in determining and comparing which tools are better and more efficient for the BisOps Logee Port Web Admin project so that the team does not need to spend time trialing and analyzing the automation tools to be used. Also, complete documentation will make it easier for future developers if they want to make improvements to the website, and developers can consider which features should be fixed, added, or even removed [8][9].

## II. METHODOLOGY

This research will be carried out individually by researchers in two ways, namely, through observation and testing by researchers related to the comparative evaluation of automated testing tools. The testing tools that will be used in this study are the cucumber and robot frameworks by adopting Jureczko's research method, namely the process evaluation tool [10].“Fig. 1” below describes the research process stages.

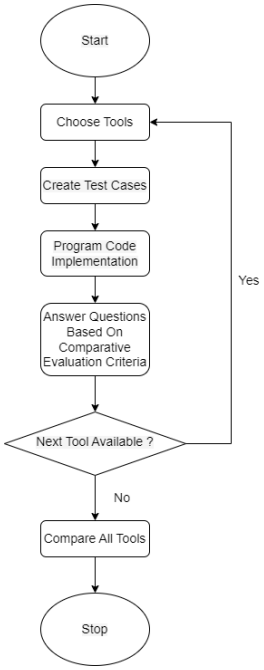


Fig. 1. Describes the Research Process Stages

The following is a detailed explanation of the stages of the research:

### *a) Choose Tools*

There are two testing tools used in this study, namely Cucumber and the Robot Framework, which were evaluated using the Eclipse IDE. An IDE (integrated development environment) is needed as a basis for writing and running test cases and generating test script code. Eclipse IDE was chosen because it is a Java platform. The Java programming language is the most popular and frequently used lately [11]. IDE is an open-source, commonly used tool that can automate many functions that are usually written manually by developers and can be integrated with Cucumber and robot frameworks [12]. There are three main folders in the Cucumber project:

- \src\test\java which contains 2 more folders, page factory, and step definitions. The page factory folder is a folder used to write the function of the step in the step definitions folder based on the page or feature page being tested. Meanwhile, the step definitions folder is a folder that is used to write down the steps of the test in detail according to the test cases that have been made in the file .feature using the BDD method. In this folder, there is also a runner that is used if the tester wants to run a test, which will later be used to get a report from the test.
- \src\test\resources which contains the features folder, where this folder is the folder to hold all the feature files.
- The target folder is the folder that contains the report.

Meanwhile, there are two main folders in the robot framework project:

- Resources, in which there is a Pagefactory folder that is used to write the functions of the test cases in the test suite. The files in this folder contain *** Settings ***, *** Variables ***, and *** Keywords ***.
- Test Suite is a folder to hold files .robot is used to store test cases made using the BDD method. Files in this folder contain *** Settings *** and *** Test Cases ***.

*b) Create Test Cases*

In making test cases, both tools use the BDD principle, where the format of the test cases will be the same. In the first test tool, namely Cucumber, test cases are created in the feature file [13]. Whereas in the second test tool, namely the robot framework, test cases are made in the script file, namely the robot file [14].

Due to time constraints, the researcher chose to implement the main features of the Logee Port admin website. The selected test cases are the result of discussions with the Logee Port QA team regarding features that are often used in using the Logee Port admin website. There are six frequently used features: login features, dashboard menus, dashboards, containers, truck orders, and transactions. Of the 6 features, 12 test cases are often used and implemented in this study.

The following test cases are:

- feature to test login functionality.
- feature to test dashboard functionality Transaksi Per Periode.
- feature to test dashboard functionality Laporan Aktivitas Hari Ini untuk fitur belum dibayar.
- feature to test dashboard functionality Laporan Aktivitas Hari Ini untuk fitur telah dibayar.
- feature to test dashboard functionality Laporan Performa Bisnis.
- feature to test dashboard functionality armada.
- feature to test dashboard2 functionality lihat detail port.
- feature to test dashboard2 functionality lihat detail truk.
- feature to test transaksi functionality cari nomor proforma.
- feature to test the pemesanan truk page functionality.
- feature to test kontainer functionality.
- feature to test logout functionality.

*c) Program Code Implementation*

In Cucumber, the program code or test script code is implemented in the definitions step, which calls the main function on the page factory using the Java language, and all test cases defined in the feature file are mapped using annotations in the step definition file. Configuration for running test cases and reports is done on a test runner or file runner using Junit and Maven. Meanwhile, in the robot framework, the program code or test script code is implemented in resources using the Python language, and all test cases defined in the robot file are mapped using annotations in the resource file. configuration in running test cases and reports using the built-in robot framework feature, namely logs.

