EKSAKTA
Journal of Sciences and Data Analysis

E-ISSN: 2720-9326
P-ISSN: 2716-0459
DOI : 10.20885/EKSAKTA.vol7.iss1.art9

Inflation Convergence Modeling Using Binary Logistic Regression With SGD-Newton Raphson Optimization Methods in Indonesia

Fatma Novalia Kussumarani[1], Nerissabila Uswatun Istiqomah[2], Siva Ifin Azzahra[3], Anggraini Puspita Sari[4*], Sischa Wahyuning Tyas[5]

[1,2,3,5] Data Science Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional "Veteran" East Java, Indonesia
[4*] Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional "Veteran" East Java, Indonesia

[*]*Corresponding author: anggraini.puspita.if@upnjatim.ac.id*

Abstract: Global economic changes have necessitated the development of inflation models that can accurately describe Indonesia's economic dynamics. This study aims to compare two optimization methods, Newton Raphson and Stochastic Gradient Descent (SGD), in binary logistic regression modeling to analyze the effectiveness of monetary policy. This study contributes to evaluating the performance of both methods in terms of convergence speed and accuracy of inflation model parameter estimation. The results of the analysis show that the Newton Raphson method is more efficient in achieving convergence with an iteration value of 0.2933 compared to SGD, while both methods produce equivalent model quality based on the Akaike Information Criterion (AIC) values of 34.4008 and 34.4254. These findings emphasize the importance of selecting the right optimization method to support more efficient monetary policy analysis.

Keywords: Inflation, binary logistic regression, Newton–Raphson, Akaike Information Criterion (AIC)

Introduction

In 2021, the pandemic caused global economic instability that affected the economies of all countries, especially rapidly developing countries such as Indonesia. The World Bank described Indonesia's -pandemic economic recovery as triggering a rise in commodity prices. In April 2022, CPI inflation rose from 1.6% to 3.5%, driven by global oil prices, which caused food prices to increase. Fuel subsidies limited energy transmission with inflation expectations remaining within the BI target [1].

Inflation is an important macroeconomic indicator in determining the stability of a country's nvarious macro variables such as interest rates, oil prices, and exchange rates. Recent empirical evidence shows that oil price shocks are one of the main determinants of inflation in the ASEAN region, including Indonesia. Previous research found that oil price shocks significantly drive inflationary pressures, especially when triggered by changes in global demand, thereby directly impacting the CPI, the national economy, and stock returns through channels inflation and exchange rates [3][4][5]. Global volatility has been shown to increase inflation forecasting errors, making traditional prediction models less accurate [6]. Furthermore, research conducted by [7] states that the Newton–Raphson method can experience convergence failure, while SGD is more stable in reaching the optimum point, especially in complex data structures such as economic growth modeling [8]. The research found that there are only a few studies or research that evaluate the convergence of inflation models using a combination of the Newton Rhapson method, Stochastic Gradient Descent (SGD), and logistic regression.

This study aims to estimate and analyze the equilibrium of the inflation model in Indonesia for the period 2021 to 2025 using dependent variables in economic dynamics, namely interest rates, world oil prices, and exchange rates. The Newton Raphson method is used to find the equilibrium point, followed by parameter optimization using the Stochastic Gradient Descent (SGD) method to ensure the convergence and stability of the model. Logistic regression is used to calculate probabilities based on dependent variables. The results will be used to assist in formulating policies and evaluating the impact of external shocks on inflation control in the future.

The novelty of this research is an innovation that applies a probabilistic approach that provides a new perspective in assessing changes in inflation status. Testing the estimation limits was carried out by integrating the two optimization methods, Newton Raphson and SGD, to produce a consistent model. Thus, this research can provide methodological contributions and enrich practical insights into the role of monetary policy in maintaining inflation stability amid uncertainty.

Materials And Methods

2.1 Research Process

This research was conducted in several stages, beginning with data collection, followed by the separation of data into training data and testing data, then modeling using binary logistic regression, followed by iterative modeling using the Newton-Raphson and SGD methods, and finally evaluating the model performance from the test data to obtain values and compare the predictive capabilities of each method, as summarized in the flowchart shown in Figure 1.

