Jurnal E-KOMTEK (Elektro-Komputer-Teknik)
Vol. 5, No. 2 (2021) pp. 190-202
https://jurnal.politeknik-kebumen.ac.id/index.php/E-KOMTEK
p-ISSN : 2580-3719 e-ISSN : 2622-3066

Classification of Poverty Reduction Program Recipients with Neural Network Algorithm in East Kotawaringin Communities

Nurahman, **Dwi Tjahjo Seabtian**
Department of Information Systems, Universitas Darwan Ali, Kotawaringin Timur, Indonesia, 13210

nurrahman.ikhtiar@gmail.com

https://doi.org/10.37339/e-komtek.v5i2.751

Published by Politeknik Dharma Patria Kebumen

Artikel Info
Submitted:
04-11-2021
Revised:
13-11-2021
Accepted:
15-11-2021
Online first :
30-12-2021

Abstract

Indonesia has a poverty rate of 24.79 million. Kotawaringin Timur is inhabited by 27.4 thousand people with an income of less than Rp. 416,777/month. The provision must be right on target, and recipients of assistance must use the assistance following the rules determined by the government. This research is to formulate a conceptual model of the Neural Network Algorithm structure that can be used to predict the use of assistance funds. This research applies the Knowledge Discovery Data methodology with Neural Network Algorithm for classification. The research has shown that the application of the Neural Network Algorithm with feature selection can improve performance with values AUC=0.974, CA=0.977, F1=0.977, Precision=0.977, Recall=0.977. The level of performance value for accuracy of Neural Network Algorithm in classifying is the excellent classification category. The recommended Neural Network parameter models are Neurons in hidden layers 100, Activation ReLu, Solver Adam, Regularization, α = 0.0001, and a Maximal number of iterations 200.

Keywords: Data mining, Classification, Neural network

Abstrak

Indonesia memiliki angka kemiskinan 24,79 juta. Kotawaringin Timur dihuni oleh 27,4 ribu jiwa dengan pendapatan kurang dari Rp. 416.777/bulan. Pemberiannya harus tepat sasaran, dan penerima bantuan harus menggunakan bantuan sesuai dengan aturan yang ditetapkan oleh pemerintah. Penelitian ini bertujuan untuk merumuskan model konseptual struktur Algoritma Neural Network yang dapat digunakan untuk memprediksi penggunaan dana bantuan. Penelitian ini menggunakan metodologi Knowledge Discovery Data dengan Algoritma Neural Network untuk klasifikasi. Hasil penelitian menunjukkan bahwa penerapan Algoritma Neural Network dengan pemilihan fitur dapat meningkatkan kinerja dengan nilai AUC=0.974, CA=0.977, F1=0.977, Precision=0.977, Recall=0.977. Tingkat nilai performansi untuk akurasi Neural Network Algorithm dalam pengklasifikasian termasuk dalam kategori klasifikasi sangat baik. Model parameter Neural Network yang direkomendasikan adalah Neuron di lapisan tersembunyi 100, ReLu Aktivasi, Solver Adam, Regularisasi, = 0,0001, dan jumlah iterasi Maksimal 200.

Kata-kata kunci: *Data mining, Klasifikasi, Jaringan syaraf*

1. Introduction

Indonesia is a country that has an uneven population distribution from Sabang to Merauke, from Miangas to Rote Island, which expands its population over time. Every year, the population increases. As the population increases, the problems also increase, including the high burden of dependents, low population quality, low population productivity, weak national economy, and the increasing poverty rates in various regions.

The government is responsible for achieving an ideal, just, and prosperous society. It can be seen from the community perspective in assessing the poverty level seen from economic factors. In September 2019, Indonesia still had a poverty rate of 24.79 million people **[1].** The Province of Central Kalimantan in September 2019 showed a poverty rate of 2.140.688/poor households **[1].** Almost all regions in Indonesia have a poverty rate. In 2019, it is evident that the poverty rate in East Kotawaringin reached a percentage of 27.4 thousand people with an income of less than Rp. 416,777/month **[2]**. The district government of East Kotawaringin continues to strive to reduce poverty. Efforts to reduce poverty have encountered various obstacles, such as covid-19 attacking the health of the Indonesian people. This condition disturbs all people. On June 25, 2020, the data increased from 1,178 to 50,187 confirmed cases of COVID-19, with a CFR reaching 5.2% of the total cases **[3]**.

