Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.4, August 2025, pp. 1201~1206
e-ISSN: 2548-6861

Determining Eligibility for Smart Indonesia Program (PIP) Recipients Using the Backpropagation Method

Ghinni Rizkya [1]*, Nurdin [2], Rini Meiyanti [3]***
* Department of Informatics, Universitas Malikussaleh
** Department of Information Technology, Universitas Malikussaleh
ghinni.200170018@mhs.unimal.ac.id [1] , nurdin@unimal.ac.id [2], rinimeiyanti@unimal.ac.id [3]

Article Info

Article history:

Received 2025-06-03
Revised 2025-06-25
Accepted 2025-07-03

Keyword:

Backpropagation Neural Network,
Classification,
Smart Indonesia Program (PIP),
Eligibility,
Data Mining.

ABSTRACT

The government provides financial assistance, educational opportunities, and expands access for students from poor or vulnerable families through the Smart Indonesia Program (PIP). At Madrasah Ibtidaiyah Negeri 20 Bireuen, the selection process for underprivileged students is still carried out manually by homeroom teachers by collecting data on students and their parents. This study aims to design, implement, and evaluate a classification method using the Backpropagation Neural Network to determine the eligibility of PIP scholarship recipients. The dataset consists of 309 entries, comprising 217 training data and 92 testing data, collected from MIN 20 Bireuen students between 2021 and 2023. The attributes used include father's occupation, mother's occupation, father's income, mother's income, number of dependents, number of vehicles, home ownership status, and card ownership status. Prior to training, the data were normalized using Min-Max scaling. The model was built with one hidden layer using a hard-limit activation function and a learning rate of 0.01. The classification results are categorized as "Eligible" and "Not Eligible". The model achieved an accuracy of 98%, precision of 100%, recall of 95%, and F1-score of 97%.

I. INTRODUCTION

In Law No. 20 of 2003 concerning the national education system, it has been regulated regarding the direction and method of implementing national education which contains the goals and functions of education in Indonesia. With the goals and functions of education that have been outlined in the law, the direction of education can be clearly seen that education in Indonesia aims to prepare a better generation of the nation [1]. Education is all learning knowledge that occurs throughout life in all places and situations that have a positive influence on the growth of each individual creature [2]. In this study it can be concluded that education can increase resources. With education that functions as a means and facility, it can develop and guide towards a better life, not only for oneself but also for others. Therefore, education must be organized as well as possible in order to produce a quality young generation in welcoming and facing the development of the era in the global era [3].

In the study it was stated that the Indonesia Pintar Program is an assistance fund given to students in the form of money which aims to ease the burden and help with the needs of the education process, especially for students who come from underprivileged or high-achieving families [4]. The Indonesia Smart Program is a government initiative to improve access and quality of education, especially at the primary level, by providing financial support to students based on certain criteria [5]. In the processing of data for determining prospective PIP scholarship recipients with the application of Data Mining in the information system, it can be used for processing data on PIP scholarship recipient participants. Because it will produce a more accurate accuracy value so that PIP scholarship recipients are right on target [6].

Poverty is included in the social problems that occur because the basic rights of poor people are not fulfilled to maintain and fulfill their lives. Everyone has the right and responsibility to have a good job. Therefore, this study will focus on people who cannot or do not want to work [7][8].

Poverty is caused by various factors, namely: status as the youngest child, living close to parents and caring for parents, no side business, low salary, long working hours, not interested in the work being done, low level of education, low economic status of parents, not persistent and not skilled [9].