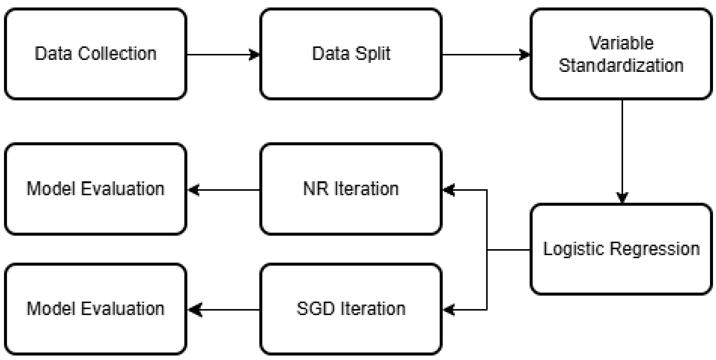

Figure 1 Research Process Flowchart

2.2 Data Types and Sources

This study uses secondary data in the form of monthly time series for the period 2021-2025 with a total of 57 observations. Inflation data was obtained from the official website of the Central Statistics Agency (BPS), interest rate data was obtained from BI Rate/BI7DRR, rupiah exchange rate data was obtained from Federal Reserve Economic Data (FRED), and world oil prices were taken from the Energy Information Administration (EIA) database.

Table 2.1 Data on Inflation, Interest Rates, World Oil Prices, and Rupiah Exchange Rates for the period 2021–2025

Year	Month	Inflation (YoY)	Interest Rate	Europe Brent Spot Price FOB	RBIDBIS
2021	January	1.55%	3.75	54.77	99.12
2021	February	1.38%	3.50%	62.28	98.93
2021	March	1.37%	3.50%	65.41	97.87
...	...		...	...	...
2025	August	2.31%	5.00%	67.87	95.38
2025	September	2.65%	4.75%	67.99	93.84

Table 2.1 is a collection of data used in research during a certain period. The data needs to be cleaned by removing spaces and percentages (%), then standardized on the predictor variables.

2.3 Variable Transformation and Standardization

Standardization is used if the data obtained has a wide range of values, as the data grouping process can become complicated and the data becomes invalid, and the parameters cannot calculate the distance between data points [9]. Standardization aims to improve model stability and performance so that the data scale is more centered and makes it easier to interpret coefficient values. This standardization process is carried out so that the variables used are on the same scale, there by facilitating a fast and accurate optimization process. Variable standardization is performed using model estimation, and all variables are standardized using the Z-score formula.

$$Z = \frac{x-\mu}{\sigma} \tag{1}$$

2.4 Binary Logistic Regression & Odds Ratio

Binary logistic regression is a regression where the dependent variable consists of two categories, namely Y=1 or Y=0 [10]. This study uses inflation as the dependent variable in binary logistic regression with the following definition.

$$Y = \begin{cases} 1 \text{ } if \text{ } inflation \text{ } rises \\ 0 \text{ } if \text{ } inflation \text{ } decreases \end{cases}$$

The general equation in the binary logistic regression model can be written as follows.

$$\log\left(\frac{P(Y = 1|X)}{1-P(Y = 1|X)}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k \quad (2)$$

In this study, the binary logistic regression model in equation 2 will be used to estimate the probability of inflation based on predictor variables, namely interest rates, oil prices, and exchange rates. Based on this equation, the binary logistic regression model estimates the logit as the probability of success as a linear function of the predictor variable [11]. The binary logistic regression model formula with the logit function.

$$\ln\left(\frac{\pi(xi)}{1-\pi(xi)}\right) = g(xi) \quad (3)$$

Next, the logit function in equation 3 is used to calculate the observed probability value, and this calculation then forms the basis for calculating the odds ratio. Odds Ratio (OR) is a model used in binary regression to interpret parameter coefficients [12]. An OR value > 1 indicates an increase in the probability of an event, while an OR < 1 indicates a decrease in the probability of an event [13]. General formula for Odds Ratio.

