

Delay Prediction in BPJS Claim Payments Using a JST Backpropagation Neural Network

Dinda Aulia Putri Br Sembiring^{1*}, Relita Buaton², Selfira³

^{1,2,3} STMIK Kaputama

dindaauliaputribrsbr@gmail.com^{1*}, bbcbuatan@gmail.com², selfira.yap@gmail.com³

Abstract

Delayed BPJS claim payments can disrupt hospital financial management, requiring prediction based on historical data. This study applies an Artificial Neural Network using the Backpropagation algorithm to predict delayed BPJS claim payment status in Binjai City. Claim data from 2023–2024 were used for training, while 2025 data were used for testing. The model employed a 4–3–1 architecture, a learning rate of 0.2, a maximum of 10,000 epochs, and a target error of 0.01 using MATLAB. Training produced a Mean Squared Error of 0.0091831 at the 87th epoch. Testing on 12 records yielded 11 correct predictions and one incorrect prediction, achieving 91.67% accuracy. These results indicate Backpropagation performs well and may support faster, objective, consistent, measurable claim payment evaluation in hospitals.

Keywords: Artificial Neural Network, Backpropagation, BPJS Health, MATLAB, Payment Delay Prediction

1. Introduction

The National Health Insurance (JKN) program, administered by BPJS Kesehatan, is one of the cornerstones of healthcare financing in Indonesia. At referral healthcare facilities, payment for services is made through the Indonesian Case-Based Groups (INA-CBGs) system, which requires complete administrative documentation, accurate coding, and a well-organized claims verification process. The accuracy of claims payments is critical because these funds directly impact cash flow, operational activities, and the sustainability of hospital service quality [1].

In practice, delays in claim payments remain a problem that can cause financial losses for hospitals [2]. These delays are generally caused by incomplete documentation, discrepancies between diagnoses and INA-CBGs codes, errors in coding medical procedures, discrepancies in verification results, and the high workload of staff [3]. These conditions prolong the claim disbursement process and can disrupt the procurement of medications, payments to medical staff, and hospital operational funding [4].

Efforts to address claim delays have so far primarily involved document reviews, administrative improvements, staff capacity building, and inter-unit coordination. However, these approaches remain reactive in nature because they are implemented only after claims have been put on hold or experienced delays [5]. Previous research has also largely focused on the causes and impacts of claim delays, but has not yet produced a system capable of predicting the risk of delays early on based on historical data patterns [6].

One method that can be used to build such a prediction system is an Artificial Neural Network using the Backpropagation algorithm [7]. This method is capable of learning nonlinear relationships and recognizing complex patterns through a process of adjusting weights based on prediction errors [8]. Previous research has shown that backpropagation has been successfully used to predict inflation rates, rainfall, and food prices with low error rates and good accuracy [9].

Nevertheless, the application of backpropagation in previous studies has primarily been used to predict economic, environmental, and commodity phenomena, whereas its application to predict the status of late payments for BPJS Kesehatan claims remains limited [10]. Furthermore, there have not been many studies that specifically use historical claim data from Djoelham Binjai General Hospital as the basis for classifying claims as late or on time [11]. Therefore, this study applies a Backpropagation Artificial Neural Network to build a predictive model for BPJS Kesehatan claim payment delays and to measure its performance based on the Mean Squared Error and accuracy rate [8].

2. Research Methods

2.1. Research Data and Variables

This study uses 36 historical BPJS Kesehatan claim records from Djoelham General Hospital in Binjai for the 2023–2025 period. A total of 24 data points from 2023–2024 were used as training data, while 12 data points from 2025 were used as testing data. The data included service type, document completeness, verification status, claim processing time, and payment delay status.

Table 1: Research variables

Symbol	Variables	Description
X1	Type of service	Outpatient or inpatient
X2	Document completeness	Complete or incomplete
X3	Verification status	Approved or requires revision
X4	Claim processing time	Payment processing time in days
Y	Delay status	Late or on time

Variables X1 through X4 are used as inputs, while variable Y is used as the model's output target.

