BAREKENG: Journal of Mathematics and Its Applications
March 2023 Volume 17 Issue 1 Page 0323–0330
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss1pp0323-0330

# LOGISTIC MODELING TO PREDICT THE INTEREST OF THE INDONESIAN PEOPLE FOR BUYING FLOOD IMPACTED INSURANCE PRODUCTS

**Yulial Hikmah***

*Vocational Education Program, Universitas Indonesia*
*Jl. Akses Vokasi UI, Kukusan, Kecamatan Beji, Kota Depok, 16424, Jawa Barat, Indonesia*

*Corresponding author's e-mail:* yulialhikmah47@ui.ac.id*

***Article History:***

*Received: 9th October 2022*
*Revised: 27th December 2022*
*Accepted: 2nd February 2023*

***Keywords:***

*Flood;*
*Risk Mitigation;*
*Logistic Regression Model;*
*Insurance Product*

***ABSTRACT***

*Indonesia is a country located on the equator and in the form of an archipelago. It has a high potential for various types of hydrometeorological-related disasters, such as floods, flash floods, droughts, extreme weather, etc. Almost all cities in Indonesia experience flooding every year, including DKI Jakarta, the capital city of Indonesia. Based on data from the National Disaster Management Agency (BNPB) in 2020, East Jakarta is a city that is prone to flooding. According to BNPB (2013), flooding is a disaster that relatively causes the most losses. Losses caused by floods, especially indirect losses, may rank first or second after an earthquake or tsunami. Floods cause so many losses, and it is necessary to have disaster mitigation efforts to minimize the possibility of flood risks. One risk mitigation due to natural disasters is buying insurance products. However, not everyone buys flood-impact insurance products due to economic and social factors. This study aims to create a model with the Logistics Regression Model to determine the factors influencing Indonesian people's interest in purchasing flood-impact insurance products. The research data is from 140 households in East Jakarta, Indonesia, using a non-probability purposive sampling technique. Furthermore, with a significance level of 10%, the logistic regression model obtained 14 significant regression coefficients. In the end, the obtained model is evaluated based on its level of accuracy. The results showed that the accuracy rate was almost excellent, namely 89.3%.*



***How to cite this article:***

Y. Hikmah., "LOGISTIC MODELING TO PREDICT THE INTEREST OF THE INDONESIAN PEOPLE FOR BUYING FLOOD IMPACTED INSURANCE PRODUCTS," *BAREKENG: J. Math. & App.*, vol. 17, iss. 1, pp. 0323-0330, March 2023.


Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

**Research Article • Open Access**

## 1. INTRODUCTION

The position of the Territory of the Unitary State of the Republic of Indonesia, which is located on the equator and in the form of an archipelago, creates a high potential for various types of disasters related to hydrometeorology, such as floods, flash floods, droughts, extreme weather (tornados), powerful waves and abrasion as well as land and forest fires. The phenomenon of climate change is also increasing the threat of hydro-meteorological disasters [1]. A flood is when water inundates an area that is usually not flooded for a certain period. Floods typically occur because rainfall continues to fall and results in an overflow of rivers, lakes, seas, or drainages because the amount of water exceeds the capacity of the water-supporting media from earlier rainfall. Besides being caused by natural factors, namely high rainfall, floods also occur due to human activities [2]. Based on the Flood Disaster Risk Study, 10 (ten) priority provinces focus on flood disaster management activities, one of which is DKI Jakarta Province [1].

Geomorphologically, most Jakarta area is formed by fluvial processes and is located in the northern part of Java Island. This condition causes the DKI Jakarta area prone to hydro-meteorological hazards, especially floods. These conditions show that DKI Jakarta Province is an area with a high threat of flooding [3]. Reporting to the CNN Indonesia website (2019), one of the major floods in Jakarta had existed since the 1600s when Jan Pieterszoon Coen served as Governor-General of the VOC [4]. In early 2020, precisely on January 1, 2020, flash floods and landslides hit DKI Jakarta, Bogor, Depok, Tangerang, Lebak, and Bekasi, which lasted for several days (Kompas, January 2, 2020) [5]. The flood disaster that occurred in early 2020 was one of the enormous floods ever in Jakarta. BMKG noted that the flood disaster in Jakarta occurred in 39 urban villages, and more than 11,000 people were evacuated [4].

