JURNAL STATISTIKA DAN APLIKASINYA

e-ISSN 2620 – 8369
VOLUME 9 (ISSUE 1) JUNE 2025: 13 – 21
https://doi.org/10.21009/JSA.09102

ANALYSIS OF DIABETES MELLITUS DISEASE USING BINARY LOGISTIC REGRESSION

Mery Anistya [1*], Putroue Keumala Intan[2*], Ahmad Hanif Asyhar[3], Wika Dianita Utami[4]

[1,2,3,4] *Mathematics Department, Faculty of Science & Technology, UIN Sunan Ampel Surabaya*
Jl. Dr. Ir. H. Soekarno No. 682, Gn. Anyar, Surabaya, 60294, Jawa Timur

Corresponding author's e-mail: * **meryanistya1532@gmail.com**, **putroue@uinsa.ac.id**

Article History:
Received: January 15, 2025
Revised: June 26, 2025
Accepted: June 29, 2025
Published: June 30, 2025
Available online.

Keywords:
Diabetes, Health, Puskesmas, Binary Logistic Regression

ABSTRACT

Diabetes mellitus has become a significant global public health concern due to its increasing prevalence and associated complications. Identifying the risk factors that contribute to the incidence of diabetes is crucial for developing targeted prevention and management strategies. This study aims to identify risk factors that affect the incidence of diabetes mellitus and evaluate the accuracy of the prediction model using binary logistic regression. The research method used secondary data from 140 patients at UPT Puskesmas Teja, Pamekasan, consisting of sixty non-diabetic patients and 80 diabetic patients. The variables analyzed included age, gender, heredity, smoking habit, body mass index (BMI), blood glucose level, cholesterol, and blood pressure. The results showed that the variables of gender and glucose levels had a significant influence on the incidence of diabetes, with significant values of 0.022 and 0.001, respectively. The gender variable has an Odds Ratio (OR) value of 0.135, indicating that female patients tend to have a lower risk of developing diabetes than men. Meanwhile, glucose levels showed a positive association with the incidence of diabetes, with each unit increase in glucose levels increasing the risk of diabetes by 1.016 times. The binary logistic regression model developed has an accuracy of 87.1% based on the Area Under Curve (AUC) value, which falls into the category of strong classification ability. This study provides important implications in supporting the development of more effective diabetes prevention and management strategies through an in-depth understanding of risk factors, so that it can be used as a basis for decision-making in public health services.

How to cite this article:
M. Anistya, P.K. Intan, A.H. Asyhar, A.H. Asyhar, "ANALYSIS OF DIABETES MELLITUS DISEASE USING BINARY LOGISTIC REGRESSION", Jurnal Statistika dan Aplikasinya, vol. 9, iss. 1, pp. 13 – 21, June 2025

Journal e-mail: *jsa@unj.ac.id*
Research Article · Open Access

1. INTRODUCTION

Non-communicable diseases (NCDs) are developing into a significant global health problem. According to a WHO report, NCDs contribute to 63% of total annual deaths, reaching more than 36 million people [1]. Diabetes is one of the critical components in the NCD category. Thamrin et al. [2] emphasized that diabetes has the potential to cause serious impacts such as cardiovascular disorders, kidney damage, blindness, and even amputation. Lin et al. [3] noted the disease as one of the top ten leading factors causing global mortality.

The prevalence of diabetes shows a consistent upward trend. Based on the International Diabetes Federation (IDF) report, in 2019 there were around 463 million people, or 9.3% of the population aged 20-70 years, with diabetes. Naim [4] projects that Indonesia ranks third in Asia with 19.5 million sufferers, with predictions of reaching 28.6 million by 2045. Complex risk factors contribute to the development of the disease, including unhealthy lifestyle, irregular diet, lack of physical activity, and heredity [5].

Medical studies show that diabetes is not just a simple metabolic issue. Yusnita et al. [6] identified the multidimensional impact, including physical, psychological, social, and economic complications. Apriyan et al. [7] emphasized that the quality of life of people with diabetes tends to decrease due to organ damage, visual impairment, neurological problems, and potential anxiety and depression. Hasibuan et al. [8] classified risk factors into two categories: non-modifiable (age, gender, heredity) and modifiable (diet, physical activity, obesity).