*d) Answering Questions Based on Comparative Evaluation Criteria*

The next step after the test cases have been implemented and executed is to answer questions based on comparative evaluation criteria. Questions are answered in two ways: through system testing and observation. The description of these two methods can be seen in "TABLE I" below. There are three modifications to the comparative evaluation questions referred to in Sandin's research [15].

The following are the three-modification explained:

- Modification of the functionality criteria in which Sandin's research reference implements the "unit testing" context while this research implements the "BDD testing" context.
- Elimination of one of the questions on the functionality criteria related to the number of methods of the two tools. This is done because based on observations, researchers do not use methods in their services. Researchers want to evaluate from a more objective standpoint. Therefore, as a substitute, the author describes the features in the analysis of the results.
- Adding questions to the usability criteria regarding the author's length of time studying and writing programs This is done because the author is using these two tools for the first time, and later this research can be useful for novice QAs who want to start automating.

"Table I" below explains modified comparison questions.

TABLE I. MODIFIED COMPARISON QUESTIONS [15].

| Criteria | Questions | Possible Answers |
|---|---|---|
| Functionality | Test tool simplicity in BDD testing implementation? (Tested in system) | • Easy<br>• Medium<br>• Hard |
| | How many lines of code need to be executed for each case study? (Tested in system) | Lines of code |

| | | |
|---|---|---|
| Reliability | Can the tool detect and perform error checking in any condition? (Tested in system) | • Possible<br>• Not possible |
| Usability | Is the tool documentation provided available and can the user rely on that documentation to understand and learn about the tool used? (Tested based on observation) | • Easy (Available with many documents provided)<br>• Medium (Available but needs more effort)<br>• Hard (Available but need hard effort or not available at all) |
| | Can the user understand the code (test script)? (Tested in system) | • Easy<br>• Medium<br>• Hard |
| | How long has it taken me to study and write test code? (Tested in system) | Time to study and write the test code. |
| Performance efficiency | How long is the execution time to perform the past and failed testing and resource used? (Tested in system) | Execution time |
| Portability | Ease of installation.? (Tested based on observation) | • Easy<br>• Medium<br>• Hard |
| | Can the tool integrate with the development environment or in other words, run on different platforms? (Tested based on observation) | • Can run in many IDE<br>• Can but in certain IDE only<br>• Not portable |

*e) Comparison of All Tools*

The comparison of the two test tools is done by comparing the performance results of the two test tools that have been executed based on test cases. After that, an analysis related to the comparison was carried out based on the questions that had been answered by the two testing tools. In the final step, the researcher makes a description and conclusion regarding the advantages and disadvantages of the testing tool.

## III. Results and Discussion

In this section, the researcher evaluates the two test tools described in the previous chapter. The following are more detailed test results and analyses of test results:

*a) Test result*

"Table II" describes the results of the evaluation of the two tools used in this study, namely the cucumber and robot frameworks, based on the comparative evaluation criteria of the implemented test cases. The perspective reviewed is a more detailed matter for each criterion reviewed.

TABLE II. The Results of Evaluation of the Two Tools

| Criteria | Aspects Considered | Tools | |
|---|---|---|---|
| | | Cucumber | Robot framework |
| Functionality | Test Implementation | Easy | Easy |
| | Lines of code | 1.379 lines | 567 lines |
| Reliability | Error checking | Possible | Possible |
| Usability | Documentation and learning | Easy | Easy |
| | Code readability | Medium | Easy |
| | Time to study and write the test code | 16 days | 10 days |
| Performance Efficiency | Execution time (seconds) | 491.314 seconds (8 minutes 11 seconds) | 276.438 seconds (4 minutes 36 seconds) |
| Portability | Ease of installation | Easy | Easy |
| | Integrated development environment | Can but in certain IDE only | Can run on many IDE |

*b) Analysis of Test Results*

*1) Functionality*

In its implementation, the test cases and test scripts for both frameworks, namely the Cucumber and Robot Frameworks, adopted the BDD principles almost the same. Cucumber requires a file .feature to store BDD test scripts, while the robot framework requires a file .robot to store BDD test scripts.

The basic difference is the language used; Cucumber uses Java language. Meanwhile, the robot framework uses Python language. "Fig. 2" below shows cucumber with Java libraries and robot framework with Python libraries.