$$OR_i = e^{\beta_i} \quad (4)$$

2.5 Newton-Raphson Method

Newton-Raphson is an iterative optimization method that updates parameters using first and second derivatives, resulting in faster and more accurate convergence in non-linear models, including logistic regression [14]. In this study, the Newton-Raphson method integrates information from the first derivative and the Hessian matrix (second derivative) to ensure a more precise and accurate convergence direction. The Newton-Raphson iteration formula can be written as follows.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (5)$$

In addition to using the Newton-Raphson method, this study also uses the SGD method as a comparison method that is appropriate for formulating inflation models. Stochastic Gradient Descent (SGD) is an effective optimization method because it updates parameters using random samples from the data, making it more efficient than full gradient methods [15]. Basically, parameter changes in SGD follow a mathematical scheme:

$$\theta t + 1 = \theta t - \eta t \nabla \theta L(\theta t; xi, yi) \quad (6)$$

2.6 Hosmer-Lemeshow

The Hosmer-Lemeshow (HL) test was applied in this study to validate the goodness-of-fit of the model by comparing the actual observation frequency with the model expectation value [16]. The testing process began by sorting all probability estimates from the logistic regression model in ascending order. Next, the data is systematically partitioned into g groups. In each group, the difference between the actual number of events and the predicted probability is calculated. Statistically, this HL value is asymptotically distributed chi-square under the null hypothesis [17]. The Hosmer-Lemeshow test formula can be written as follows.

$$C = \sum_{g=1}^{G} \frac{\left(O_{gl}-E_{gl}\right)^2}{E_{gl}} + \frac{\left(O_{g0}-E_{g0}\right)^2}{E_{g0}} \tag{7}$$

2.7 Pseudo R² using McFadden

The contribution of predictor variables in explaining response variance is evaluated using McFadden's R-Square index [18]. The calculation mechanism involves comparing the log-likelihood values between the null model (model without predictors) and the full model (complete model). This ratio is used to measure the predictive power of the model, where a value close to one indicates a higher level of precision and reliability of the estimation [19]. The following is the formula for Pseudo R^2

$$R^2_{McF} = 1 - \frac{\ln(L1)}{\ln(L0)} \tag{8}$$

Meanwhile, the formula for the likelihood function value is as follows.

$$F(\theta) = \sum_{i=1}^{n} \ln f\, i(yi \mid \theta) \tag{9}$$

2.8 Wald Test

The significance of each parameter in the model is evaluated individually through the Wald test procedure. Technically, this method works by calculating the ratio of the square of the coefficient estimate to its standard error to obtain the W statistic. The result is then compared with the chi-square distribution to determine the significance value. If the resulting p-value is below the threshold of 0.05, then the related independent variable is declared to have a significant effect on the response variable partially [20]. The Wald test statistical formulation is stated as follows:

$$W = \frac{\hat{\beta}}{SE(\hat{\beta})} \tag{10}$$

2.9 Akaike Information Criterion (AIC)

The most optimal model among several alternatives is determined using the Akaike Information Criterion (AIC) metric [21]. The evaluation mechanism works by penalizing the log-likelihood value based on the number of estimated variables to minimize the risk of overfitting. Through this procedure, the model that produces the lowest AIC score is determined to be the most accurate and efficient representation of the data [22]. The AIC formulation is stated as follows:

$$AIC = -2\ln(L) + 2k \tag{11}$$

Results And Discussion

In this study, Python programming was used to calculate and visualise the inflation convergence model. The use of Python enabled the analysis process to be carried out efficiently, from parameter estimation to model performance evaluation using various optimisation methods. The results of the modelling and evaluation are presented and discussed in the following section.

3.1 Variable Standardization

The first stage of analysis in this study was to clean the data and standardise the variables to equalise the variable scales without eliminating the original distribution patterns. All independent variables were standardised to obtain the data presented in Table 3.1.

Table 3.1 Results of Data Cleaning and Standardization of Predictor Variables

Interest Rate	Europe Brent Spot Price FOB	RBIDBIS
-1.34063044	-1.25701991	-1.03780985
-1.34063044	-0.56466499	-0.06351371
-1.34063044	-1.44416285	-0.31002237
...	...	...
0.47387843	-1.28364894	-1.773423
-0.20656239	-1.02179676	-1.02179676

Table 3.1 shows the results of standardizing according to equation 1 all predictor variables with a value range of –2 to 1. Extreme values such as –1.77 and 0.70 indicate significant deviations from the standard deviation of the data. These values indicate the presence of outlier observations that have the potential to greatly influence model formation. Conversely, values clustered around the range of –1 to 0.5 indicate that the data distribution remains optimal.