Backpropagation for classification is quite effective, the level of accuracy in classifying phishing websites with the backpropagation algorithm provides a better level of accuracy compared to ELM where the prediction results on the testing set produced by the backpropagation algorithm [10]. Backpropagation has been proven to have a very good level of accuracy in performing classification [11]. The Backpropagation Neural Network (BPNN) was chosen in this study due to its capability to capture complex nonlinear patterns within data. Although the dataset size is relatively small, BPNN has been shown in prior studies to perform well for classification tasks involving socioeconomic attributes, such as eligibility determination for scholarship programs. Artificial neural network (ANN) is widely used in the field of optimization and classification. Backpropagation is utilized as a training method for neural networks with multiple layer [12]. Neurons in neigh-boring layers are connected, but neurons in every layer are not continuous. A BPNN canaccurately represent any continuous function [13]. Backpropagation is a gradient reduction method to minimize the squared output error. There are three stages that must be done in network training, namely the forward propagation phase, backpropagation, and weight and bias changes. The network architecture consists of an input layer, a hidden layer, and an output layer.

Data mining is a science that is used to handle the problem of retrieving information from large databases by combining techniques from statistics, machine learning, data visualization, pattern recognition, and databases. The goal of data mining is to intelligently extract information from a data set and then transform the information into a structure that can be understood for further use [14]. The successful development, implementation, and management of data mining projects in such organisations requiresa structured and repeatable approach [15]. According to [16] Data mining is a method used to extract hidden predictive information in databases, this is a very potential technology for companies in empowering data warehouses.

Classification according to [17] is a job of assessing data objects to include them in a certain class from the number of available classes. Classification builds a model based on existing training data, then uses the model to classify on new data. In order to develop an specific, research is required to investigate what should be the minimum impairment criteria [18]. Classification can be defined as a job that performs training or learning on a target function that maps each set of attributes to a single number of available class labels.

II. METHOD

Research steps use the waterfall method. Each step is arranged sequentially and systematically from start to finish to ensure that the system development process runs in a structured and logical manner. All of these methodological stages will be presented in Figure 1 below.

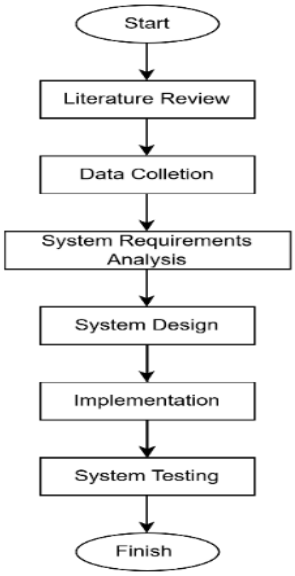

Figure 1 Research Step

Figure 1 shows the steps of the research methodology which begin with literature review, data collection, system needs analysis, system design, implementation, and system testing.

A. System Scheme Backpropagation

The Backpropagation Algorithm is an algorithm that functions to reduce the error rate by adjusting the weight based on the desired output and target [19]. It operates through an iterative learning process, where the weights in the neural network are updated in response to the error produced in the output layer. The main goal is to minimize this error by propagating it backward through the network and calculating gradients for each weight. Backpropagation allows the network to learn from mistakes and progressively improve its predictions over time.The system scheme is illustrated in Figure 2.

Figure 2 illustrates how Backpropagation processes input through a network to produce output, compares it with the target, and adjusts weights based on the error gradient to minimize future errors. The BPNN consists of three layers: input, hidden, and output. The input layer transmits signals to the hidden layer, while computations involving weights, biases, and activation functions occur in the hidden and output layers [20]. In this study, the Backpropagation model was built with 8 input nodes representing student socioeconomic attributes. The network architecture includes 1 hidden layer and 1 output node. The activation function used is hard-limit. Before training, the data were normalized using Min-Max Scaling to ensure all numerical attributes were on the same scale.

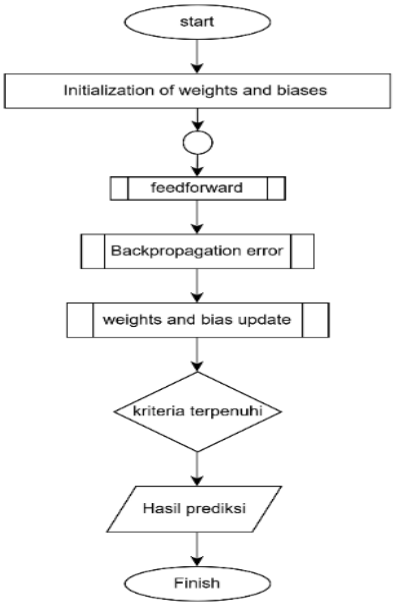

Figure 2 System Scheme Backpropagation

The normalization process used the following formula:

$$Normalisasi = \frac{data\ asli - min}{max - min} \quad (1)$$

This process was applied to features such as parents' income, number of dependents, and number of vehicles, both in the training and testing datasets, to improve model performance and training convergence.