Table 2: BPJS Health Claims Data

Years	ID	Type of service	Document completeness	Verification	Claim processing time	Target
2023	K001	Inpatient Care	Complete	Approved	10	It's Not Too Late
2023	K002	Outpatient Care	Complete	Approved	12	It's Not Too Late
...	...	...	...	...	...	...
2023	K012	Outpatient Care	Complete	Approved	15	It's Not Too Late
2024	K001	Inpatient Care	Complete	Approved	9	It's Not Too Late
...	...	...	...	...	...	...
2024	K012	Outpatient Care	Complete	Approved	14	It's Not Too Late
2025	K051	Inpatient Care	Complete	Approved	10	It's Not Too Late
...	...	...	...	...	...	...
2025	K062	Inpatient Care	Incomplete	Correction	19	It's Too Late

Source: BPJS Kesehatan claims data from Djoelham General Hospital in Binjai for the 2023–2025 period.

2.2. Research Data and Variables

Claim data first undergoes transformation and normalization so that it can be processed by the Artificial Neural Network. Categorical variables are converted to numerical values, while the claim processing duration variable is normalized to the range 0–1. This step is performed to ensure that all variables are in a format compatible with the binary sigmoid activation function.

Table 3: Research Data Transformation

Variables	Category	Value
Type of service	Outpatient	0
	Inpatient	1
Completeness of Documents	Complete	0
	Incomplete	1
Verification Status	Under Review	0
	Approved	1
Delay Status	Not Late	0
	Late	1

Normalization was performed on the claim processing duration variable using the Min-Max Normalization method as follows:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

For example, a claim processing time of 10 days—with a minimum of 8 days and a maximum of 20 days—results in the following normalized value:

$$X' = \frac{10 - 8}{20 - 8} = 0,17 \quad (2)$$

The results of the transformation and normalization are then used as input data for the model training and testing processes. Examples of the processed data are shown in Table 4.

Table 4: Research Data Transformation

ID Claim	X1	X2	X3	X4	Y
K001	1	0	1	0,17	0
K002	0	0	1	0,33	0
K003	0	0	1	0,00	0
K004	1	1	0	1,00	1
...	...	...	...	...	...
K012	0	0	1	0,58	0

Source: Analysis of BPJS Kesehatan claim data from Djoelham General Hospital in Binjai.

2.3. Backpropagation Artificial Neural Network Architecture

The artificial neural network architecture used is a 4–3–1 architecture, consisting of four neurons in the input layer, three neurons in the hidden layer, and one neuron in the output layer. The four input neurons represent service type, document completeness, verification status, and claim processing time. The single output neuron generates a prediction of the claim payment status—whether it is late or not.

Table 5: Network architecture and parameters

Components	Configuration
Input layer	4 neuron
Hidden layer	3 neuron
Output layer	1 neuron
Network architecture	4–3–1
Activation functions	Sigmoid biner
Learning rate	0,2
Maximum epoch	10.000
Target error	0,01
Software	MATLAB

The binary sigmoid activation function is used because it produces outputs in the range of 0 to 1, making it suitable for binary classification. In this study, output values close to 0 are categorized as on-time claims, while values close to 1 are categorized as late claims. The model architecture can be described as follows:

$$(X1, X2, X3, X4) \rightarrow (Z1, Z2, Z3) \rightarrow Y \quad (3)$$

Description:

X1 : service type
 X2 : document completeness
 X3 : verification status
 X4 : claim processing time
 Z1–Z3 : neurons in the hidden layer
 Y : claim payment delay status

A learning rate of 0.2, a maximum of 10,000 epochs, and a target error of 0.01 were used in the model training process using MATLAB.

2.4. Training and Testing Process

The training process was conducted using 24 claim data points from 2023–2024, while 12 data points from 2025 were used in the testing phase. The data was split by period so that the model could be trained using historical patterns and tested on new data it had not previously encountered. The entire training and testing process was implemented using MATLAB.