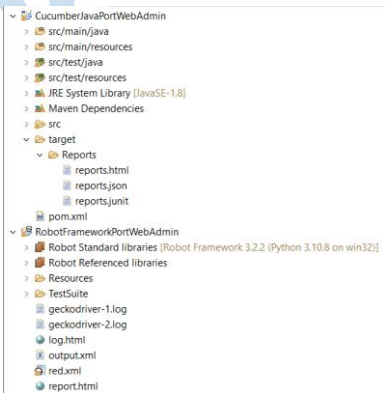


Fig. 2. Cucumber with Java Libraries and Robot Framework with Python Libraries

Another difference lies in the use of the library. The Cucumber library facilitates the implementation of annotations and methods that are implemented in test cases and test scripts. There are 10 annotations used in this study, which are the basic annotations needed so that test cases can be run. Meanwhile, the robot framework uses a keyword-based standard library,

which in its implementation allows BDD test scripts that have been made before to be used as keywords and called back by other functions without the need for a constructor like the one in Cucumber. The keywords used in this study are 17; these are the basic keywords needed for the test cases to run.

In running one test case, the cucumber requires 3 files, namely:

- File .feature (for writing BDD scripts, scenarios, or test scripts). "Fig. 3" below shows the cucumber file .feature.

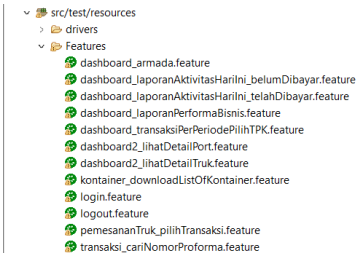


Fig. 3. Cucumber File.Feature

- file.java in the step definition folder (to write the implementation of BDD scripts that have been made in the file .feature). "Fig. 4" below shows the cucumber step definitions folder.

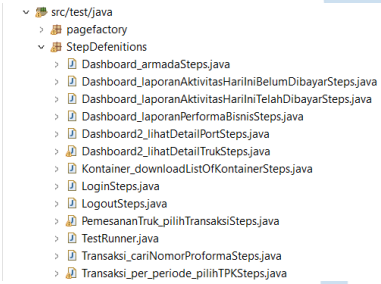


Fig. 4. Cucumber Step Definitions Folder

- file.java in the page factory folder (to write the functions that will be called in implementing the BDD script in the step definition folder). "Fig. 5" below shows the cucumber PageFactory folder.

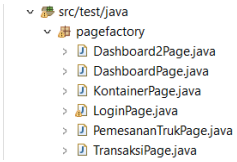


Fig. 5. Cucumber PageFactory Folder

To run all test cases, one additional file is needed, namely the test runner (file.java, which contains the runner file so that the test cases can be run). The configuration of this test runner requires an annotation, namely @CucumberOption, which is shaped like an array of associations and can be changed according to needs. "Fig. 6" below shows the cucumber test runner.

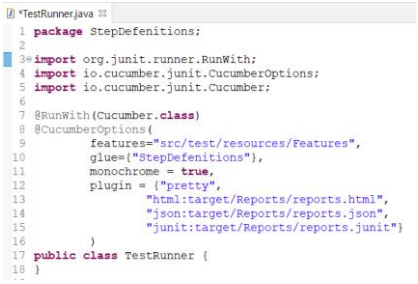


Fig. 6. Cucumber Test Runner

Meanwhile, in running one or all the test cases on the robot framework, only two files .robot are needed, namely:

- File .robot in the test suite folder (to write BDD scripts, scenarios, or test scripts). "Fig. 7" below shows the Robot Framework Test Suite Folder.

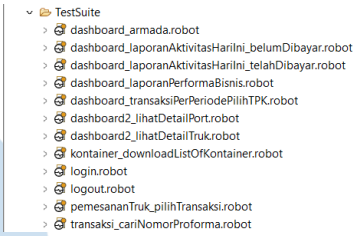


Fig. 7. Robot framework TestSuite Folder

- File .robot in the resource folder (to write the implementation of the BDD script that was created in File .robot in the previous test suite folder, as well as the implementation of the functions needed in the BDD test script). "Fig. 8" below shows the robot framework Resources/PageFactory Folder.

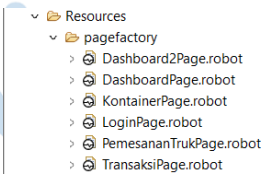


Fig. 8. Robot framework Resources/PageFactory Folder

The last difference lies in the reporting. In Cucumber, in terms of generating reports, it is still done manually, namely by creating a manual folder and manually creating a .html file that will later be included in the test runner file so that later the report can be generated, and every time you want to open a report, you have to refresh it first so that the report or reporting can be updated to the file that was run most recently. "Fig. 9" below shows the cucumber Report Folder.