B. Indonesian Smart Program

The Indonesia Smart Program (PIP) is an enhancement of the previous Poor Student Assistance (BSM) program, aimed at supporting students from underprivileged families. PIP recipients include holders of the Prosperous Family Card (KKS), participants of the Family Hope Program (PKH), persons with disabilities, orphans, and victims of disasters. They receive the Indonesia Smart Card (KIP) as proof of eligibility for the assistance. [21].

The system scheme is illustrated in Figure 3.

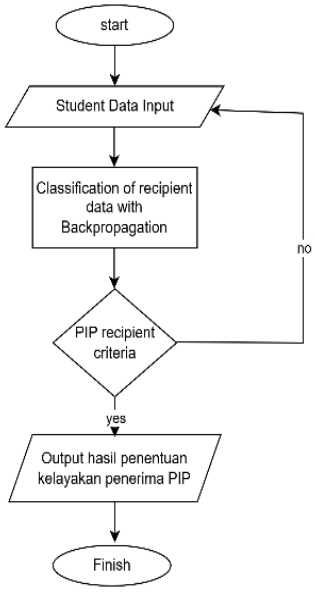

Figure 3 Scheme Smart Indonesian Program

Figure 3 explains the process of the Indonesia Smart program, starting with inputting student data, followed by classification using the backpropagation method. After that, the system will process the eligibility criteria and then output the results of students who are eligible to receive PIP. If the criteria entered do not match the data that has been input, the system will return to the data input section.

III. RESULT AND DISCUSSION

System design is the process of creating a system that serves as a clear and comprehensive representation or plan to meet users' needs. The system id designed using UML (unified Modeling Language) as a depicition of the system processes, making the system's operation clearer and easier to implement in programming.

A. Manual Calculation of the Backpropagation Method

In this system, there are several criteria used for assessment, namely: father's occupation, mother's occupation, father's income, mother's income, number of dependents, number of vehicles, home ownership status, card ownership status. Can be seen in table 1.

TABLE I
CRITERIA ASSESMENT

Criteria	Code
Father's occupation	x1
Mother's occupation	x2
Father's income	x3
Mother's income	x4
Number of dependents	x5
Number of vehicles	x6
Home ownership status	x7
Card ownership status	x8

TABLE II
INIZIALITATION STUDENT NAME

Student Name	Code
Juwita Aprilia	A1
Kanzul Akbar	A2
Khiara Arnita	A3
Muhammad Alif	A4
Muhammad Ouzan	A5
M Zayyan	A6
Raisya Mulya A	A7
Syafira	A8
Ummi Yati	A9
Zazila Azza	A10

Test data functions to provide an overview of how well the created model can generalize and perform effectively on previously unseen data. The test data in the PIP classification system can be seen in the following table 3.

TABLE III
DATA TESTING

	x1	x2	x3	x4	x5	x6	x7	8x
A1	2	9	3	8	1	2	3	0
A2	2	9	3	8	1	2	3	0
A3	2	9	3	1	1	2	3	0
A4	8	9	8	8	2	1	3	1
A5	8	9	5	8	3	2	3	1
A6	8	9	7	8	3	1	3	1
A7	8	10	7	9	3	1	3	1
A8	2	9	3	8	1	2	3	0
A9	8	9	8	8	2	1	3	1
A10	2	9	3	8	1	1	3	0

The next step is to determine the minimum and maximum values of the data attributes obtained from the values of each cell/column, then perform normalization of the testing data. The results of the normalization of the testing data can be seen in table 4.