3.2 Iteration Results Using the Newton Raphson Method

The second step is to optimise the first logistic regression model using the Newton-Raphson method, which aims to determine the estimated value and the optimal value maximised by the logit function. The test results show very rapid convergence, as shown in Table 3.2 below.

Table 3.2 Convergence of Newton Raphson Log-loss

Iteration	Log-loss
1	0.344269

2	0.301883
3	0.293979
4	0.293349
5	0.293343
6	0.293343
7	0.293343

Based on Table 3.2 and using equation 5, the Newton–Raphson log-loss value decreases rapidly and converges to 0.293343 from the fifth iteration onwards, indicating that the model parameters have been optimised. This rapid convergence indicates that the relationship between the convergence status of regional inflation and its predictor variables such as GDP growth, interest rates, and exchange rates is stable and well-identified. Consequently, this model can be utilised by Bank Indonesia to map provinces with inflation that converges or diverges from the national target, whilst also serving as an early warning system for inflation disparities between regions.

3.3 Iteration Results Using the Stochastic Gradient Descent (SGD) Method

The third stage involves optimising the SGD method used as a comparison to Newton – Raphson. This method is tested to find optimisation solutions using random samples. The final results of the test are equivalent to the Newton-Raphson method, as presented in Table 3.3 below.

Table 3.3 Iteration Results of the SGD (*Stochastic Gradient Descent*) Method

Epoch	Log-loss
50/300	0.296465
100/300	0.294187
150/300	0.293796
200/300	0.293492
250/300	0.293591
300/300	0.293515

Based on Table 3.3, the log-loss value in the Stochastic Gradient Descent (SGD) method shows a gradual downward trend throughout the iteration process. Although there is a small increase from epoch 200/300 (0.293492) to epoch 250/300 (0.293591), this phenomenon is common because this method is stochastic and does not indicate divergence. Overall, the log-loss value decreases again and stabilizes at 0.293515, indicating that the model has reached convergence. This condition indicates that the resulting logistic regression parameters are optimal and capable of representing the inflation convergence pattern in Indonesia.

3.4 Hosmer-Lemeshow Test

After performing logistic regression model estimation, a goodness of fit test was conducted using the Hosmer-Lemeshow test. This test is used to evaluate the suitability of the model with the observed data. The following are the results of the Hosmer-Lemeshow test presented in Table 3.4 below.

Table 3.4 Hosmer-Lemeshow Test

Method	Chi-Square	df	p – value
NR	10.700519739790373	8	0.21925226918335183
SGD	10.36985469863977	8	0.2400202915597116

The results of the Hosmer-Lemeshow test using equation 7 in Table 3.4 show a Newton-Raphson p-value of 0.219 and an SGD value of 0.240, both of which are greater than the general significance level of 0.05. Therefore, the logistic model is deemed appropriate for the data and is able to accurately explain the inflation convergence pattern.

3.5 Pseudo R^2 Test Mc Fadden

The fifth step is to perform a coefficient of determination test evaluated using Pseudo R^2 based on log-likelihood comparison. This test value is used to describe the variation in the probability of events that can be explained by the predictor variables in the model, as shown in the following test result table 3.5 below.

Table 3.5 Coefficient of Determination Test

Method	LL	LLnull	Pseudo R^2
NR	-13.2004	-30.2855	0.5641
SGD	-13.2127	-30.2855	0.5637

Based on Table 3.5, it is known that both methods produce almost the same log-likelihood value, namely -13.2 with LLnull -30.2855 as the model without predictors. The pseudo R^2 values for Newton Raphson 0.5641 and SGD 0.5637 indicate that the model is able to explain approximately 56% of the log-likelihood variation compared to the null model, and the performance results of both are equivalent, with Newton Raphson being slightly better numerically. This shows that both methods can be used effectively in modeling inflation convergence patterns.