The training stages of the Backpropagation Artificial Neural Network are shown in Table 6.

Table 6: Model training stages

Stage	Process
1	Input the transformed and normalized data
2	Initialize the network's weights and biases
3	Perform the feedforward process
4	Calculate the difference between the network's output and the target
5	Perform backpropagation to update the weights
6	Repeat the process until the target error or the maximum number of epochs is reached
7	Save the trained weights for testing

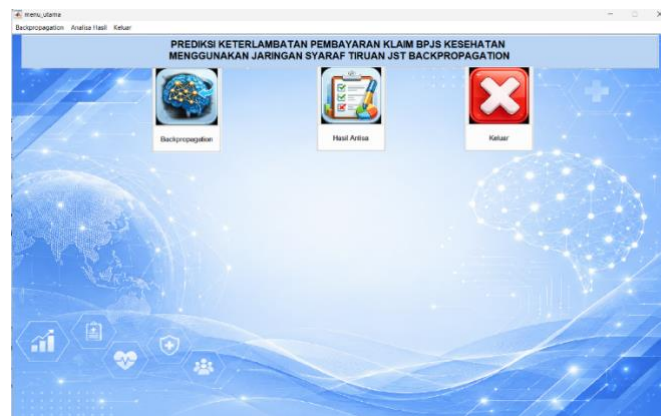
In the feedforward process, input values are passed to the hidden layer and the output layer to obtain a predicted output. Next, the output value is compared with the target to obtain an error value. This error is backpropagated from the output layer to the hidden layer to update the network's weights and biases. The process is repeated until the error value reaches the target of 0.01 or the number of training epochs reaches a maximum of 10,000.

Once the training process is complete, the resulting network weights are used to test the 2025 claims data. The model's output is then classified into two categories: a value of 0 for claims that are not late and a value of 1 for late claims. The predicted results are compared with the actual values to assess the model's ability to predict delays in BPJS Kesehatan claim payments.

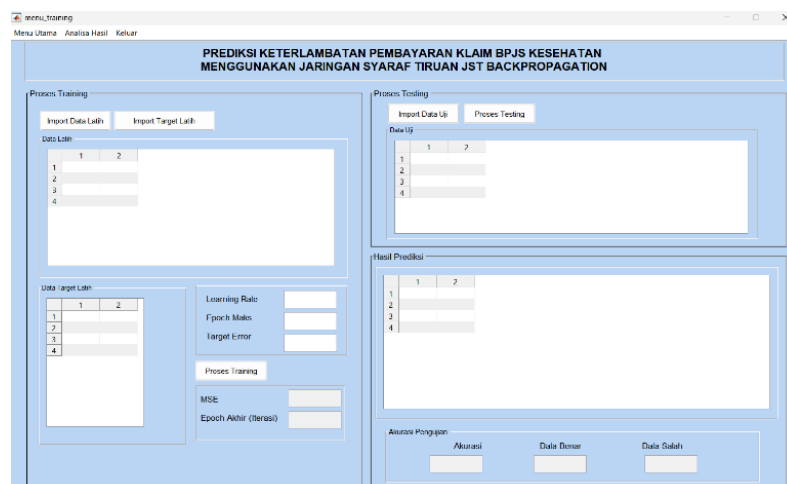
3. Results and Discussion

3.1. System Implementation

The The BPJS Kesehatan claim payment delay prediction system was developed using MATLAB R2015b. The system interface consists of a main menu, a Backpropagation menu, and an analysis results menu. The main menu serves as a navigation page for accessing the training and testing processes, viewing prediction results, and exiting the application.

**Fig. 1:** The system's main menu screen

The Backpropagation menu is the main page for running the model training and testing processes. In the training section, users can import training data and targets, and specify the learning rate, maximum number of epochs, and target error. In the testing section, users can import test data, run the prediction process, and view the accuracy rate, the number of correct predictions, and the number of incorrect predictions.