TABLE IV
RESULT NORMALIZATION DATA TESTING

	x1	x2	x3	x4	x5	x6	x7	8x
A1	0,11	0,89	0,25	0.88	0,00	0,25	1,00	0,00
A2	0,11	0,89	0,25	0,88	0,00	0,25	1,00	0,00
A3	0,11	0,89	0,25	0,88	0,00	0,25	1,00	0,00
A4	0,78	0,89	0,88	0,88	0,33	0,00	1,00	1,00
A5	0,78	0,89	0,50	0,88	0,67	0,25	1,00	1,00
A6	0,78	0,89	0,75	0,88	0,67	0,00	1,00	1,00
A7	0,78	1,00	0,75	1,00	0,67	0,00	1,00	1,00
A8	0,11	0,89	0,25	0,88	0,00	0,25	1,00	0,00
A9	0,78	0,89	0,88	0,88	0,33	0,00	1,00	1,00
A10	0,11	0,89	0,25	0,88	0,00	0,00	1,00	1,00

The next stage is to determine the learning rate (which ranges from 0 to 1) at 0.01, then initialize the weights and biases of the first data with small numbers. The results of the analysis of weights and biases can be seen in the following table 5.

TABLE V
INITIALIZE WEIGHTS AND BIASES

w1	0,032
w2	0,0023
w3	-0,02
w4	-0,006
w5	-0,003
w6	-0,008
w7	-0,006
w8	0,03
b	0

In the feedforward phase, all outputs in the hidden layer units will be computed using the formula:

$$\sum = b + x1.w1 + x2.w2 + \cdots + x_n.w_2 \qquad (2)$$

Next, we calculate the output in the output layer using the binary activation function (hard limit), output layer calculation result can be seen in table 6 :

$$y(o) = \begin{cases} 0, jika\ o \leq 0 \\ 1, jika\ o > 0 \end{cases} \qquad (3)$$

TABLE VI
OUTPUT LAYER CALCULATION RESULTS

Code	∑	O
A1	-0,0127	0
A2	-0,0127	0
A3	-0,0127	0
A4	0,0272	1
A5	0,0317	1
A6	0,0287	1
A7	0,0282	1
A8	-0,0127	0
A9	0,0272	1
A10	-0,0107	0

In the table above are the results of the output layer calculations where ∑ is the result of summing all outputs from the hidden layer units and o is the value such that if ∑ is greater than or equal to 0, o equals 1, and if ∑ is less than 0, o equals 0. Then, the error in the hidden layer is calculated using the formula:

$$e = y - o \qquad (4)$$

The table below shows the results of the error calculation in the hidden layer where y represents the initial class, o is the calculated output layer result, and the error is the result of $\boldsymbol{y - o}$. Result of calculating the erroe in the Hidden Layer can be seen in table 7.

TABLE VII
RESULT OF CALCULATING THE ERROR IN THE HIDDEN LAYER

Code	y	o	error
A1	0	0	0
A2	0	0	0
A3	0	0	0
A4	1	1	0

A5	1	1	0
A6	1	1	0
A7	1	1	0
A8	0	0	0
A9	1	1	0
A10	0	0	0

The next stage is the weight update phase to update the new weights and biases. Then, the final results of the backpropagation classification method will be obtained. The results can be seen in Table 8.

TABLE VIII
RESULTS OF CLASSIFICATION USING THE BACKPROPAGATION METHOD

Code	∑	o	Result
A1	-0,0127	0	Not Eligible
A2	-0,0127	0	Not Eligible
A3	-0,0127	0	Not Eligible
A4	0,0272	1	Eligible
A5	0,0317	1	Eligible
A6	0,0287	1	Eligible
A7	0,0282	1	Eligible
A8	-0,0127	0	Not Eligible
A9	0,0272	1	Eligible
A10	-0,0107	0	Not Eligible

B. System Implementation

The purpose of this test page is to assess the accuracy of the model that has been trained using the backpropagation algorithm, by providing test data input and observing the model's prediction results, comparing the predicted results with the target data (ground truth), as well as evaluating the model's performance such as accuracy level, error (MSE, RMSE, etc.).