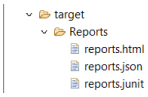


Fig. 9. Cucumber Report Folder

Meanwhile, in the robot framework, reports are generated automatically, so after installing the robot framework, the console will automatically have the message log and execution view options. In fact, after running the test case on the default console, a link will be given to see more detailed reporting. "Fig. 10" below shows the robot framework message log dan execution view console.

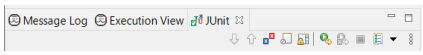


Fig. 10. Robot framework Message Log dan Execution View Console

"Fig. 11" below shows the robot framework default console.

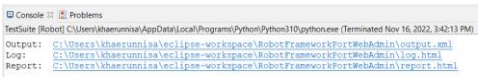


Fig. 11. Robot framework Default Console

There are 4 types of output in Cucumber: the default console, Junit console, HTML, and JSON. Whereas in the robot framework, there are six types of output: default console, console execution view, console message log, output.xml, log.html, and report.html. Cucumber only displays the BDD test script in the .html format output for cases with a green tick or a red cross symbol indicating whether a test case was successful or not. Whereas in the robot framework, there is a very informative log option, where in the output detailed information, such as time and others, is given, and if there is an error, a screenshot of the error page, along with the error code and some other additional information, is displayed.

*2) Reliability*

When implementing the library in test cases and code, there are no problems with the Cucumber and Robot frameworks. The problem that was found was the Selenium web driver, which executed the program too quickly while the browser was still in a loading state, which caused the test case to fail to run and the console to display the web driver used to reach out. This can be overcome by using the Selenium function itself, namely WebDriverWait(), which waits for a while before executing the next step, or you can also write code in the cucumber, namely Thread .sleep(millisecond); and on the robot framework, namely the sleep second keyword, which functions to wait for a few seconds before executing the next step according to the number of seconds' input.

*3) Usability*

Cucumber and robot frameworks have learning documentation or user guides regarding how to use them on their respective official websites.

Cucumber documentation can be accessed at the following link: https://docs.cucumber.io/docs/guides/ and robot framework documentation can be accessed at the following link: https://robotframework.org/robotframework/latest/RobotFrameworkUserGuide.html. Both documents have good structure and detailed information. Both also have guides for beginners, such as in Cucumber's introduction and Robot Framework's getting started, and in both, there are examples of program code to give users a better understanding. When implementing the program code, both tools have end-to-end documentation. So, when hovering over the library or keywords, information about the libraries or keywords used will be displayed, which helps writers in writing program code using these two tools. As a result, it is possible to conclude that these two test tools are simple to document and learn. "Table III" below shows forum activeness comparison on stack overflow.

TABLE III. FORUM ACTIVENESS COMPARISON ON STACK OVERFLOW

| Feature | Cucumber | Robot Framework |
|---|---|---|
| The whole question | 10.585 | 6.499 |
| Questions without answers | 1.777 | 1.089 |
| Questions without upvotes or answers are accepted | 4.205 | 2.515 |

"Table IV" below shows a forum activeness comparison on GitHub.

TABLE IV. FORUM ACTIVENESS COMPARISON ON GITHUB

| Feature | Cucumber | Robot Framework |
|---|---|---|
| Stars | 2.507 | 7.500+ |
| Watchers | 223 | 484 |
| Forks | 2.000+ | 2.000+ |
| Last Commit | 15 November 2022 5.970 commits | 12 November 2022 13.771 commits |

In both tools, the program code can be read properly, as explained in the functionality sub-chapter. The difference lies in the difference in language, which makes the robot framework shorter because the syntax uses Python and because the robot framework doesn't require a test runner to run its test cases. The author takes 16 days to learn and write using the Cucumber tools and 10 days using the Robotframework tools.

*4) Performance Efficiency*

In the Sandin research, the test cases were run three times [15]. The author also does the same thing to find out the average execution time value of each tool or framework. Robot Framework executes test cases 2 times faster than Cucumber, this includes when Cucumber and Robot Framework generate reports automatically.

*5) Portability*

These two tools use different languages; therefore, the libraries and dependencies used are also different. The installation process or setup of the first test tool, namely Cucumber, can be seen as follows:

- Created a new Maven project.

- Adding Maven dependencies Cucumber Java | Cucumber JUnit | JUnit | Selenium Java. "Fig. 12" below shows cucumber dependency.


Fig. 12. Cucumber Dependency

The installation process or setup of the first test tool, namely the robot framework, can be seen as follows:

- Check if the device has Python, if not then install Python."Fig. 13" below shows Python version 3.10.8.

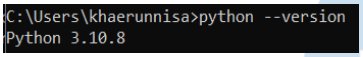


Fig. 13. Python Version 3.10.8

- Check if the device has pip, if not then install pip."Fig. 14" below shows pip version 22.3.