The COVID-19 pandemic is very disturbing for the community, especially the poor. To keep the poverty rate under control, the government launched various programs to help the community. Poverty reduction is carried out in collaboration with many parties down to the neighborhood unit (RT) level. The varied types of social programs launched by the government require assistance. It can be from the local government collaborating with the RT. Family Hope Program (PKH) is one of the programs that require assistance.

The PKH assistance program aims to help the community and is carried out in a well-coordinated manner so that it is right on target. PKH beneficiaries are determined based on several components. Those who fulfill these components will receive PKH assistance based on the verification carried out by PKH facilitators. The beneficiaries must also use the assistance, according to the provisions set by the government, so that it can be beneficial.

The development of technology makes data collected by humans in electronic form more and more. Data is growing very fast, and almost all online media provide data that can be conveyed easily by all groups **[4]**. A collection of data will be meaningless if it is not managed properly. Data is processed to find a new pattern, knowledge, or information needed by all

parties, which is often referred to as data mining [5]. Data mining can be categorized into several roles, including Clustering, Prediction, Classification, Estimation, and Association. The implementation of the data mining is by choosing one of the algorithms. Neural Network Algorithm is an algorithm that can perform data classification. Neural Networks are also beneficial for solving complex cases, such as prediction, identification, pattern recognition, etc. [6]. The CFS method can significantly improve the accuracy performance of the classifiers algorithm [7].

Based on the problems above and the results of the literature review, the research was carried out using a neural network algorithm to classify PKH rock recipient data. The results of the research can provide a structure of a neural network that can be used as a parameter to predict PKH ROC recipients who are consistent in using aid according to government regulations.

2. Method

Information technology has changed all circles so that all aspects connect with technology [8]. Technology has made it easier for all parties to share and obtain information [9]. According to the press council records, in Indonesia, 43,000 sites are news sites [10]. This data shows that there are more and more websites in Indonesia. Therefore, organizations can use network techniques to increase productivity, reduce costs, and more [11].

The increasing use of the web and the abundance of data make experts explore knowledge from the data. Data mining is a process of inferring knowledge from a large data set [12]. Data mining is done to find patterns and models applied in scientific findings to business analysis [13]. In this research, pattern discovery was carried out using the Neural Network Algorithm with feature selection. According to [14], feature selection is used to lighten the processing load in the data mining model. The data mining roles include estimation, classification, clustering, prediction, and association [15]. This research applied the role of data mining, namely Classification. To determine the performance of the algorithm in classifying datasets, cross-validation, confusion matrix, and ROC curves can be used, as well as the accuracy of the classification process carried out in data mining [16]. Neural Network is an artificial intelligence algorithm that can learn, grow, and adapt to a dynamic environment [17]. The results of data processing were obtained by comparing the results of Neural networks without selection features and with selection features. When the results of the different tests on

statistical changes, or when there is a difference in value if the P-value <0.05, it can be stated that the change in value is significant [18]. Data processing and analysis were done scientifically based on the Knowledge Discovery Data (KDD) Methodology [19]. This methodology consists of 5 steps. Broadly speaking, this methodology is as follows:

1) Data selection
2) Pre-processing/ Cleaning
3) Transformation
4) Data mining
5) Interpretation/Evaluation