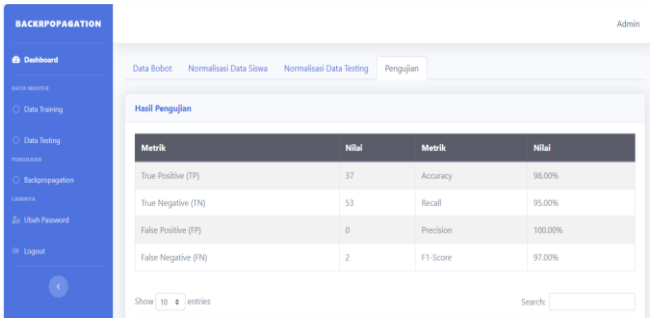

Figure 4 Testing Page

Figure 4 shows the results of the testing such as false positives (FP), false negatives (FN), true positives (TP), true negatives (TN), accuracy, recall, precision, and f1-score. This page is used after the model has finished training, when you want to see how well the model performs on new data, and for demonstration, presentation, or system validation purposes.

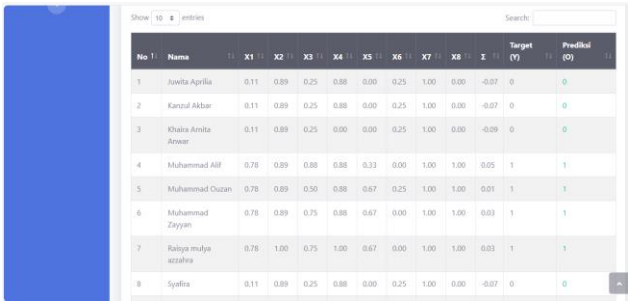

Figure 5 Continuation of the test page

Figure 5 shows the calculation results of Target (y) and Prediction (o) to produce the expected outcome.

TABLE IX
CONFUSION MATRIX

n = 92	Actual Eligible (1)	Actual Not Eligible (0)
Predicted Eligible (1)	TP = 37	FP = 0
Predicted Not Eligible (0)	FN = 2	TN = 53
	39	53

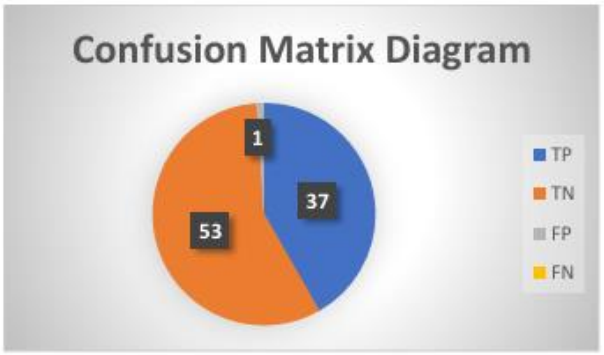

Figure 6. Diagram Confusion Matrix

TABLE X
MODEL EVALUATION CONFUSION MATRIX

Metric	Formula	Calculation	Result
Accuracy	(TP + TN) / (TP + TN + FP + FN)	(37 + 53) / (37 + 53 + 0 + 2) = 90 / 92	98%
Precision	TP / (TP + FP)	37 / (37 + 0) = 37 / 37	100%
Recall	TP / (TP + FN)	37 / (37 + 2) = 37 / 39	95%
F1-Score	2 × (Precision × Recall) / (Precision + Recall)	2 × (1.00 × 0.95) / (1.00 + 0.95) = 1.9 / 1.95	97%