Approaches to diabetes prevention and management require a comprehensive strategy. Wahyurin et al. [9] recommended education on healthy lifestyle modifications, including increased physical activity, consumption of balanced meals, and weight control. Safaruddin and Permatasari [10] highlighted the importance of mathematical modeling, particularly binary logistic regression, in identifying risk factors and predicting the likelihood of diabetes occurrence. This method allows the analysis of categorical and numerical variables to understand the probability of disease.

A retrospective study at Puskesmas Teja revealed the complexity of the diabetes problem. In the July-December 2024 period, out of 140 patients served, 57.1% had a history of diabetes. Various previous studies, such as Ramadhan [11], Diksa & Fithriasari [12], and Susanti et al. [13], have used binary logistic regression to explore risk factors. Key variables that consistently affect diabetes risk include age, heredity, physical activity, diet, smoking habits, and metabolic health status [14], [15].

This study aims to identify and analyze the factors that influence the incidence of diabetes mellitus in patients at Puskesmas Teja, Pamekasan. Using a binary logistic regression approach, this study focuses on variables such as age, gender, heredity, smoking habit, body mass index (BMI), blood glucose level, cholesterol, and blood pressure as the main predictors. The study also aims to evaluate the strength of each variable's association with diabetes risk and provide a basis for more targeted and evidence-based health intervention strategies. With the expected results of the analysis, this study can support the prevention and management of diabetes more effectively at the public health service level.

2. METHODS

The study used quantitative methods to analyze the factors of diabetes mellitus in 140 patients at UPT Puskesmas Teja, Pamekasan. Using secondary data consisting of 60 negative and 80 positive patients with diabetes, this study aimed to identify the contribution of variables such as age, gender, blood pressure, smoking habits, BMI, cholesterol, and glucose levels to the risk of diabetes through a binary logistic regression approach.

$$log\left(\frac{P(Y=1)}{1-P(Y=1)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \quad (1)$$

Where $P(Y=1)$ is the probability of a person developing diabetes and β_0, β_k is the regression coefficient for variable X_k.

Statistical testing was conducted through two methods:

1. Concurrent Test (G-test): Using G statistics to evaluate the overall effect of the independent variable on the dependent variable. The formula used is:

$$G = -2ln\left(\frac{L_0}{L_1}\right) \quad (2)$$

Where L_0 is the likelihood model without predictor and L_1 is the likelihood model with predictor.
Test Criteria:
   a. Reject H_0 if $G > \chi^2_{df,\alpha}$ or if *Sig.* $< \alpha$, where χ^2 is the critical value of the chi-square distribution.
2. Partial Test (Wald Test): Used to assess the significance of each predictor variable individually. The formula used is:

$$W = \frac{\hat{\beta}_k^2}{Var(\hat{\beta}_k)} \quad (3)$$

Where $\hat{\beta}_k$ is the estimated regression coefficient of variable X_k and $Var(\hat{\beta}_k)$ is the variance of the estimated coefficient.

Test Criteria:
   a. Reject H_0 if $W > \chi^2_{1,\alpha}$ or if *Sig.* $< \alpha$.

In addition, to understand the risk impact of each variable on the incidence of diabetes, the Odds Ratio (OR) was calculated using the following equation:

$$OR = e^{\beta k} \quad (4)$$

Where βk is the regression coefficient of variable X_k. When $OR > 1$ is the variable increases risk and $OR < 1$ is the variable reduces risk.

This study also used the AUC-ROC curve to assess the accuracy of the model with the formula:

a. True Positive Rate (TPR):

$$TPR = \frac{True\ Positives}{True\ Positives + False} \quad (5)$$

b. False Positive Rate (FPR):

$$FPR = \frac{False\ Positives}{False\ Positives + True} \quad (6)$$

The resulting AUC (Area Under Curve) gives a value between 0 and 1 to evaluate how well the model predicts diabetes.