A flood is also a disaster that relatively causes the most losses. Losses caused by floods, especially indirect losses, may rank first or second after the earthquake or tsunami [5]. Flood damage can be divided into four types: direct tangible damage (e.g., physical damage from contact with water), indirect (e.g., loss of production and income), direct intangible (e.g., loss of life), and indirect (e.g., trauma) [6]. Losses due to flooding can be in the form of damage to buildings, loss of valuables, and losses that result in not going to work and school. Floods cannot be prevented, but they can be controlled, and the impact of losses can be reduced [7]. Therefore, flood mitigation and control efforts must be carried out to reduce disaster risk and minimize material losses and fatalities. According to BNPB in the Handbook of Responding to Disasters (2017), one of the steps for the community in pre-disaster in dealing with flood disasters is to consider flood insurance [2].

Insurance is an insurance transaction that involves two parties, namely the insured and the insurer. The insurer guarantees the insured party that the insured will get compensation for a loss that may be suffered due to an event that was not necessarily going to occur or which at first could not be determined when it happened [6]. One of the insurance losses due to flooding is property insurance, a type of protection for property assets such as houses, apartments, and offices. The goal is to anticipate financial losses due to unexpected events that occur in property damage. These unexpected events can be catastrophic. However, not everyone buys flood-impact insurance products due to economic and social factors. In 2022, a similar study was conducted, although the variables were not precisely the same, and the model used was the probit regression model. The study's results showed that the accuracy obtained was 84.3%, PPV 27.3%, and NPV 89.1% [8].

This study aims to determine the social and economic factors that influence the interest of the Indonesian people in buying flood-impact insurance products. Various methods can be used to determine the factors that influence the community's interest, especially those who experience flooding and are affected by floods, to buy property insurance products due to flooding. The logistic model is one of the statistical probability models that can be used when the dependent variable is a variable with two categories. The logistic regression model, also called the nonlinear regression model, analyzes the relationship between a dependent variable (response variable) and several independent variables. The response variable is dichotomous qualitative data, namely 0 and 1. Logistics Regression can determine the factors influencing Indonesian people's interest in buying flood insurance products.

## 2. RESEARCH METHODS

### 2.1 Research Samples

The sample in this study was 140 householders in the city of East Jakarta, Indonesia. This study uses a non-probability purposive sampling technique, which is a sampling technique with specific considerations [9]. Based on the BNPB (National Disaster Management Agency) report, East Jakarta City was the worst affected by floods [10], [11]. Furthermore, the Head of the BNPB Disaster Data, Information, and Communication Center, Agus Wibowo, said eight sub-districts in East Jakarta with the highest number of flood victims, namely 752 families (2476 people) [12]. Due to limited time, cost, and resources, this research selected several Neighborhood Units that did and are still experiencing flooding in early 2020. The Neighborhood Units selection was based on news articles in some mass media [13]–[15].

### 2.2 Method of Collecting Data

The number of research samples was 140 families. Data were collected through interviews and also by filling out online questionnaires. The reason for conducting interviews is that several questions must be given directly to see the truth and validity of the respondent's answers. However, although it was completed with an interview, this research still complies with the health protocol (covid-19 pandemic). The data collection process through interviews and online filling.

### 2.3 Research Questionnaire

A research instrument is a tool used to measure the observed natural and social phenomena 28. One of the research instruments is a questionnaire [16]. The questionnaire is a data collection technique that gives respondents questions or written statements to answer [17]. To find out the factors that influence the interest of the Indonesian population to buy property insurance products due to floods, the researchers formulated several questions which were research variables and were built using Google Form (online questionnaire). Research variables are everything the researcher determines to be studied so that information is obtained 30. Twenty-nine questions become independent variables in this study, while the dependent variable is the respondent's interest in buying flood-impact insurance products (Y). The list of questionnaire questions can be seen in Table 1.