**Fig. 2:** Backpropagation menu display

3.2. Model Training Results

The training process was conducted using 24 BPJS Kesehatan claim datasets from 2023–2024. The model used a 4–3–1 architecture, a learning rate of 0.2, a maximum of 10,000 epochs, and a target error of 0.01. The training process in the system began by importing the data and targets, entering the network parameters, and then clicking the “Start Training” button.

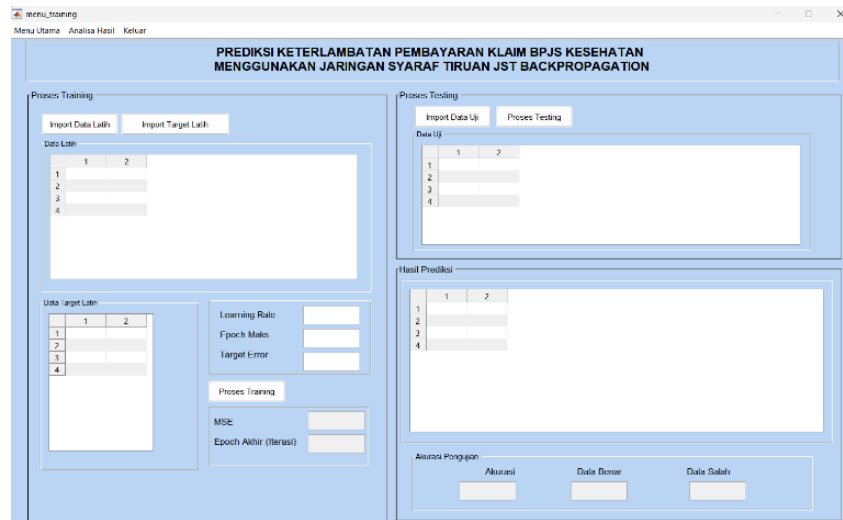


Fig. 3: Implementation of the model training process

The training results show that the network achieved an MSE value of 0.0091831 at the 87th epoch. This value is smaller than the target error of 0.01, so training was automatically stopped with the status “Performance Goal Met.” This result indicates that the network successfully converged without using the entire maximum limit of 10,000 epochs.

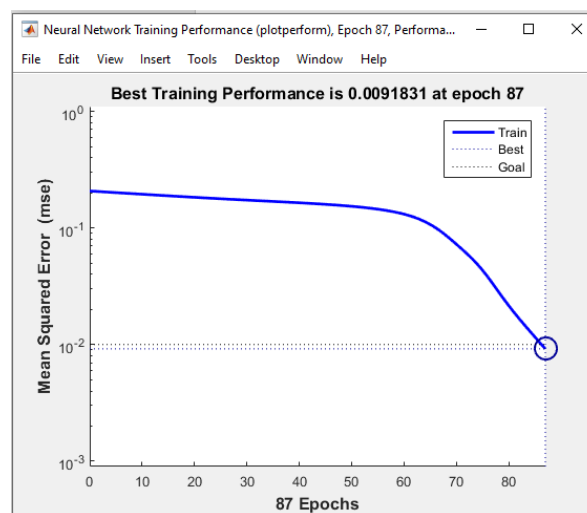


Fig. 4: Model training performance graph

The performance graph shows a gradual decrease in MSE until it reaches its best value of 0.0091831 at epoch 87. This decrease indicates that the process of updating the weights and biases brings the network’s output closer to the target. Additionally, the regression coefficient of 0.99177 indicates a very strong relationship between the network’s output and the target value.

3.3. Section headings

The testing process was conducted using 12 BPJS Kesehatan claim data points from 2025 that had not been used in training. The test data was imported via the “Import Test Data” button; the system then loaded the trained model, performed normalization, and generated predictions based on a threshold value of 0.5. Output values ≥ 0.5 were classified as “late,” while outputs < 0.5 were classified as “not late.”