3.6 Wald Test

As a continuation of the previous maximum logit estimation and goodness-of-fit test processes, the next step in hypothesis testing is to apply the Wald test to evaluate the relative contribution of independent variables such as interest rates, world oil prices and the rupiah exchange rate to the logit model. The results of the Wald test are presented in Table 3.6 below.

Table 3.6 Parameter Estimates

Variable	Coef	Std. Error		z – value		p-value		OR	
		NR	SGD	**NR**	SGD	**NR**	SGD	**NR**	SGD
Intercept	β_0	**0.6136**	0.6146	**0.4955**	0.4619	**0.6203**	0.6441	**1.3553**	1.3197
Interest Rate	$\beta_1 X_1$	**0.6742**	0.6749	**2.2524**	2.2274	**0.0243**	0.0259	**4.5652**	4.3464
Oil Price	$\beta_2 X_2$	**0.8876**	0.8754	**1.9209**	1.9374	**0.0547**	0.0527	**5.5019**	5.3993
RBIDBIS	$\beta_3 X_3$	**0.6689**	0.6762	**1.9224**	1.9347	**0.0545**	0.0530	**3.6176**	3.5243

Based on Table 3.6, interest rates are the most influential variable on inflation with a z-value of around 2.2 and a p-value of around 0.02, as well as an odds ratio of around 4.3–4.6, making it statistically significant in increasing the likelihood of inflation. Oil prices and exchange rates show positive coefficients with z-values close to 2 and p-values around the 5% threshold, as well as odds ratios of around 5.4 and 3.5, indicating that they have a fairly strong influence on inflation increases, although their significance values are weaker than those of interest rates. The Wald test results indicate that interest rates, oil prices, and the exchange rate influence inflation dynamics in Indonesia. Therefore, controlling interest rates and stabilizing the exchange rate and energy prices are crucial for maintaining inflation convergence.

3.7 Akaike Information Criterion (AIC) Test

Based on equation 11, the AIC calculation results show that both models have relatively similar performance, with AIC values of 34.4008 and 34.4254, respectively. The difference of approximately 0.02 points indicates that the quality of the adjustment of both models is almost the same. The AIC principle states that the model with a lower value is preferred, because it provides a balance between model fit and complexity. Therefore, the Newton–Raphson model is considered more efficient and more suitable as the main reference in the analysis. Furthermore, this model is able to capture the relationship between observed economic variables and inflation more effectively, thus enabling a better understanding of inflation patterns and supporting decision-making related to inflation control.

3.8 Simulation and Visualization of Probability

As the final stage of various tests conducted, including convergence and model validity tests, a probabilistic simulation of Indonesia's inflation increase over the next six months was carried out to determine the probability of the estimated logit model. The projection results are presented in Figure 2 below.

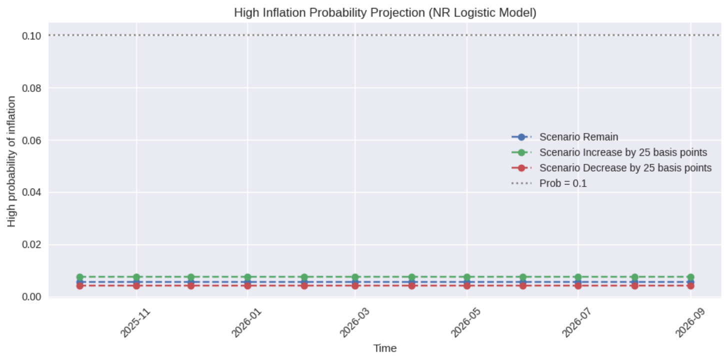

Figure 2 Visualization of High Inflation Probability Projections

Figure 2 is a visualization of inflation probability with a decision threshold or probability of 0.1. The graph illustrates the projection of high inflation probability from the Newton Raphson logistic regression model for several scenarios of fixed interest rates, a 25 bps increase, and a 25 bps decrease. It can be seen that all scenario lines are at a very low probability level, around 0.5–0.7 percent, and are relatively flat throughout the period, so that small changes in interest rate policy in this simulation do not significantly alter the likelihood of high inflation.