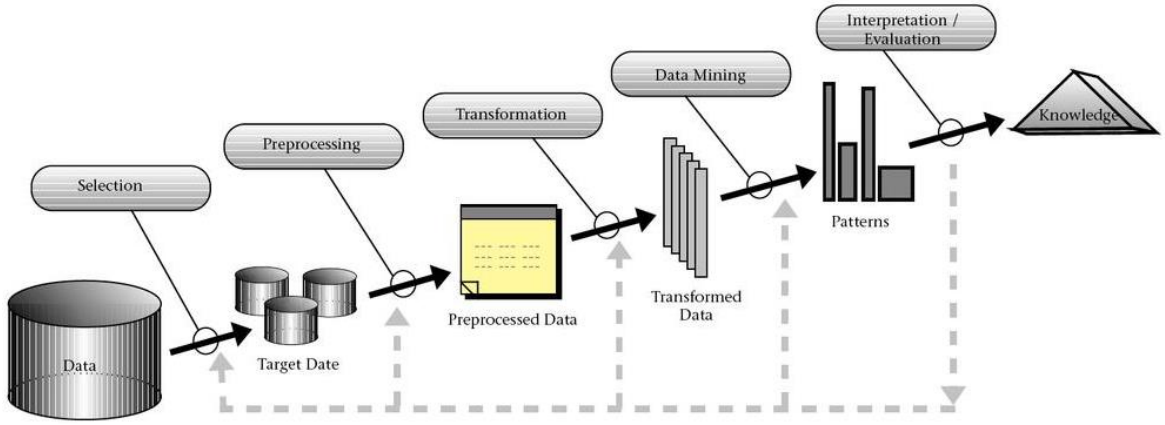


Figure 1. Knowledge Discovery Data (KDD)

The Knowledge Discovery Data (KDD) methodology will produce knowledge. So, by using these stages, new knowledge can be beneficial in the decision-making process by an agency or government. For the user to gain new knowledge, data mining needs to pay attention to analysis from data collection to getting unexpected pattern relationships and must be able to summarize data in different ways so that the results can be easily understood [20].

3. Results and Discussion

a. Data Source

The research data were collected from the East Kowaringin District Social Service. The data to analyze were the Family Hope Program (PKH) aid recipients, which were obtained gradually. Firstly, it was observing the data needs. Then, it was determining the type of data to analyze. Furthermore, it was the withdrawal of the data recorded on the social welfare of the e-PKH application by accessing the East Kowaringin District Social Service website in https://epkh.kemsos.go.id/login.xhtml.

b. Research Dataset

The dataset is an essential part of the research obtained from the object of the research case. The dataset is a set of data to be trained from a collection of data or past information. The information is used to find out the data's patterns and models. The dataset used in this research was as many as 11,229 records. The dataset has 20 columns as attributes and 1 column as the label. The label on the dataset is shown in the suspension column, which means that each participant will be between suspended or not suspended conditions. Each participant who complies with the stipulated conditions belongs to a non-suspended condition, and those who do not will be in a suspended condition. The dataset obtained from the data source in the initial conditions is in **Table 1** below.

Table 1. Initial Dataset Condition

No	Attributes	Missing	Distinct	Type	Unique
1	No	0	11229	Numeric	11229
2	Nomor Peserta (Participant's Number)	0	11229	Numeric	11229
3	Nama Pengurus (Management's Name)	0	6350	Text	5051
4	Tempat Lahir (Birthplace)	209	946	Text	501
5	Tanggal Lahir (Date of Birth)	99	6370	Date	3965
6	Ibu Kandung (Biological Mother)	124	5102	Text	3698
7	Alamat (Address)	0	4327	Text	2631
8	Jumlah Anak SD (Number of Children attending Elementary School)	0	3	Numeric	0
9	Jumlah Anak SMP (Number of Children attending Junior High School)	0	3	Numeric	0
10	Jumlah Anak SMA (Number of Children attending High School)	0	3	Numeric	0
11	Jumlah Bumil (Number of Pregnant Women)	0	2	Numeric	0
12	Jumlah Usia Dini (Number of Early Age)	0	3	Numeric	0
13	Jumlah Lansia (Number of Elderly)	0	3	Numeric	0
14	Jumlah Disabilitas (Number of Disabled People)	0	3	Numeric	0
15	Jumlah ART (Number of Housemaids)	0	7	Numeric	1
16	Nama Propinsi (Province's Name)	0	1	Nominal	0
17	Nama Kabupaten (Regency's Name)	0	1	Nominal	0
18	Nama Kecamatan (Sub-district's Name)	0	17	Nominal	0
19	Nama Kelurahan (Village's Name)	0	183	Text	1
20	Nama Pendamping (Assistance's Name)	0	47	Nominal	0
21	Penangguhan (Suspension)	0	2	Nominal	0

The initial dataset needs to be improved because it has a lot of missing data. The missing data in this research was 0.1% of the data from 15 features and 0.6% from 5 meta-attributes. Datasets that are still missing are considered irrelevant for processing. Thus, action is needed to make the data more relevant. The dataset can be improved into the relevant data to the research case raised, namely carrying out the attribute selection and data pre-processing stages. The condition of the initial data before pre-processing when entered the orange 3.30 tools is presented in **Figure 2** below.