**Tabel 1. Questionnaire Questions**

| No | Questions | Variables |
|---|---|---|
| 1. | Respondent's interest in buying flood impact insurance products (No, Yes) | Y |
| 2. | Respondents' Marital Status (Single, Married, Divorced) | X1 |
| 3. | Respondents' House Status (Contract/Lease, Family-Owned, Private) | X2 |
| 4. | Age of Householder Head | X3 |
| 5. | Last Education Level of Householder Head (No School, Primary School, Junior High School, Senior High School, D3, Bachelor Degree, Master/Doctoral) | X4 |
| 6. | Monthly Income Type (IDR) (0, 0-2 million, 2-4 million, 4-6 million, 6-8 million, 8-10 million, 10-12 million, >12 million) | X5 |
| 7. | Average Monthly Expenses | X6 |
| 8. | The number of dependents | X7 |
| 9. | Flood Experienced (No, Yes) | X8 |
| 10. | Has there been a flood in the past year? (No, Yes) | X9 |
| 11. | How many times have you experienced floods? | X10 |
| 12. | Flood Height (cm) | X11 |
| 13. | Is your house near a river? (No, Yes) | X12 |
| 14. | How long does the flood take until it recedes (days)? | X13 |
| 15. | Have you ever been evacuated? (No, Yes) | X14 |
| 16. | How many times do you evacuate each year? | X15 |
| 17. | Do you always evacuate if you experience a flood? (No, Yes) | X16 |
| 18. | Has your family member ever died in a flood? (No, Yes) | X17 |
| 19. | Has a flood ever caused damage to your property? (No, Yes) | X18 |

| No | Questions | Variables |
|---|---|---|
| 20. | How much is the loss of property damage due to flooding (in IDR)? | X19 |
| 21. | Do you know about insurance? (No, Yes) | X20 |
| 22. | Do you agree with the following statement: Insurance is an agreement between the insured and the insurer that requires the insured to pay a premium to provide reimbursement for the risks that will occur? (No, Yes) | X21 |
| 23. | Do you agree with the following statement: The function of insurance is to control the risk that will occur? (No, Yes) | X22 |
| 24. | Do you agree with the following statement: Insurance is one way to minimize losses from loss or damage to valuable objects due to flooding? (No, Yes) | X23 |
| 25. | Do you know the types of insurance? (No, Yes) | X24 |
| 26. | Do you know some insurance companies in Indonesia? (No, Yes) | X25 |
| 27. | Do you know the purpose of insurance? (No, Yes) | X26 |
| 28. | Do you know that insurance products can be purchased to cover the losses you will receive from flooding each year? (No, Yes) | X27 |
| 29. | Do you know how to buy insurance products? (No, Yes) | X28 |
| 30. | Does Your Family Have Insurance? (No, Yes) | X29 |

### 2.4 Data Analysis Method

One of the analytical methods to obtain the factors that influence the interest of the Indonesian population to buy flood impact insurance products is modeling. To predict a problem, the general model is a linear regression model. Logistic regression can be considered a particular linear regression case when the predicted is categorical. Logistics Regression is a statistical model capable of estimating the probability of an event occurring. Some of the data of interest in a regression study may be ordinal or nominal in scale. Because regression analysis requires numerical data, coding is carried out on variables **[18]**.

Let 1 represent when the event occurs, and 0 define when the event does not happen. Therefore, $P(1)$ is the probability of the event and $P(0) = 1 - P(1)$. Suppose the occurrence of an event depends on the independent variables $X_1, X_2, \dots, X_n$. In Logistic Regression, the probability log is a linear function of $X_i$ with $i = 1, \dots, n$ which is described in the following equation: **[19]**

$$\ln\left(\frac{P(1)}{1-P(1)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (1)$$

If **Equation (1)** is solved, then we get:

$$P(1) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \quad (2)$$

where $\beta_i$ with $i = 0,1, \dots, n$ is the coefficient of logistic regression **[18]**.

In logistic regression, it is necessary to carry out several tests, including the Omnibus test, which is a test of the significance of the effect of all independent variables on the dependent variable simultaneously, and the Hosmer and Lemeshow test, which is a Goodness of Fit test, to determine whether the model formed is correct or not. The model is appropriate if there is no significant difference between the model and the observed values in this study. The T-test is used to test the significance of the effect of each variable partially. **[20]**.

The dependent variable code in this study is 0 for not interested and 1 for interested in buying flood impact insurance products. After the model is formed, it is evaluated. This study measures the results of the evaluation of the logistic regression model based on the level of accuracy, PPV, and NPV. The formula for the level of accuracy, PPV, and NPV can be calculated based on the results of the confusion matrix as follows: **[21]**, **[22]**

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$PPV = \frac{TP}{TP+FP} \quad (4)$$

$$NPV = \frac{TN}{TN+FN} \quad (5)$$

The three measurements can be calculated after obtaining the confusion matrix first. The confusion matrix can be interpreted to analyze whether the classifier recognizes tuples from different classes. True Positive (TP) and True Negative (TN) values provide information when classifying data is correct. In

contrast, False Positive (FP) and False Negative (FN) provide information when the classifier is wrong in categorizing data [21]. This study uses IBM SPSS Statistics 21 software for logistic regression analysis and Microsoft Excel to calculate the level of accuracy, PPV, and NPV.