Conclusion

Based on the research conducted using interest rate, rupiah exchange rate, and oil price variables on the inflation variable, it was found that the Newton Raphson method converged more quickly at the seventh iteration with a value of 0.293343, while the SGD method converged at the 300th iteration with a value of 0.293515. The Wald test confirms that interest rates are the most significantly influential variable, while oil prices and exchange rates also make a fairly strong positive contribution despite being around the significance threshold, making all three the main determinants of the probability of high inflation in the model. The Pseudo R^2 value shows that both methods are able to explain around 56% of the variation in the probability of inflation in the model. The AIC values of the two approaches are of nearly equal quality, but the Newton–Raphson model with an AIC of 34.4008 is slightly better than the SGD model with an AIC of 34.4254, so Newton–Raphson is chosen as the main model because it provides the most efficient balance of fit and complexity. Further research is recommended to expand the data coverage with longer periods and higher frequencies, adding non-monetary variables such as the producer price index and the food security index. Then consider modeling with seasonal effects and explore machine learning methods that can capture *structural breaks* such as regime-switching or segmented regression models. Thus, future inflation modeling can be more accurate and relevant to support effective policy formulation.

References

[1] W. Bank, "Deepening the Financial Sector for Stronger Growth and Sustainable Recovery Executive Summary," *Indones. Econ. Prospect.*, 2022, [Online]. Available: https://www.worldbank.org/in/country/indonesia/publication/indonesia-economic-

prospects-iep-june-2022-financial-deepening-for-stronger-growth-and-sustainable-recovery
[2] Syifatul Husna, "Monetary Policy Strategies for Inflation in Indonesia," *J. Cent. Publ.*, vol. 1, no. 8, pp. 888–895, Sep. 2024, doi: 10.60145/jcp.v1i8.183.
[3] D. Y. Aharon, M. I. Azman Aziz, and I. Kallir, "Oil price shocks and inflation: A cross-national examination in the ASEAN5+3 countries," *Resour. Policy*, vol. 82, p. 103573, May 2023, doi: 10.1016/j.resourpol.2023.103573.
[4] H. Handri, H. D. Mulyaningsih, A. K. Hidayat, R. Kurniawan, and A. W. Rachmawati, "The impact of Indonesian oil price (CPI) and macroeconomics on investments in the manufacturing sector in Indonesia," *F1000Research*, vol. 10, p. 338, May 2021, doi: 10.12688/f1000research.27958.1.
[5] W. Thorbecke, "How oil prices impact the Indonesian economy: Evidence from the stock market," *Asia Glob. Econ.*, vol. 5, no. 2, p. 100122, Dec. 2025, doi: 10.1016/j.aglobe.2025.100122.
[6] A. Kliber, M. Szyszko, M. Próchniak, and A. Rutkowska, "Impact of uncertainty on inflation forecast errors in Central and Eastern European countries," *Eurasian Econ. Rev.*, vol. 13, no. 3–4, pp. 535–574, Dec. 2023, doi: 10.1007/s40822-023-00237-9.
[7] A. R. N. Fauzi, S. Abusini, and C. Karim, "Comparing Newton Raphson and Stochastic Gradient Descent Methods for Traffic Accident in Malang," *CAUCHY J. Mat. Murni dan Apl.*, vol. 10, no. 2, pp. 519–532, Jun. 2025, doi: 10.18860/cauchy.v10i2.33177.
[8] N. Lianingsih, Y. Audina, and S. Purwani, "Implementation of Newton-Raphson Iterative Method for Solving Non-Linear Equations in the Solow Economic Growth Model," *World Sci. News*, vol. 199, no. November 2024, pp. 90–98, 2025.
[9] R. Risawandi and F. Fadlisyah, "Recommendation System for Zakat Recipients at Baitul Mal Kota Lhokseumawe Using the Z-Score Method," *J. Teknol. Terap. Sains 4.0*, vol. 4, no. 3, p. 136, Nov. 2023, doi: 10.29103/tts.v4i3.14887.
[10] R. Wadui, J. S. Kekenusa, and D. Hatidja, "Binary Logistic Regression Analysis to Determine Inpatient Satisfaction with the Quality of Services at the Tobelo Regional General Hospital," *d'Cartesian*, vol. 13, no. 1, pp. 43–48, Mar. 2024, doi: 10.35799/dc.13.1.2024.52520.
[11] D. P. Handayani and K. Kismiantini, "Binary Logistic Regression Analysis of Quarter-Life Crisis Symptoms on Sleep Difficulties in Early Indonesian Adulthood," *J. Mat. Stat. dan Komputasi*, vol. 22, no. 1, pp. 61–77, Sep. 2025, doi: 10.20956/j.v22i1.44765.
[12] A. A. Putri, D. T. Salaki, and J. Titaley, "Binary Logistic Regression Model of Gastric Symptom Tendencies in Mathematics Students at the FMIPA, UNSRAT," *J. Mat. dan Apl.*, vol. 11, no. 1, pp. 39–43, 2022, [Online]. Available: https://ejournal.unsrat.ac.id/v3/index.php/decartesian/article/view/37847
[13] A. J. Scott, D. W. Hosmer, and S. Lemeshow, *Applied Logistic Regression.*, vol. 47, no. 4. 1991. doi: 10.2307/2532419.
[14] S. Bahrami and K. Amini, "An efficient two-step trust-region algorithm for exactly determined consistent systems of nonlinear equations," *J. Comput. Appl. Math.*, vol. 367, p. 112470, Mar. 2020, doi: 10.1016/j.cam.2019.112470.
[15] Y. Tian, Y. Zhang, and H. Zhang, "Recent Advances in Stochastic Gradient Descent in Deep Learning," *Mathematics*, vol. 11, no. 3, p. 682, Jan. 2023, doi: 10.3390/math11030682.
[16] D. T. Ailobhio and J. A. Ikughur, "A Review of Some Goodness-of-Fit Tests for Logistic Regression Model," *Asian J. Probab. Stat.*, vol. 26, no. 7, pp. 75–85, Jun. 2024, doi: 10.9734/ajpas/2024/v26i7631.
[17] N. Surjanovic and T. M. Loughin, "Improving the Hosmer-Lemeshow goodness-of-fit test in large models with replicated Bernoulli trials," *J. Appl. Stat.*, vol. 51, no. 7, pp. 1399–1411, May 2024, doi: 10.1080/02664763.2023.2272223.
[18] Syarifah fatimah tuz zahro and Khusnul Khuluqi, "The Effect of Audit Fees, Management