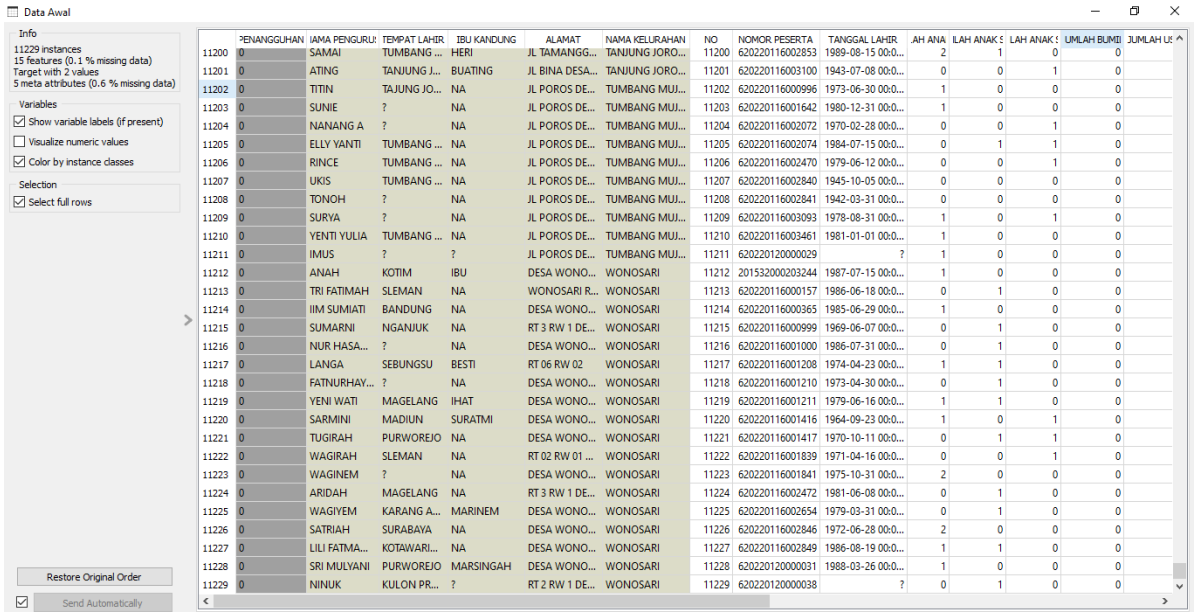


Figure 2. Initial Data View

c. Data Selections

Data Selection is to select the required attributes/columns. Unnecessary attributes will be removed at this stage. In **Table 1**, the total number of attributes owned was 21. Then, the column selection stage was carried out to find out which columns were considered relevant for analysis. At this stage also determine which column is an attribute or label. The dataset that has been selected will be used as a dataset that enters the mining process. In the selection process, several attributes are omitted. The selection results obtained as many as 16 columns of attributes. Attributes have labels that are commonly referred to as target columns. So, the dataset has 15 criteria or attributes with one label column. It can be observed that the attributes omitted are the attributes of No, Participant's Number, Management's Name, Province's Name, and Regency's Name. **Table 2** shows the dataset that has been carried out by data selection.

Table 2. Data selection attribute

No	*Initial Attribute*	*Selected Attributes*	*Note*
1	No	*Skip*	-
2	Nomor Peserta (Participant's Number)	*Skip*	-
3	Nama Pengurus (Management's Name)	*Skip*	-
4	Tempat Lahir (Birthplace)	Tempat Lahir	Attribute
5	Tanggal Lahir (Date of Birth)	Tanggal Lahir	Attribute
6	Ibu Kandung (Biological Mother)	Ibu Kandung	Attribute
7	Alamat (Address)	Alamat	Attribute
8	Jumlah Anak SD (Number of Children attending Elementary School)	Jumlah Anak SD	Attribute
9	Jumlah Anak SMP (Number of Children attending Junior High School)	Jumlah Anak SMP	Attribute
10	Jumlah Anak SMA (Number of Children attending High School)	Jumlah Anak SMA	Attribute
11	Jumlah Bumil (Number of Pregnant Women)	Jumlah Bumil	Attribute
12	Jumlah Usia Dini (Number of Early Ages)	Jumlah Usia Dini	Attribute
13	Jumlah Lansia (Number of Elderly)	Jumlah Lansia	Attribute
14	Jumlah Disabilitas (Number of Disabled People)	Jumlah Disabilitas	Attribute
15	Jumlah ART (Number of Housemaids)	Jumlah ART	Attribute
16	Nama Propinsi (Province's Name)	*Skip*	-
17	Nama Kabupaten (Regency's Name)	*Skip*	-
18	Nama Kecamatan (Sub-district's Name)	Nama Kecamatan	Attribute
19	Nama Kelurahan (Village's Name)	Nama Kelurahan	Attribute
20	Nama Pendamping (Assistance's Name)	Nama Pendamping	Attribute
21	Penangguhan (Suspension)	Penangguhan	Lable