## 3. RESULTS AND DISCUSSION

The data analysis results in a summary of the data and overall model testing using the IBM SPSS Statistics 21 software can be seen in Tables 2 to 5. Table 2 summarizes the output of the number of samples, which is 140 samples, and there are no missing cases. Table 3 is an Omnibus test, a simultaneous test of all independent variables on the dependent variable. It can be seen that the significance value is 0.005. This study uses a significance level of 10%, so it can be said that the independent variable can significantly influence the model (fit).

**Table 2. Data Summary**

| Cases | N | Percent |
|---|---|---|
| Data | 140 | 100 |
| Missing | 0 | 0 |
| Total | 140 | 100 |

**Table 3. Omnibus Test Results**

| | Chi-square | df | Sig. |
|---|---|---|---|
| Model | 68.286 | 41 | 0.005 |

Table 4 shows the ability of the independent variable to explain the dependent variable. It can be seen that the value of Nagelkerke R Square is 0.581 and Cox & Snell R Square is 0.386. This shows that the dependent variable (interest in purchasing flood impact insurance products) can be explained by 29 independent variables in this study of 58.1%. In contrast, the remaining 41.9% is explained by other independent variables outside the research model.

**Table 4. Model Summary**

| | Nagelkerke R Square |
|---|---|
| Model | 0.581 |

Table 5 is the output of the Hosmer & Lemeshow test results. The test is a Goodness of Fit test to determine whether the model formed is correct or not. The model is appropriate if there is no significant difference between the model and the observed values in this study. Because the significance value is 0.393, with a significance level of 10%, it can be said that the model can be accepted, and hypothesis testing can be carried out because there is a significant difference between the model and the observed values in this research.

**Table 5. Hosmer and Lemeshow Test Results**

| | Chi-square | df | Sig. |
|---|---|---|---|
| Model | 8.426 | 8 | 0.393 |

Furthermore, parameter estimation will be carried out to form a logistic regression model. The parameter estimation results can be seen in Table 6. In Table 6, each variable's significance test results can be obtained. It shows that with a significance level of 10%, there are 14 significant regression coefficients, namely X1 (Married), X2 (Contract/Lease and Family Owned), X7 (The number of dependents), X8 (Has experienced a flood), X10 (Frequency of flood experience), X11 (Flood height), X12 (Distance to the river), X13 (Flood receding time), X18 (Property Loss due to flooding), X20 & X21 (Insurance knowledge), X27 (Flood-impacted insurance product knowledge), and X29 (Family participation in insurance).

**Table 6. Parameter estimation**

| Variables | B | df | Sig. |
|---|---|---|---|
| X1 | | 2 | .642 |
| X1(1) | -1.230 | 1 | .039 |
| X1(2) | .652 | 1 | .678 |
| X2 | | 2 | .010 |
| X2(1) | -.849 | 1 | .095 |
| X2(2) | .935 | 1 | .277 |
| X3 | .000 | 1 | .991 |
| X4 | | 6 | .374 |
| X4(1) | 20.352 | 1 | .999 |
| X4(2) | 20.115 | 1 | .999 |
| X4(3) | 19.263 | 1 | .999 |
| X4(4) | 18.442 | 1 | .999 |
| X4(5) | 22.523 | 1 | .999 |
| X4(6) | 18.661 | 1 | .999 |
| X5 | | 6 | .985 |
| X5(1) | -53.485 | 1 | .999 |
| X5(2) | -54.788 | 1 | .999 |
| X5(3) | -54.415 | 1 | .999 |
| X5(4) | -54.439 | 1 | .999 |
| X5(5) | -74.987 | 1 | .999 |
| X5(6) | -30.548 | 1 | 1.000 |
| X6 | .000 | 1 | .112 |
| X7 | .415 | 1 | .053 |
| X8(1) | 18.974 | 1 | .009 |
| X9(1) | 17.964 | 1 | .219 |
| X10 | .286 | 1 | .019 |
| X11 | .002 | 1 | .080 |
| X12(1) | 3.326 | 1 | .015 |
| X13 | -.655 | 1 | .052 |
| X14(1) | -.250 | 1 | .784 |
| X15 | -.313 | 1 | .423 |
| X16(1) | .826 | 1 | .449 |
| X17(1) | 16.678 | 1 | 1.000 |
| X18(1) | 2.139 | 1 | .025 |
| X19 | .000 | 1 | .398 |
| X20(1) | .267 | 1 | .074 |
| X21(1) | -1.878 | 1 | .096 |
| X22(1) | -.739 | 1 | .498 |
| X23(1) | -.777 | 1 | .437 |
| X24(1) | -.934 | 1 | .306 |
| X25(1) | .795 | 1 | .369 |
| X26(1) | .409 | 1 | .691 |
| X27(1) | -2.144 | 1 | .041 |