3. RESULTS AND DISCUSSION

This research utilizes secondary data collected from 140 patients at UPT Puskesmas Teja, Pamekasan. The dataset consists of 60 non-diabetic patients and 80 diabetic patients. The data was collected over the period from July to December 2023. The study analyzes eight variables, including age, gender, heredity, smoking habits, body mass index (BMI), glucose levels, cholesterol levels, and blood pressure. These variables were categorized into two types: categorical data and continuous data.

Categorical data includes gender (0 = male, 1 = female), heredity (0 = no family history, 1 = family history of diabetes), smoking (0 = non-smoker, 1 = smoker), and blood pressure (0 = normal, 1 = high blood pressure). Continuous data includes age, BMI, glucose level, and cholesterol levels.

A descriptive statistical analysis was performed on the data, providing key statistics including minimum, maximum, mean, and standard deviation for each variable. The summary of this analysis is presented in Table 1 below:

Table 1. Research Data Description

Variable	N	Minimum	Maximum	Mean	Std. Deviation
Age	140	33	83	57.29	9.360
Gender	140	0	1	0.61	0.490
Descent	140	0	1	0.41	0.494
Smoking	140	0	1	0.29	0.457
BMI	140	14.8	37.1	25.203	4.0239
Glucose	140	0	601	244.59	124.594
Cholesterol	140	0	455	202.59	107.685
Blood Pressure	140	0	1	0.71	0.453
Diagnose	140	0	1	0.57	0.497

This research uses two types of data, namely categorical data and continuous data. Before conducting further analysis, the first step is descriptive statistical analysis. For example, this analysis was applied to diabetes diagnosis data recorded between July and December 2023. The results of this descriptive analysis include information regarding the minimum, maximum, mean, and standard deviation values for each variable used in the study, which are presented in Table 1.

Table 1 shows the characteristics of predictor variables that may affect diabetes. The mean age (X1) was 57.29 years with a standard deviation of 9.360, indicating a patient age range of 33-83 years. The mean gender (X2) was 0.61, with slightly more females (1 = Female, 0 = Male) and a standard deviation of 0.490. The mean heredity (X3) was 0.41, with 41% of respondents having a family history of diabetes and a standard deviation of 0.494. Average smoking (X4) was 0.29, indicating 29% of respondents were smokers, with a standard deviation of 0.457. The mean body mass index (X5) was 25.203, with a standard deviation of 4.0239, indicating a range of 14.8-37.1, where 1 = Obese and 0 = Normal. The average blood sugar level (X6) was 244.59 with a standard deviation of 124.594, assuming 1 = Abnormal and 0 = Normal.

The average cholesterol (X7) is 202.59 with a standard deviation of 107.685, assuming 1 = Not Normal and 0 = Normal. The average blood pressure (X8) is 0.71 with a standard deviation of 0, 453, assuming 1 = High Blood Pressure and 0 = Normal. The average diabetes diagnosis in this study was 0.57, indicating 57% of the sample was diagnosed with diabetes, with a standard deviation of 0.497. A 1 indicates a positive patient with diabetes, and a 0 indicates a negative. The data in Table 1 shows the independent and dependent variables are categorical.

Parameter Estimation

Parameter estimation is performed to test the statistical significance of independent variables in predicting the dependent variable. At this stage, all variables in the diabetes dataset are tested to determine which variable X has a significant impact on variable Y, based on the significance value of each variable. The results of β parameter estimation for all variables can be seen in Table 2.