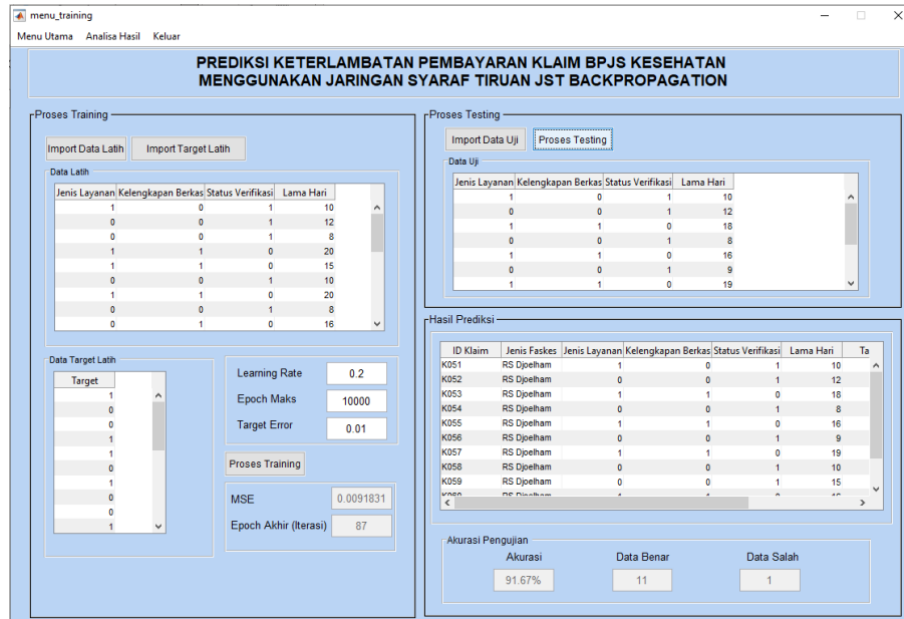


Fig. 5: Implementation of the model testing process

The test results show that 11 out of 12 data points were predicted correctly, while one data point resulted in a prediction error. Thus, the model's accuracy rate is 91.67%. These results indicate that the model has good generalization ability in recognizing patterns of late claim payments in new data.

3.4. Analysis of Prediction Results

The analysis results menu displays a prediction table, the number of test data points, the number of correct and incorrect predictions, the accuracy rate, a bar chart, and a pie chart. This display makes it easy for users to evaluate the model's results without performing manual calculations.