Change and Company Growth on Auditor Change," *J. Akunt. Keuang. Dan Perpajak. | E-ISSN 3063-8208*, vol. 1, no. 3, pp. 366–376, Feb. 2025, doi: 10.62379/jakp.v1i3.277.

[19] A. Miranda, S. Suyitno, and M. Fauziyah, "Modelling the Probability of River Water Pollution Using Geographically Weighted Logistic Regression Model (Case Study: River Water DO Data in East Kalimantan)," *J. Mat. Stat. dan Komputasi*, vol. 21, no. 2, pp. 408–430, Jan. 2025, doi: 10.20956/j.v21i2.40346.

[20] B. Akturk, U. Beyaztas, H. L. Shang, and A. Mandal, "Robust functional logistic regression," *Adv. Data Anal. Classif.*, vol. 19, no. 1, pp. 121–145, Mar. 2025, doi: 10.1007/s11634-023-00577-z.

[21] L. I. Harlyan, E. S. Yulianto, Y. Fitriani, and Sunardi, "Application of the Akaike Information Criterion (AIC) in Calculating the Technical Efficiency of Purse Seine Fishing in Tuban, East Java," *Mar. Fish. J. Mar. Fish. Technol. Manag.*, vol. 11, no. 2, pp. 181–188, Dec. 2021, doi: 10.29244/jmf.v11i2.38550.

[22] E. Civitelli, M. Lapucci, F. Schoen, and A. Sortino, "An effective procedure for feature subset selection in logistic regression based on information criteria," *Comput. Optim. Appl.*, vol. 80, no. 1, pp. 1–32, Sep. 2021, doi: 10.1007/s10589-021-00288-1.