In **Table 2**, several attributes were not used in the data processing. The "No" was used as an attribute that showed the data records' number, so it was not needed in the mining process. The "Participant's Number" and "Management's Name" were attributes that indicated objects or records of data, not a criterion to determine a goal (label), so they were also not used in classifying data. Then the attributes of Province's Name and Regency's Name were also not used in data classification because the focus of this research is in the Kotawaringin Timur Regency, so all recorded data contains the same data.

d. Pre-Processing Data

At the pre-processing stage, the reparations were made to the damaged or inconsistent data. In addition, accuracy and caution are vital here because the research results' quality depends on the quality of the data itself. So, to get a good data quality, the Pre-Processing stage carried out in this research repaired and deleted data so that the dataset became qualified. As mentioned previously, the initial data of this research amounted to 11,229 records. After Pre-Processing, it became 10,961 records, which were considered qualified data. **Figure 3** shows the view of the attribute selection results and data pre-processing.

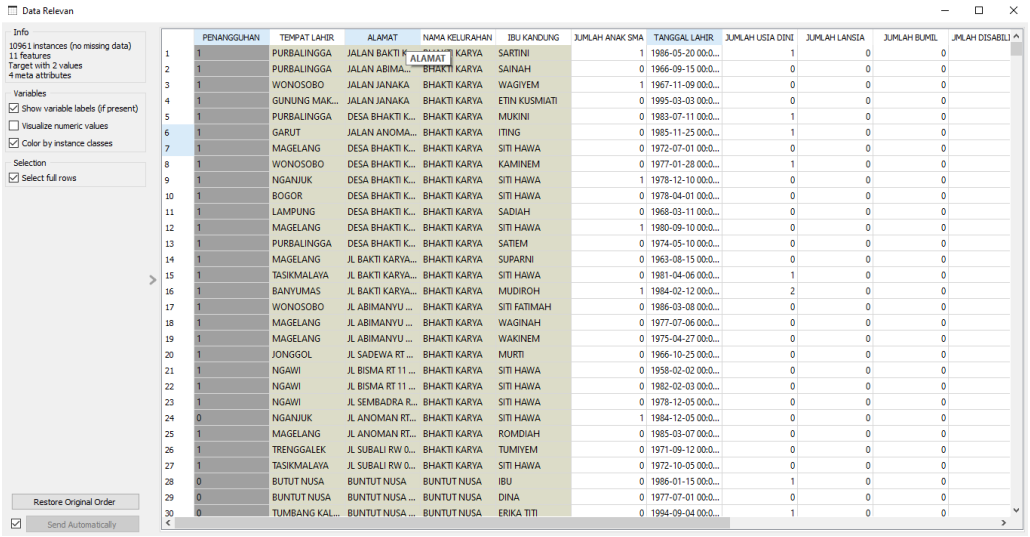


Figure 3. View Attribute Selection Results and Pre-Processing (Relevant Data)

e. Transformations

The Transformation stage aims to change the value data into more ideal data to facilitate the data's mining process. Performance testing was carried out in two stages, namely testing without feature selection and testing with feature selection. To make the dataset better, data transformation stages were also done. The steps that can be done in the data transformation are normalizing the data.

f. Data Mining

Datasets that have gone through the Pre-Processing stage and other stages are ready for mining processes. The mining process is to get new knowledge or information. At this stage, it is primary to determine the methods and algorithms used in the research. The Methods Data processing was conducted with data classification using the Neural Network Algorithm. The mining process was done twice. The first mining process was carried out by looking for patterns or information using the Neural Network Algorithm based on the attributes needed in

the search for data patterns. Then, in the second mining process, the Neural Network algorithm was used by selecting attributes (Features) using the Information Gain Algorithm.