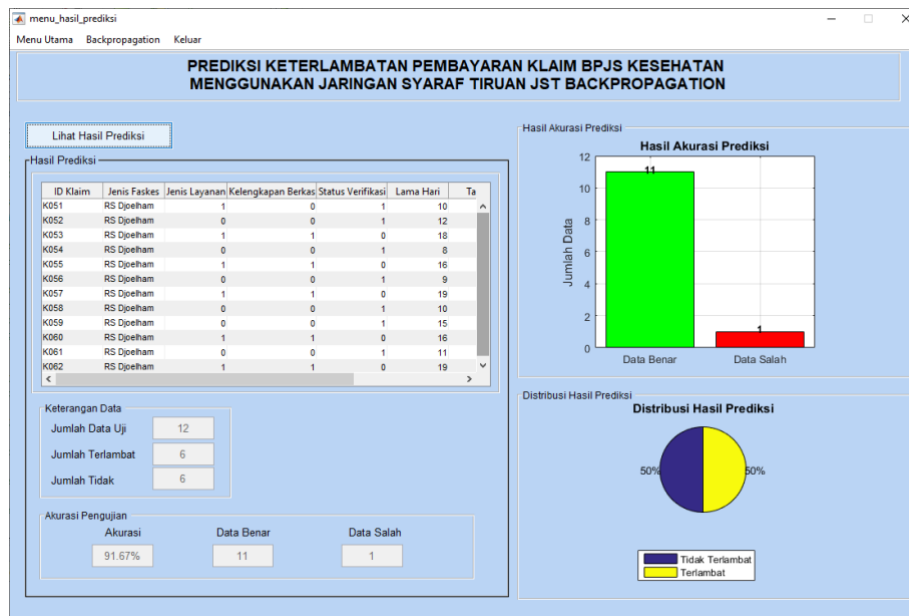


Fig. 6: Display of prediction analysis results

The analysis results show that delays tend to occur in claims with incomplete files, verification statuses that still require correction, and relatively long claim processing times. Conversely, complete files, approved verification statuses, and shorter processing times tend to result in a "not delayed" classification. This pattern indicates that the model is capable of learning the relationship between the input variables and the claim payment delay status.

3.5. Black-Box Testing Results

Black-box testing was conducted to ensure that each function in the application could receive input and produce output according to user requirements. The testing focused on navigation, data import, model training, data testing, and result presentation without examining the internal program code structure.

Table 7: Black-box testing results

No.	Menu/Function	Test Scenario	Expected Result	Result
1	Backpropagation menu	Select the Backpropagation menu	The training and testing page is displayed	Successful
2	Import training data	Select the training data file	The training data is displayed in the table	Successful
3	Import training target	Select the training target file	The target data is displayed in the table	Successful
4	Training process	Run the training process after entering the parameters	The system displays the MSE and epoch values and saves the model	Successful
5	Import testing data	Select the testing data file	The testing data is displayed in the table	Successful
6	Testing process	Run the testing process	The system produces predictions, accuracy, and the numbers of correct and incorrect predictions	Successful
7	Analysis results	Select the prediction results menu	The system displays the result table and visualisations	Successful
8	Main menu	Select the main menu	The system returns to the application's main page	Successful
9	Exit menu	Select the exit menu	The application closes successfully	Successful

Source: Application testing results using the black-box testing method.

The testing results demonstrate that all application functions operated according to the predefined scenarios. The system successfully imported data, trained the model, tested new data, calculated prediction accuracy, and presented the results in tables and graphical visualisations. No functional errors were identified during testing. Therefore, the application met the specified functional requirements and was considered suitable as a supporting tool for predicting delays in BPJS Kesehatan claim payments.

4. Conclusion

This study successfully implemented an Artificial Neural Network using the Backpropagation algorithm to predict delays in BPJS Kesehatan claim payments at RSU Djoelham Binjai. The model employed a 4–3–1 network architecture, a learning rate of 0.2, a maximum of 10,000 epochs, and a target error of 0.01. The training process achieved an MSE of 0.0091831 at the 87th epoch and a regression coefficient of 0.99177, indicating a strong relationship between the network output and the target values.

Testing on 12 claim records from 2025 produced an accuracy of 91.67%, with 11 correct predictions and one incorrect prediction. The model identified that incomplete documents, claims requiring verification corrections, and longer processing times were associated with a higher risk of payment delay. Black-box testing also confirmed that all application functions operated successfully. Therefore, the developed model can support early identification and evaluation of potential delays in BPJS Kesehatan claim payments.