| Variables | B | df | Sig. |
|---|---|---|---|
| X28(1) | -.522 | 1 | .547 |
| X29(1) | -2.602 | 1 | .008 |
| Constant | 20.113 | 1 | 1.000 |

Next is to create a confusion matrix and calculate the level of accuracy, PPV, and NPV. The confusion matrix of this study can be seen in **Table 7**. Based on **Table 7**, TP, FP, FN, and TN values are 22, 4, 11, and 103, respectively.

**Table 7. Parameter estimation**

| | | Predicted Class | |
|---|---|---|---|
| | | **Interest** | **Not Interest** |
| **Actual Class** | **Interest** | 22 | 11 |
| | **Not Interest** | 4 | 103 |

The following is the result of calculating the level of accuracy:

$$Accuracy = \frac{22 + 103}{22 + 103 + 4 + 11} = \frac{125}{140} = 89.3\%$$

Therefore, the logistic regression model has an accuracy level in the good classification category, even approaching the excellent classification category. The following are the results of PPV and NPV calculations:

$$PPV = \frac{22}{22 + 4} = \frac{22}{26} = 84.6\%$$

$$NPV = \frac{103}{103 + 11} = \frac{103}{114} = 90.4\%$$

## 4. CONCLUSIONS

In this study, with a significance level of 10%, the logistic regression model obtained was appropriate, and the independent variables explained the dependent variable well. In addition, it was also found that 14 significant regression coefficients, namely X1 (Married), X2 (Contract/Lease and Family Owned), X7 (The number of dependents), X8 (Has experienced a flood), X10 (Frequency of flood experience), X11 (Flood height), X12 (Distance to the river), X13 (Flood receding time), X18 (Property Loss due to flooding), X20 & X21 (Insurance knowledge), X27 (Flood-impacted insurance product knowledge), and X29 (Family participation in insurance). The logistic regression model is then evaluated based on its level of accuracy. The results show that the level of accuracy is almost excellent, namely 89.3%. In addition, this study also obtained PPV and NPV measures which were valued at 84.6% and 90.4%, respectively. The evaluation results from applying the logistic regression model with 29 independent variables in this study are better than previous studies involving 19 independent variables using the probit regression model.

## REFERENCES

[1] BNPB RI, "Rencana Nasional Penanggulangan Bencana 2015-2019," 2014. [Online]. Available: https://www.bnpb.go.id//uploads/renas/1/BUKU RENAS PB.pdf.

[2] BNPB RI, "Buku Saku Tanggap Tangkas Tangguh Menghadapi Bencana," 2017. [Online]. Available: https://siaga.bnpb.go.id/hkb/po-content/uploads/documents/Buku_Saku-10Jan18_FA.pdf.

[3] S. Dahlia, N. H. Tricahyono, and W. F. Rosyidin, "Analisis Kerawanan Banjir Mengunakan Pendekatan Geomorfologi di DKI Jakarta," *J. Alami J. Teknol. Reduksi Risiko Bencana*, vol. 2, no. 1, pp. 1–8, 2018.

[4] E. Y. Gunawibawa and H. Oktiani, "Politik & Bencana Banjir Jakarta 2020: Analisis Peta Percapakan #JakartaBanjir," *Expo. J. Ilmu Komun.*, vol. 3, no. 1, pp. 60–75, 2020, doi: https://doi.org/10.33021/exp.v3i1.989.