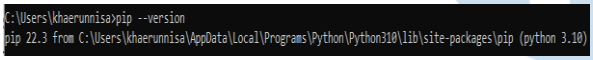


Fig. 14. Pip Version 22.3

- In the command prompt, type "pip install robot framework to install robot framework" to install the robot framework. "Fig. 15" below shows robot framework version 6.0.

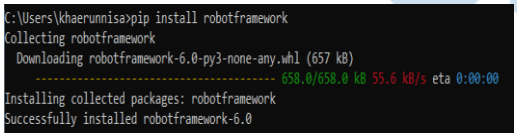


Fig. 15. Robot Framework Version 6.0

- Download Eclipse RED—the Robot Editor—from the Eclipse Marketplace.
- Added the path to RED in Eclipse / Windows / Preferences / RF / Installed FWs

The author was confused when trying to install the robot framework on the Eclipse IDE because, according to the robot framework user guide, the way to install it is to use the command prompt and input the pip install robot framework command. The version to be installed is the latest, namely version 6.0. Meanwhile, RED or the Eclipse robot editor can only support robot framework 3. x and Eclipse IDE version 2020-06 (4.16). Overall, the installation process for these two tools is easy. "Fig. 16" below shows the RED user guide.

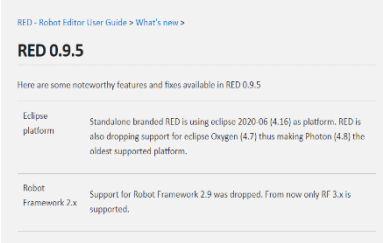


Fig. 16. RED User Guide

Regarding integration with other platforms, based on what is stated on each tool's official website, Cucumber can only be implemented in the Eclipse IDE and IntelliJ IDEA and can also be implemented in the NetBeans IDE, which may still be under development. Meanwhile, robot framework can be implemented in almost any IDE, such as RIDE, sublime plugin, atom plugin, notepad++, IntelliSense for Visual Studio Code, and many more.

## IV. Conclusion

Robot Framework is easier to use than Cucumber. This can be seen in the table of test results in Sub-Chapter 4. Where 5 differences make the robot framework superior, namely:

1. On the functionality criteria. The test script's lines of code require only 567 lines of code to run the 14 test cases discussed previously. Meanwhile, cucumber requires 1,379 lines of code to run the same test case.
2. On usability criteria. The readability of the program code on the cucumber has an intermediate status, while the robot framework has an easy status. This is because Cucumber uses the Java language and the robot framework uses Python, which makes the syntax shorter.
3. Still on usability criteria. The author's time spent studying and writing test code, where learning and writing code on Cucumber took the writer approximately 16 days, but only 10 days on the Robot Framework. This is because, apart from the short syntax of Python, you don't need to build a constructor to call one function to another like the one in Cucumber. With the keyword facility in the robot framework, it makes it easier for writers to call functions.
4. On Performance Efficiency criteria. The execution time of the Robot framework is quite short, at 4 minutes and 36 seconds, whereas the cucumber takes 8 minutes and 11 seconds to execute the same test case.
5. On the portability criteria, the robot framework can be integrated with almost any IDE

platform. Meanwhile, Cucumber can only be integrated into certain IDEs.

Although the Cucumber forum is more active on the Stack Overflow platform, with a total of 10,585 questions, the RobotFramework forum only has 6,499 questions. However, on the GitHub platform, the robot framework is far more popular, with more than 7,500 stars, while Cucumber only has 2,507 stars as of November 16, 2022. This shows that Robotframework is no less competitive than Cucumber in terms of popularity.

In terms of reporting, automatic reporting from the robot framework has very clear details, time information, error messages, and logs, even when an error occurs, a screenshot of the error page will be displayed, and many other features. which is not owned by the cucumber's automatic reporting.

The two tools in this study, Cucumber, and Robot Framework can be recommended for novice QAs who want to learn the basics of automation and implement automated BDD testing easily. However, specifically for QAs who have done automation before and want to do more in-depth configuration and reporting, the authors recommend using a robot framework because, in addition to having a short syntax, it also has many keywords that make it easier for testers and can make the testing system shorter but more specific.

Suggestions for future work are to make comparisons with different criteria and different case studies and to use more test cases to get a longer total time result so that the comparison can be seen more clearly.

## ACKNOWLEDGMENT

This paper and the research behind it would not have been possible without the exceptional support of my lecturers, Nungki Selviandro, S. Kom, M. Kom, Ph.D., and Rosa Reska Riskiana, S.T., M.T.I. Their enthusiasm, knowledge, and exacting attention to detail have been an inspiration and kept my work on track. Their generosity and expertise have improved this study in innumerable ways and saved me from many errors.