Table 2. Variables in Parameter Estimation

Variable	β	S. E	Wald	Sig.	Exp(β)
Age	-0.048	0.025	3.580	0.058	0.953
Gender	-2.005	0.877	5.223	0.022	0.135
Heredity	-0.272	0.600	0.205	0.650	0.762
Smoking	-1.689	0.881	3.678	0.055	0.185
BMI	-0.021	0.058	0.131	0.718	0.979
Glucose	0.016	0.004	21.394	0.001	1.016
Cholesterol	0.00	0.002	0.027	0.870	1.000
Blood Pressure	0.250	0.526	0.225	0.635	1.284
Diagnose	1.580	2.348	0.453	0.501	4.856

From Table 2, a binary logistic regression model will be formed involving eight variables contained in the β column with the following details:

$$g(x_i) = ln\frac{\pi(x_i)}{1-\pi(x_i)} = 1.580 - 0.048X_1 - 2.005X_2 - 0.272X_3 - 1.689X_4 - 0.021X_5$$
$$+0.016X_6 + 0.00X_7 + 0.250X_8$$

The model shows that the constant is 1.580, which describes the value of g(x) when all X variables are zero. The coefficients associated with the X variables show the changes in g(x) as follows: for X1, each 1-unit increase leads to a -0.048 decrease in g(x); for X2, a -2.005 decrease; for X3, a -0.272 decrease; for X4, a -1.689 decrease; for X5, a -0.021 decrease; and for X6, a 0.016 increase.

Probability

Table 3. The Probability of Having Diabetes Mellitus

No	X1	X2	X3	X4	X5	X6	X7	X8	Probability
1	36	1	1	0	24.14	212	0	1	0.329
2	59	1	0	0	26.31	112	0	0	0.882
3	63	1	0	0	28.23	136	287	1	0.834
4	60	1	1	0	22.68	410	252	1	0.059
5	63	1	1	0	19.53	532	369	1	0.009
-	-	-	-	-	-	-	-	-	-
136	53	1	0	0	31.2	163	231	0	0.734
137	56	1	0	0	30.1	229	253	1	0.457
138	78	0	0	1	20.4	141	248	0	0.883
139	65	1	0	0	24.7	296	261	1	0.284
140	57	0	0	0	24.8	141	196	1	0.303

Table 3 shows the probability of a person having or not having diabetes based on independent variables. The threshold value is the limit used to classify an observation in one of the two categories based on the model's predicted probability. In binary logistic regression, the threshold value is usually 0.5. That is, if the prediction probability is more than 0.5, the observation is categorized as positive (y = 1); if it is less than or equal to 0.5, the observation is categorized as negative (y = 0) (Riyani et al., 2024).

For example, to calculate the probability of someone suffering from diabetes mellitus at the age of 59, female, without family history, non-smoking, with a BMI of 26.31 kg/m², glucose level of 112 mg/dL, no cholesterol, and low blood pressure, the probability can be calculated with the equation that has been formed.

$$g(x_i) = 1.580 - 0.048(59) - 2.005(1) - 0.272(0) - 1.689(0) - 0.021(26.31) + 0.016(112) + 0.00(0) + 0.250(0)$$
$$= 1.580 - 2.832 - 2.005 - 0 - 0 - 0.552 - 1.792 + 0 + 0$$
$$= -2.0175$$
$$P = \frac{1}{1 + \exp\left(-(-2.0175)\right)}$$
$$= \frac{1}{1 + 0.1329}$$
$$= \frac{1}{1.1329}$$
$$= 0.8826$$

Thus, the probability of someone with these criteria experiencing diabetes mellitus is 88.26%, because the probability value is greater than the threshold value (0.5).

Simultaneous Test

Table 4. Model Summary

-2 Log likelihood	Chi-Square
127.829	15.507

Based on Table 4, the overall test results using the likelihood ratio test show the -2-log likelihood G value of 127.829, while the chi-square (X^2) value at $\alpha = 0.05$ with 8 degrees of freedom is 15.507313. The decision-making criterion is to reject H0 if the value of G is greater than the chi-square. Since the value of G > Chi-Square, it can be concluded that the independent variables collectively have an effect on variable Y, or at least one variable X has an effect on variable Y.