[5] A. Rosyidie, "Banjir: Fakta dan Dampaknya, serta Pengaruh dari Perubahan Guna Lahan," *J. Reg. City Plan.*, vol. 24, no. 3, pp. 241–249, 2013, doi: https://doi.org/10.5614/jpwk.2013.24.3.1.

[6] P. Sidi, *Pemanfaatan Ilmu Aktuaria dalam Mewujudkan Jaminan Risiko Banjir di dalam Konsep Smart City*. Universitas Terbuka (UT), 2017.

[7] A. Findayani, "Kesiap Siagaan Masyarakat dalam Penanggulangan Banjir di Kota Semarang," *J. Geogr. Media Inf.*

*Pengemb. dan Profesi Kegeografian*, vol. 2, no. 1, pp. 102–114, 2015, doi: https://doi.org/10.15294/jg.v12i1.8019.
[8] Y. Hikmah, V. Yuristamanda, I. R. Hikmah, and K. A. Safitri, "Probit Modeling of Indonesian Economic and Social Factors to the Interest in Purchasing Flood-Impacted Insurance Products," *Int. J. Ind. Eng. Prod. Res.*, vol. 33, no. 2, pp. 1–12, 2022, doi: 10.22068/ijiepr.33.2.13.
[9] Sugiyono, *Memahami Penelitian Kualitatif*. Bandung: Alfabeta, 2014.
[10] BNPB RI, "Hujan Lebat Sebabkan 23 Kecamatan DKI Jakarta Terdampak Banjir," 2020. https://bnpb.go.id/berita/hujan-lebat-sebabkan-23-kecamatan-dki-jakarta-terdampak-banjir (accessed Jun. 01, 2020).
[11] K. Nisa, "Rekapitulasi Data Banjir DKI Jakarta dan Penanggulangannya Tahun 2020," 2020. https://statistik.jakarta.go.id/rekapitulasi-data-banjir-dki-jakarta-dan-penanggulangannya-tahun-2020/ (accessed Jun. 01, 2020).
[12] CNN Indonesia, "BNPB Sebut Banjir Rendam 23 Kecamatan di DKI, Jaktim Terparah," Feb. 08, 2020. https://www.cnnindonesia.com/nasional/20200208215738-20-472935/bnpb-sebut-banjir-rendam-23-kecamatan-di-dki-jaktim-terparah (accessed Jun. 01, 2020).
[13] I. Komara, "Warga di 10 Lokasi Jakarta Timur Dievakuasi dari Banjir," *detikNews*, Jan. 01, 2020. https://news.detik.com/berita/d-4842307/warga-di-10-lokasi-jakarta-timur-dievakuasi-dari-banjir (accessed Jun. 01, 2020).
[14] Kumparan News, "Titik Banjir Jakarta Bertambah, Berikut Daftar 46 RW yang Terendam," 2020. https://kumparan.com/kumparannews/titik-banjir-jakarta-bertambah-berikut-daftar-46-rw-yang-terendam-1ssTJoGx9Qv/3 (accessed Jun. 01, 2020).
[15] N. P. Putra, "Cek Titik Banjir di Ruas Jalan Jakarta Timur," *Liputan6*, 2020. https://www.liputan6.com/news/read/4186976/cek-titik-banjir-di-ruas-jalan-jakarta-timur (accessed Jun. 01, 2020).
[16] Syahraeni, "Analisis Tingkat Pemahaman Mahasiswa Jurusan Ilmu Perpustakaan Fakultas Adab dan Humaniora UIN Alauddin Makassar terhadap Sistem Klasifikas DDC," UIN Alauddin Makassar, 2016.
[17] Sugiyono, *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: Alfabeta, 2018.
[18] D. L. Olson and D. Wu, *Predictive Data Mining Models*, 2nd ed. Singapore: Springer, 2020.
[19] H. Zhou, *Learn Data Mining Through Excel: A Step-by-Step Approach for Understanding Machine Learning Methods*. USA: Apress, 2020.
[20] D. Hosmer Jr, S. Lemeshow, and R. Sturdivant, *Applied Logistic Regression*, 3rd ed. New York: John Wiley & Sons, 2013.
[21] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, "Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking," *IEEE Access*, vol. 8, pp. 90847–90861, 2020, doi: 10.1109/ACCESS.2020.2994222.
[22] M. Stojanoviü *et al.*, "Understanding Sensitivity, Specificity, and Preductive Values," *Vojnosanit. Pregl.*, vol. 71, no. 11, pp. 1062–1065, 2014.