## REFERENCES

[1] SVS College of Engineering, Institute of Electrical and Electronics Engineers. Madras Section, and Institute of Electrical and Electronics Engineers, *Proceedings of 2017 Second IEEE International Conference on Electrical, Computer, and Communication Technologies : 22-24, February 2017, SVS College of Engineering, Coimbatore, Tamil Nadu, India*.

[2] T. Akhir, “Evaluasi Komparatif Automated Behavior Driven Development Testing Tool Framework pada Aplikasi Berbasis Web menggunakan Comparative Evaluation Criteria.

[3] T. Couto, S. Marczak, and F. Gomes, “On the Understanding of How to Measure the Benefits of Behavior-Driven Development Adoption: Preliminary Literature Results from a Grey Literature Study,” in *ACM International Conference Proceeding Series*, Dec. 2020. doi 10.1145/3439961.3440000.

[4] S M. Härlin, “Testing and Gherkin in agile projects.

[5] V. Österholm, “Overview of Behaviour-Driven Development tools for Web applications,” 2021.

[6] J. Jokio, “Test automation tools Robot Framework vs. Selenium-cucumber,” 2020.

[7] N. Huu and N. Triem, “RESEARCH AND COMPARE SOME FRAMEWORKS COMMONLY USED IN AUTOMATION TESTING GRADUATION THESIS SUMMARY INFORMATION TECHNOLOGY Student : Luong Khac Tuan Anh-ID: 17IT001,” 2017.

[8] U. Nugraha, S. Atikah Nurduha Robaiah, and D. Rospinoedji, “TESTING THE INFORMATION SYSTEM SOFTWARE USING BEHAVIOR DRIVEN DEVELOPMENT METHOD Ucu Nugraha, Siti Atikah Nurduha Robaiah, Djoko Rospinoedji. Testing The Information System Software Using Behavior Driven Development Method-Palarch’s Journal Of Archaeology Of Egypt,” vol. 17, no. 10, 2020.

[9] H. Lopes Tavares, G. G. Rezende, V. Mota Dos Santos, R. Soares Manhães, and R. Atem De Carvalho, “A tool stack for implementing Behaviour-Driven Development in Python Language.” [Online]. Available: http://www.renapi.org/biblioteca-digital/ferramentas.

[10] M. Jureczko and M. Mlynarski, “Automated acceptance testing tools for web applications using Test-Driven Development,” 2010.

[11] K. Kumar and S. Dahiya, “Programming Languages: A Survey International Journal on Recent and Innovation Trends in Computing and Communication Programming Languages: A Survey”, [Online]. Available: http://www.ijritcc.org.

[12] “geer2005”.

[13] M. Wynne and A. Hellesøy, “What Readers Are Saying About..

[14] Sumit. Bisht, *Robot Framework Test Automation.* Packt Publishing, 2013.

[15] R. Mohamad and N. M. Yassin, “Comparative Evaluation of Automated Unit Testing Tool for PHP Development of Dengue-Entomological Surveillance System View project SBSE, meta-heuristic search algorithms and Artificial immune system View project.” [Online]. Available: https://www.researchgate.net/publication/313208886.

# AUTHOR GUIDELINES

## 1. Manuscript criteria

- The article has never been published or in the submission process on other publications.
- Submitted articles could be original research articles or technical notes.
- The similarity score from plagiarism checker software such as Turnitin is 20% maximum.
- For December 2021 publication onwards, Ultimatics : Jurnal Teknik Informatika will be receiving and publishing manuscripts written in English only.

## 2. Manuscript format

- Article been type in Microsoft Word version 2007 or later.
- Article been typed with 1 line spacing on an A4 paper size (21 cm x 29,7 cm), top-left margin are 3 cm and bottom-right margin are 2 cm, and Times New Roman's font type.
- Article should be prepared according to the following author guidelines in this template. Article contain of minimum 3500 words.
- References contain of minimum 15 references (primary references) from reputable journals/conferences

## 3. Organization of submitted article

The organization of the submitted article consists of Title, Abstract, Index Terms, Introduction, Method, Result and Discussion, Conclusion, Appendix (if any), Acknowledgment (if any), and References.