Partial Test

The individual or partial test results are as follows:

Table 5. Partial Test Result

Model	t	Sig.
(Constant)		
X1 (Age)	3.580	0.058
X2 (Gender)	5.223	0.022
X3 (Heredity)	0.205	0.650
X4 (Smoking)	3.678	0.055
X5 (BMI)	0.131	0.718
X6 (Glucose)	21.394	0.001
X7 (Cholesterol)	0.027	0.870
X8 (Blood Pressure)	0.225	0.635

Partial test results show that only gender (X2) and glucose (X6) variables have a significant influence on diabetes diagnosis, with sig. values of 0.022 and 0.001, respectively, which are smaller than 0.05. Meanwhile, other variables such as age (X1), heredity (X3), smoking (X4), BMI (X5), cholesterol (X7), and blood pressure (X8) showed no significant effect, as their sig. values were greater than 0.05. This indicates that the factors of gender and glucose level are more decisive in the diagnosis of diabetes compared to other factors.

Model Fit Test

Table 6. Hosmer and Lemeshow

Chi-Square	df	Sig.
11.449	8	0.178

From Table 6, the sig. value is 0.178, which means the sig. value. > α (0.05) and the Chi-Square count value (11.449) < Chi-Square table value (15.507). Because the Chi-Square count value < than the Chi-Square table value, accept H0. It can be concluded that the predicted binary logistic regression model is appropriate.

Odds Ratio (OR)

Odds Ratio (OR) is a statistical measurement used to assess the relationship between independent and dependent variables in a logistic model. Based on the calculations in Table 2, the OR value for variable X2 (Gender) of 0.135 indicates a negative correlation to the likelihood of developing diabetes. This means that each one-unit increase in the sex variable reduces the chance of developing diabetes by 0.135 times. Similarly, the variable X3 (Heredity) has an OR value of 0.762, indicating that a one-unit increase in the heredity variable will reduce the chance of developing diabetes by 0.762 times. This negative correlation is also seen in variable X4 (Smoking), with an OR value of 0.185, which means that an increased smoking frequency will reduce the risk of developing diabetes by 0.185 times.

On the other hand, variable X8 (Blood Pressure) showed a positive association with diabetes incidence, with an OR value of 1.284. This means that every one-unit increase in blood pressure will increase a person's chance of developing diabetes by 1.284 times. Thus, variables that have OR values less than 1 tend to decrease the risk, while OR values greater than 1 indicate an increased risk of diabetes. This interpretation helps us understand how risk factors contribute to the likelihood of disease occurrence.

Model Accuracy

Table 7. Confusion Matrix

Actual	Prediction	
	+	-
+	62	18
-	13	47

From the table above, the model correctly predicted 62 data that were positive, while 18 positive data were predicted as negative. In addition, 13 negative data were predicted as positive, and 47 negative data were correctly predicted as negative. These results show that the distribution of positive and negative classes is relatively balanced, which allows the model to maintain good performance on both classes.