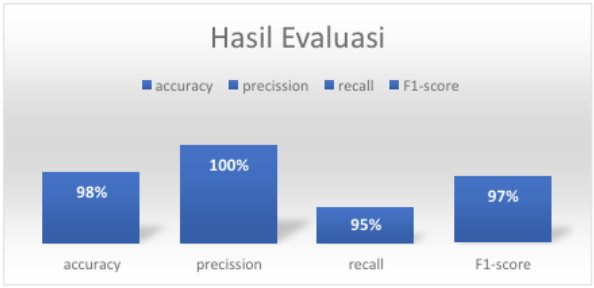

Figure 6. Evaluation Results

IV. CONCLUSION

A web-based system built with PHP has been successfully developed to classify the eligibility of recipients of the Indonesia Smart Program using the Backpropagation Neural Network method, with classification results consisting of two categories: Eligible and Not Eligible. The testing results conducted using the Backpropagation method on 92 test data yielded a true positive rate of 37, a false positive rate of 0, a true negative rate of 53, and a false negative rate of 2. Based on the testing results for determining the eligibility of recipients of the Indonesia Smart Program (PIP) using the Backpropagation method, there is a 98 percent accuracy rate. However, since the dataset only consists of 309 records, there is a risk of overfitting when using neural networks. In future research, cross-validation techniques such as **k-fold cross validation** should be implemented to improve the model's generalization and to mitigate the risk of overfitting. A web-based system built with PHP has been successfully developed to classify the eligibility of recipients of the Indonesia Smart Program using the Backpropagation Neural Network method, with classification results consisting of two categories: Eligible and Not Eligible.

REFERENCES

[1] D. Oleh, "Penentuan Kelayakan Penerima Program Indonesia Pintar (Pip) Menggunakan Metode Backpropagation Proposal SkripsI Disusun Sebagai Syarat Memperoleh Gelar Sarjana Komputer Prodi Teknik Informatika Fakultas Teknik Universitas Malikussaleh."

[2] S. Ujud, T. D. Nur, Y. Yusuf, N. Saibi, and M. R. Ramli, "Penerapan Model Pembelajaran Discovery Learning Untuk Meningkatkan Hasil Belajar Siswa Sma Negeri 10 Kota Ternate Kelas X Pada Materi Pencemaran Lingkungan," *J. Bioedukasi*, vol. 6, no. 2, pp. 337–347, 2023, doi: 10.33387/bioedu.v6i2.7305.

[3] Fajriana, Safriana, and Nurdin, "Students' Perceptions of Learning Using Powtoon Based on Gender in SMP/MTS," *Int. J. Eng. Sci. InformationTechnology (IJESTY*, vol. 2, no. 1, pp. 76–83, 2022, doi: 10.52088/ijesty.v1i1.205.

[4] A. Iskandar, "Sistem Pendukung Keputusan Kelayakan Penerima Bantuan Dana KIP Kuliah Menggunakan Metode ROC-EDAS," *Build. Informatics, Technol. Sci.*, vol. 4, no. 2, Sep. 2022, doi: 10.47065/bits.v4i2.2265.

[5] N. Aulia, N. Suarna, and W. Prihartono, "Klasifikasi Penentuan Penerima Program Indnesia Pintar Di Krwilbidikcam Greged Menggunakan Algoritma C4.5," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 6, pp. 3913–3919, 2024, doi: 10.36040/jati.v7i6.8294.

[6] A. Razi, "Klasifikasi Penerima Beasiswa Aceh Carong (Aceh Pintar) Di Universitas Malikussaleh Menggunakan Algoritma Knn (K-Nearest Neighbors)," *Jurnal Tika*, vo. 7, no.1, pp. 79–84, 2022. https://doi.org/10.51179/tika.v7i1.1116

[7] M. Qamal and I. Sahputra, "Sistem Pendukung Keputusan Penentuan Penerimaan PKH Menggunakan Metode Naïve Bayes."

[8] C. Maharani, D. A. Ningrum, A. E. Fatmawati, and A. Fadilla, "Dampak Kemiskinan terhadap Kualitas Pendidikan Anak di Indonesia: Rekomendasi Kebijakan yang Efektif," *J. Macroecon. Soc.*, vol. 1, no. 3, pp. 1–10, 2024, doi: 10.47134/jmsd.v1i3.199.