- Title
  The maximum words count on the title is 12 words (including the subtitle if available)
- Abstract
  Abstract consists of 150-250 words. The abstract should contain logical argumentation of the research taken, problem-solving methodology, research results, and a brief conclusion.
- Index terms
  A list in alphabetical order in between 4 to 6 words or short phrases separated by a semicolon ( ; ), excluding words used in the title and chosen carefully to reflect the precise content of the paper.
- Introduction
  Introduction commonly contains the background, purpose of the research, problem identification, research methodology, and state of the art conducted by the authors which describe implicitly.
- Method
  Include sufficient details for the work to be repeated. Where specific equipment and materials are named, the manufacturer's details (name, city and country) should be given so that readers can trace specifications by contacting the manufacturer. Where commercially available software has been used, details of the supplier should be given in brackets or the reference given in full in the reference list.
- Results and Discussion
  State the results of experimental or modeling work, drawing attention to important details in tables and figures, and discuss them intensively by comparing and/or citing other references.
- Conclusion
  Explicitly describes the research's results been taken. Future works or suggestion could be explained after it
- Appendix and acknowledgment, if available, could be placed after Conclusion.
- All citations in the article should be written on References consecutively based on its' appearance order in the article using Mendeley (recommendation). The typing format will be in the same format as the IEEE journals and transaction format.

## 4. Reviewing of Manuscripts

Every submitted paper is independently and blindly reviewed by at least two peer-reviewers. The decision for publication, amendment, or rejection is based upon their reports/recommendations. If two or more reviewers consider a manuscript unsuitable for publication in this journal, a statement explaining the basis for the decision will be sent to the authors within six months of the submission date.

## 5. Revision of Manuscripts

Manuscripts sent back to the authors for revision should be returned to the editor without delay (maximum of two weeks). Revised manuscripts can be sent to the editorial office through the same online system. Revised manuscripts returned later than one month will be considered as new submissions.

## 6. Editing References

- **Periodicals**
  J.K. Author, “Name of paper,” Abbrev. Title of Periodical, vol. x, no. x, pp. xxx-xxx, Sept. 2013.

- **Book**
  J.K. Author, “Title of chapter in the book,” in Title of His Published Book, xth ed. City of Publisher, Country or Nation: Abbrev. Of Publisher, year, ch. x, sec. x, pp xxx-xxx.

- **Report**
  J.K. Author, “Title of report,” Abbrev. Name of Co., City of Co., Abbrev. State, Rep. xxx, year.

- **Handbook**
  Name of Manual/ Handbook, x ed., Abbrev. Name of Co., City of Co., Abbrev. State, year, pp. xxx-xxx.

- **Published Conference Proceedings**
  J.K. Author, “Title of paper,” in Unabbreviated Name of Conf., City of Conf., Abbrev. State (if given), year, pp. xxx-xxx.

- **Papers Presented at Conferences**
  J.K. Author, “Title of paper,” presented at the Unabbrev. Name of Conf., City of Conf., Abbrev. State, year.

- **Patents**
  J.K. Author, “Title of patent,” US. Patent xxxxxxx, Abbrev. 01 January 2014.

- **Theses and Dissertations**
  J.K. Author, “Title of thesis,” M.Sc. thesis, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year. J.K. Author, “Title of dissertation,” Ph.D. dissertation, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.

- **Unpublished**
  J.K. Author, “Title of paper,” unpublished. J.K. Author, “Title of paper,” Abbrev. Title of Journal, in press.

- **On-line Sources**
  J.K. Author. (year, month day). Title (edition) [Type of medium]. Available: http:// www.(URL) J.K. Author. (year, month). Title. Journal [Type of medium]. volume(issue), pp. if given. Available: http://www.(URL) Note: type of medium could be online media, CD-ROM, USB, etc.

## 7. Editorial Adress

Universitas Multimedia Nusantara
Jl. Scientia Boulevard, Gading Serpong
Tangerang, Banten, 15811
Email: **ultimatics@umn.ac.id**

# Paper Title

## Subtitle (if needed)

Author 1 Name[1], Author 2 Name[2], Author 3 Name[2]
[1] Line 1 (of affiliation): dept. name of organization, organization name, City, Country
Line 2: e-mail address if desired
[2] Line 1 (of affiliation): dept. name of organization, organization name, City, Country
Line 2: e-mail address if desired

Accepted on mmmmm dd, yyyy
Approved on mmmmm dd, yyyy

***Abstract*—This electronic document is a "live" template which you can use on preparing your ULTIMATICS paper. Use this document as a template if you are using Microsoft Word 2007 or later. Otherwise, use this document as an instruction set. Do not use symbol, special characters, or Math in Paper Title and Abstract. Do not cite references in the abstract.**

***Index Terms*—enter key words or phrases in alphabetical order, separated by semicolon ( ; )**

## I. INTRODUCTION

This template, modified in MS Word 2007 and saved as a Word 97-2003 document, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. Margins, column widths, line spacing, and type styles are built-in here. The authors must make sure that their paper has fulfilled all the formatting stated here.