References

- [1] Yurika, "JEVNAH Journal of Evidence-based Nursing and Public Health JEVNAH 2(2): 107-120 Factors contributing to delayed health insurance outpatient claims and administrative optimization strategies," *Journal of Evidence-based Nursing and Public Health*, vol. 2, no. 2, 2025, doi: 10.61511/jevnah.v2i2.2025.2211.
- [2] R. Y. Ahdinur, R. Semiarty, and R. Fahmy, "Penyebab Pengembalian Berkas Klaim Pelayanan Kesehatan JKN di Rumah Sakit Ibu Dan Anak Mutiara Bunda Padang," *JIK JURNAL ILMU KESEHATAN*, vol. 7, no. 1, p. 77, Apr. 2023, doi: 10.33757/jik.v7i1.633.
- [3] P. Setiawan, S. Setiatin, and Y. Ayu Nuraeni, "Analisis Penyebab Penundaan Klaim BPJS Kesehatan Pasien Rawat Inap Di Rumah Sakit X," *JURNAL ILMIAH PEREKAM DAN INFORMASI KESEHATAN IMELDA Vol.10 No.2*, Agustus 2025, pp. 220-229 *JURNAL ILMIAH PEREKAM DAN INFORMASI KESEHATAN IMELDA*, vol. 10, no. 2, pp. 2502–7786, 2025, [Online]. Available: <http://jurnal.uimedan.ac.id/index.php/JIPIKI> □ 220Journalhomepage:<http://jurnal.uimedan.ac.id/index.php/JIPIKI>
- [4] P. Natalia Napitupulu, A. Rahim Damanik, and J. Evonella Napitupulu, "Jurnal JPILKOM (Jurnal Penelitian Ilmu Komputer) IMPLEMENTASI ALGORITMA BACKPROPAGATION JARINGAN SYARAF TIRUAN UNTUK PREDIKSI ANGKA HARAPAN HIDUP DI KOTA JAMBI," *Jurnal JPILKOM*, vol. 1, no. 1, 2023.
- [5] F. Alfallah, Y. Yuhandri, and S. Sumijan, "Penerapan Artificial Neural Network untuk Memprediksi Persediaan Obat Esensial," *Jurnal komtekinfo*, pp. 63–72, Mar. 2025, doi: 10.35134/komtekinfo.v12i1.630.
- [6] D. Anyaprita, K. N. Siregar, B. Hartono, M. Fachri, and F. Ariyanti, "DAMPAK KETERLAMBATAN PEMBAYARAN KLAIM BPJS KESEHATAN TERHADAP MUTU PELAYANAN RUMAH SAKIT ISLAM JAKARTA SUKAPURA," *MUHAMMADIYAH PUBLIC HEALTH JOURNAL*, vol. 1, no. 1, 2020.
- [7] Putri Dewita Sari and Faiz Ahyaningsih, "Jaringan Saraf Tiruan Backpropagation untuk Prediksi Harga Bahan Pangan di Wilayah Kabupaten Deli Serdang," *Algoritma: Jurnal Matematika, Ilmu pengetahuan Alam, Kebumihan dan Angkasa*, vol. 2, no. 6, pp. 105–117, Nov. 2024, doi: 10.62383/algoritma.v2i6.287.
- [8] G. Zaida Muflih, "Penentuan Parameter Pembelajaran Jaringan Saraf Tiruan Backpropagation dan Pengaruhnya Terhadap Proses Pelatihan," *Jurnal Riset Teknologi Informasi dan Komputer (JURISTIK)*, vol. 1, no. 2, 2021, doi: 10.53863/juristik.v1i02.363.
- [9] L. Anas Zahir, M. Muhlis, J. Kimangunsarkoro Beji, K. Boyolangu, and K. Tuiungagung, "JARINGAN SARAF TIRUAN (NEURAL NETWORK) UNTUK ESTIMASI PRODUKTIVITAS PEKERJA KONSTRUKSI (ARTIFICIAL NEURAL NETWORKS FOR ESTIMATING CONSTRUCTION WORKER PRODUCTIVITY)," *Jurnal Teknik Sipil Universitas Tulungagung*, 2022, [Online]. Available: <http://journal.unita.ac.id/index.php/daktilitas/>
- [10] M. Thoriq, A. E. Syaputra, and Y. S. Eirlangga, "Prediksi Peningkatan Kunjungan Pasien Dimasa Mendatang Menggunakan Jaringan Saraf Tiruan Backpropagation," *JURNAL FASILKOM*, Apr. 2024.
- [11] P. Padilah, "PREDIKSI LUAS LAHAN PERTANIAN KECAMATAN HAURWANI MENGGUNAKAN PROGRAM MATLAB DAN JARINGAN SARAF TIRUAN," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 2, Apr. 2023, doi: 10.23960/jitet.v11i2.2890.