[9] E. Adawiyah, "Kemiskinan Dan Fakor-Faktor Penyebabnya," vol. 1, no. April, pp. 43–50, 2020.

[10] O. Putra Barus, "Perbangingan Metode Extreme Learning Machine dan Backpropagation untuk Mengklasifikasi Phising Websites," 2019.

[11] W. A. Adzani and P. S. Sasongko, "Klasifikasi Status Gizi Balita Menggunakan Metode Backpropagation Dengan Algoritma Levenberg-Marquardt dan Inisialisasi Nguyen Widrow," *J. Masy. Inform.*, vol. 12, no. 1, pp. 29–43, 2021, doi: 10.14710/jmasif.12.1.41020.

[12] S. K. Mondal *et al.*, "Glacial lakes outburst susceptibility and risk in the Eastern Himalayas using analytical hierarchy process and backpropagation neural network models," *Geomatics, Nat. Hazards Risk*, vol. 16, no. 1, p., 2025, doi: 10.1080/19475705.2024.2449134.

[13] R. Huang, B. Zeng, D. Ai, J. Yuan, and H. Xu, "Landslide dynamic hazard prediction based on precipitation variation trend and backpropagation neural network," *Geocarto Int.*, vol. 39, no. 1, p., 2024, doi: 10.1080/10106049.2024.2322058.

[14] Waldi Setiawan, Dedy Hartama, Muhammad Ridwan Lubis, Ihsan Syajidan, and Agus Perdana Windarto, "Klasifikasi Peminatan Topik Keilmuan Dalam Penyelesaian Studi Menggunakan Algoritma Naive Bayes," *J. Comput. Informatics Res.*, vol. 3, no. 2, pp. 188–198, 2024, doi: 10.47065/comforch.v3i2.1200.

[15] V. Plotnikova, M. Dumas, A. Nolte, and F. Milani, "Designing a data mining process for the financial services domain," *J. Bus. Anal.*, vol. 6, no. 2, pp. 140–166, 2023, doi: 10.1080/2573234X.2022.2088412.

[16] N. Noviyanto, "Penerapan Data Mining dalam Mengelompokkan Jumlah Kematian Penderita COVID-19 Berdasarkan Negara di Benua Asia," *Paradig. - J. Komput. dan Inform.*, vol. 22, no. 2, pp. 183–188, 2020, doi: 10.31294/p.v22i2.8808.

[17] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.

[18] O. R. Runswick, R. H. J. C. Ravensbergen, P. M. Allen, and D. L. Mann, "Expert opinion on classification for footballers with vision impairment: Towards evidence-based minimum impairment criteria," *J. Sports Sci.*, vol. 39, no. sup1, pp. 30–39, 2021, doi: 10.1080/02640414.2021.1881301.

[19] D. Hutabarat, Solikhun, M. Fauzan, A. P. Windarto, and F. Rizki, "Penerapan Algoritma Backpropagation dalam Memprediksi Hasil Panen Tanaman Sayuran," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 21–29, 2021, doi: 10.37148/bios.v2i1.18.

[20] J. R. Simanungkalit, H. Haviluddin, H. S. Pakpahan, N. Puspitasari, and M. Wati, "Algoritma Backpropagation Neural Network dalam Memprediksi Harga Komoditi Tanaman Karet," *Ilk. J. Ilm.*, vol. 12, no. 1, pp. 32–38, 2020, doi: 10.33096/ilkom.v12i1.521.32-38.

[21] A. S. Gutama, M. Fedryansyah, and E. Nuriyah, "Implementasi Program Kartu Indonesia Pintar (Kip) Berdasarkan Basis Nilai Keadilan Dalam Kebijakan Implementation of the Kartu Indonesia Pintar (Kip) Based on the Value Basis of Justice in Social Policy," *Penelit. dan Pengabdi.*, vol. 2, no. 3, p. 394, 2021.