Introduction commonly contains the background, purpose of the research, problem identification, and research methodology conducted by the authors which been describe implicitly. Except for Introduction and Conclusion, other chapter's title must be explicitly represent the content of the chapter.

## II. EASE OF USE

### *A. Selecting a Template*

First, confirm that you have the correct template for your paper size. This template is for ULTIMATICS. It has been tailored for output on the A4 paper size.

### *B. Maintaining the Integrity of the Specifications*

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them.

## III. PREPARE YOUR PAPER BEFORE STYLING

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not add any kind of pagination anywhere in the paper. Please take note of the following items when proofreading spelling and grammar.

### *A. Abbreviations and Acronyms*

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Abbreviations that incorporate periods should not have spaces: write "C.N.R.S.," not "C. N. R. S." Do not use abbreviations in the title or heads unless they are unavoidable.

### *B. Units*

- Use either SI (MKS) or CGS as primary units (SI units are encouraged).
- Do not mix complete spellings and abbreviations of units: "Wb/m2" or "webers per square meter," not "webers/m2." Spell units when they appear in text: "...a few henries," not "...a few H."
- Use a zero before decimal points: "0.25," not ".25."

### *C. Equations*

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number the equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop.

$$\int_0^{r_2} F(r,\phi)\, dr\, d\phi = [\sigma r_2 / (2\mu_0)] \qquad (1)$$

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the

equation. Use "(1)," not "Eq. (1)" or "equation (1)," except at the beginning of a sentence: "Equation (1) is ..."

### D. Some Common Mistakes

- The word "data" is plural, not singular.
- The subscript for the permeability of vacuum $\mu_0$, and other common scientific constants, is zero with subscript formatting, not a lowercase letter "o."
- In American English, commas, semi-/colons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an "inset," not an "insert." The word alternatively is preferred to the word "alternately" (unless you really mean something that alternates).
- Do not use the word "essentially" to mean "approximately" or "effectively."
- In your paper title, if the words "that uses" can accurately replace the word using, capitalize the "u"; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones "affect" and "effect," "complement" and "compliment," "discreet" and "discrete," "principal" and "principle."
- Do not confuse "imply" and "infer."
- The prefix "non" is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the "et" in the Latin abbreviation "et al."
- The abbreviation "i.e." means "that is," and the abbreviation "e.g." means "for example."

## IV. Using the Template

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention as below

ULTIMATICS_firstAuthorName_paperTitle.

In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper. Please take note on the following items.

### A. Authors and Affiliations

The template is designed so that author affiliations are not repeated each time for multiple authors of the same affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### B. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include ACKNOWLEDGMENTS and REFERENCES, and for these, the correct style to use is "Heading 5."

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles, named "Heading 1," "Heading 2," "Heading 3," and "Heading 4", are prescribed.

### C. Figures and Tables

Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation "Fig. 1," even at the beginning of a sentence.

TABLE I. TABLE STYLES

| **Table Head** | **Table Column Head** | | |
|---|---|---|---|
| | ***Table column subhead*** | ***Subhead*** | ***Subhead*** |
| copy | More table copy | | |

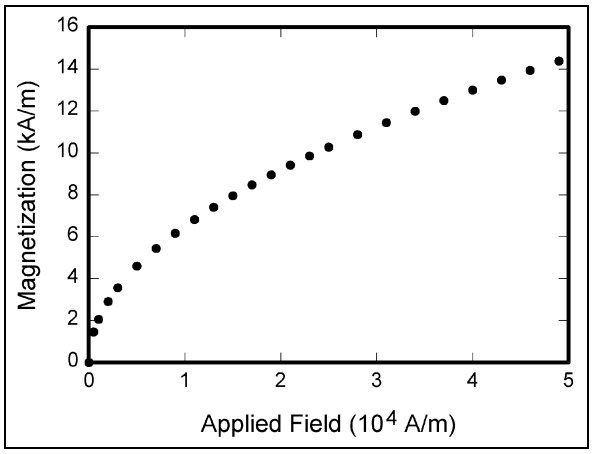


Fig. 1. Example of a figure caption

## V. Conclusion

A conclusion section is not required. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

## Appendix

Appendixes, if needed, appear before the acknowledgment.

## ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Avoid expressions such as "One of us (S.B.A.) would like to thank ... ." Instead, write "F. A. Author thanks ... ." You could also state the sponsor and financial support acknowledgments here.

## References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

[1] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (*references*)

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[3] I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

ISSN 2085-4552



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