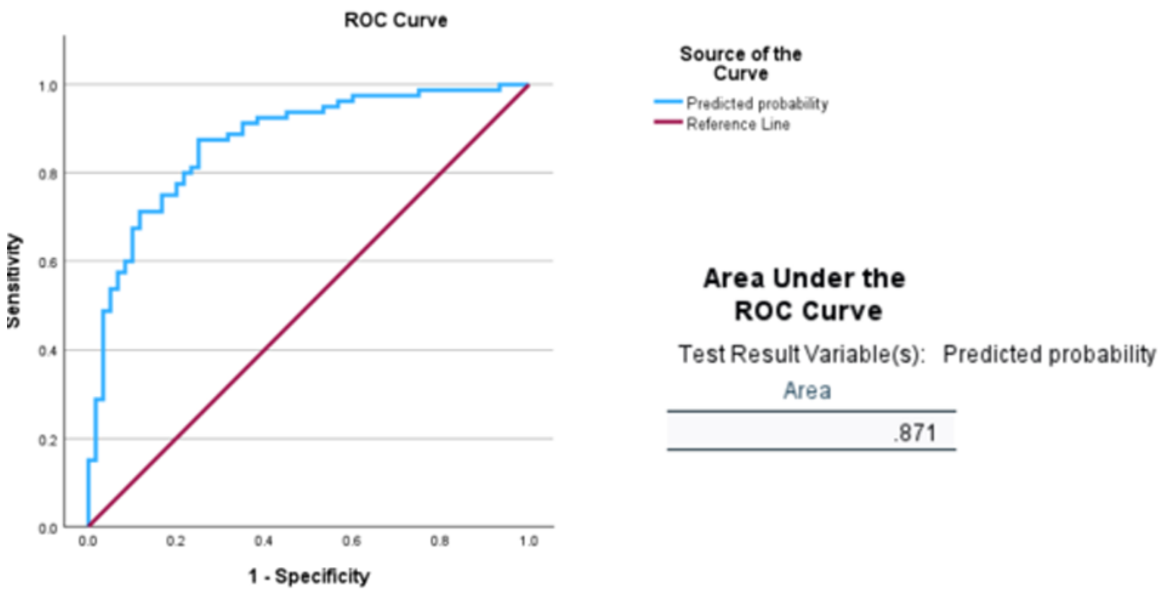

Figure 1. ROC Curve and Area Under the ROC Curve

The binary logistic regression model used to predict the likelihood of diabetes occurrence, considering risk factors, is shown in Figure 1. This model has an accuracy rate of 87.1% based on the Area Under Curve (AUC) value. With an AUC of 87.1%, which falls within the range of 0.8 to 0.9, the model proved to be robust in predicting the incidence of diabetes mellitus.

4. CONCLUSION

This study successfully identified risk factors that influence the incidence of diabetes mellitus using a binary logistic regression approach. From the analysis, it was found that the variables of gender and glucose levels had a significant influence on the incidence of diabetes. The gender variable showed that female patients had a lower risk than males, with an Odds Ratio (OR) value of 0.135. Meanwhile, glucose levels have a positive relationship with the incidence of diabetes, where each unit increases in glucose levels increases a person's chance of developing diabetes by 1.016 times. In contrast, other variables such as age, family history (heredity), smoking habits, Body Mass Index (BMI), cholesterol levels, and blood pressure did not show a significant influence on the incidence of diabetes in this study, with significance values greater than 0.05.

The binary logistic regression model developed has an accuracy rate of 87.1% based on the Area Under Curve (AUC) value, which falls into the strong classification ability category (AUC is in the range of 0.8-0.9). This indicates that the model can effectively predict the likelihood of a person having diabetes mellitus based on the variables studied. These results reinforce that binary logistic regression is a relevant and reliable statistical approach to analyze the relationship between risk variables and diabetes incidence.

5. REFERENCES

[1] Andi Mariani, A. Sauddin, Adiatma, and H. Nufus, "ANALISIS REGRESI LOGISTIK BINER UNTUK MENGIDENTIFIKASI FAKTOR-FAKTOR YANG MEMPENGARUHI INDEKS PRESTASI KUMULATIF MAHASISWA MATEMATIKA UNIVERSITAS ISLAM NEGERI ALAUDDIN MAKASSAR," Jurnal MSA (Matematika dan Statistika serta Aplikasinya), vol. 11, no. 1, 2023, doi: 10.24252/msa.v11i1.34860.

[2] N. Apriyan, A. Kridawati, and T. B. W. Rahardjo, "Hubungan Diabetes Mellitus Tipe 2 Dengan Kualitas Hidup Pralansia Dan Lansia Pada Kelompok Prolanis," Jurnal Untuk Masyarakat Sehat (JUKMAS), vol. 4, no. 2, 2020, doi: 10.52643/jukmas.v4i2.1028.

[3] G. J. Azmi, P. Q. Dewi, S. Y. C. Anugerah, S. D. Oktarina, A. Rizki, and M. R. Nurhambali, "Analisis Faktor yang Memengaruhi Tingkat Diabetes Melitus pada Masyarakat di Kota/Kabupaten di Provinsi Jawa Timur Tahun 2018," Jurnal Sains dan Kesehatan, vol. 5, no. 5, 2023, doi: 10.25026/jsk.v5i5.1896.

[4] P. A. S. Banilai and M. Sakundarno, "SYSTEMATIC REVIEW: FAKTOR-FAKTOR YANG BERHUBUNGAN DENGAN KEJADIAN DIABETES MELITUS (DM) PADA PENDERITA TUBERKULOSIS (TB)," Healthy Tadulako Journal (Jurnal Kesehatan Tadulako), vol. 9, no. 2, 2023, doi: 10.22487/htj.v9i2.739.

[5] R. E. Dewi, S. E. D. Putra, L. Aditama, and H. Wijono, "Pengaruh Edukasi Perubahan Gaya Hidup Sehat terhadap Clinical Outcome pada Pasien di Rumah Diabetes Universitas Surabaya," MPI (Media Pharmaceutica Indonesiana), vol. 3, no. 2, 2020, doi: 10.24123/mpi.v3i2.2982.

[6] I. G. B. N. Diksa and K. Fithriasari, "Analisis Faktor Resiko Penyebab Diabetes Mellitus dengan Regresi Logistik Biner," Inferensi, vol. 4, no. 1, 2020, doi: 10.12962/j27213862.v4i1.8480.

[7] A. Fauziah, "Pengaruh Beras Merah (Oryza Nivara) Dan Beras Jagung (Zea Mays L.) Terhadap Kadar Glukosa Darah Pada Penderita Diabetes Melitus Tipe 2 Di Puskesmas Antang Kota Makassar," Universitas Hasanudin Makassar, 2022.

[8] A. N. Fauziani, A. Adelia, O. B. Ardika, R. Himayani, and S. Rahmanisa, "Pengenalan Diabetes Mellitus Tipe 1 , Mutasi Genetik," Medula, vol. 14, no. 2, 2024.

[9] R. E. Fitri, E. Setiawan, M. Usman, and D. Aziz, "Analisis Regresi Logistik Biner terhadap Data Indeks Kedalaman Kemiskinan di Indonesia Tahun 2020," Jurnal Siger Matematika, vol. 3, no. 2, 2022.

[10] M. S. Fitria et al., “Edukasi Pencegahan Penyakit Diabetes Melitus dan Pemeriksaan Kadar Gula Darah Sewaktu di Panti Asuhan,” JURNAL INOVASI DAN PENGABDIAN MASYARAKAT INDONESIA, vol. 2, no. 3, 2023, doi: 10.26714/jipmi.v2i3.130.

[11] N. Hadianto, H. B. Novitasari, and A. Rahmawati, “KLASIFIKASI PEMINJAMAN NASABAH BANK MENGGUNAKAN METODE NEURAL NETWORK,” Jurnal Pilar Nusa Mandiri, vol. 15, no. 2, 2019, doi: 10.33480/pilar.v15i2.658.

[12] N. K. Hasibuan, S. Dur, and I. Husein, “Faktor Penyebab Penyakit Diabetes Melitus dengan Metode Regresi Logistik,” G-Tech: Jurnal Teknologi Terapan, vol. 6, no. 2, 2022, doi: 10.33379/gtech.v6i2.1696.

[13] U. R. Inayah, S. Suyitno, and M. Siringoringo, “Upaya Pencegahan Pencemaran Air Sungai Mahakam melalui Pemodelan Geographically Weighted Logistic Regression pada Data BOD,” EKSPONENSIAL, vol. 12, no. 1, 2021, doi: 10.30872/eksponensial.v12i1.755.

[14] Ardisson Korat, A. V., Willett, W. C., & Hu, F. B. (2014). Diet, lifestyle, and genetic risk factors for type 2 diabetes: a review from the Nurses' Health Study, Nurses' Health Study 2, and Health Professionals' Follow-up Study. Current nutrition reports, 3(4), 345–354. https://doi.org/10.1007/s13668-014-0103-5

[15] Wu, Y., Ding, Y., Tanaka, Y., & Zhang, W. (2014). Risk factors contributing to type 2 diabetes and recent advances in the treatment and prevention. International journal of medical sciences, 11(11), 1185–1200. https://doi.org/10.7150/ijms.10001