The first mining process was carried out with the Neural Network Algorithm, which applied several parameter provisions to produce an AUC value of 0.500 and a classification accuracy of 0.976. The performance of the Neural Network Algorithm in classifying datasets is illustrated in **Table 3**. This research also presents some performance values generated from the classification using the Neural Network Algorithm.

Table 3. Neural Network Algorithm Performance

Model	*AUC*	*CA*	*F1*	*Precision*	*Recall*
Nural Network	0,500	0,976	0,965	0,953	0,976

In the second mining process, the Information Gain Algorithm was applied as a feature selection algorithm. The implementation of the Information Gain Algorithm was done by looking for the Entropy value and the Gain value. The formula for finding the Entropy value can be seen in equation **(1)**, and that for obtaining the gain value can be seen in equation **(2)** below.

$$Entropy(S) = \sum_{i=i}^{n} -p_i * log_2\, p_i \quad (1)$$

$$Gain\,(S, A) = Entropy(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * \; Entropy(S_i) \quad (2)$$

Based on the Neural Network Algorithm model by selecting features using the Information Gain Algorithm first and applying the parameters of the Neural Network Algorithm, the algorithm performance value was obtained. The resulting performance was an AUC value of 0.970 and an accuracy of 0.977. **Table 4** shows the performance value of the Neural Network Algorithm.

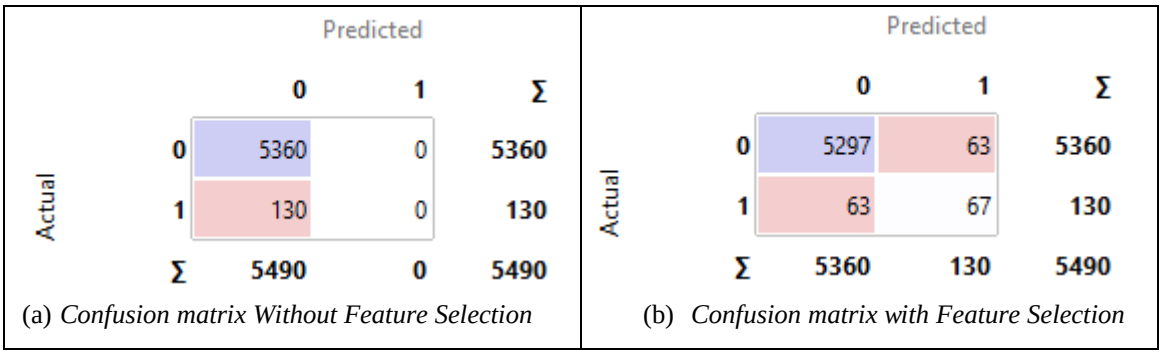
Table 4. Neural Network Algorithm performance with feature selection

Model	*AUC*	*CA*	*F1*	*Precision*	*Recall*
Information Gain + Nural Network	0,974	0,977	0,977	0,977	0,977

g. Evaluations

Algorithm performance evaluation is very important in the data classification process. In data mining and the Decision Support System, the Confusion matrix value can be used to

determine the accuracy value that functions in analyzing whether the classifier can recognize different classes (labels). Figure 4 shows the results of the Confusion matrix from the results of data processing. Figure 4.a shows that the correct prediction is 5360+0 = 5360 and the wrong prediction is 130+0 = 130 out of a total of 5490 testing data. So that the results of the accuracy level are known to be 5360/5490 *100% = 0.976. Then, in **Figure 4.** b, it is known that the correct prediction is 5297+67 = 5364 and the wrong prediction is 63+63 = 126. So, the result of the correct prediction accuracy in Figure 4.b is 5364/5490*100% = 0.977.



(a) *Confusion matrix Without Feature Selection* (b) *Confusion matrix with Feature Selection*

Figure 4. Confusion matrix value

Based on the results of the mining process, the evaluation results show that the Neural Network Algorithm has significantly increased the AUC value if the feature selection is carried out with the Information Gain Algorithm. It is just that in the process of training and testing, it takes longer; the training time was 197.872 seconds, and the testing time was 0.702 seconds. The results of the comparison of algorithm performance are shown in **Table 5**.

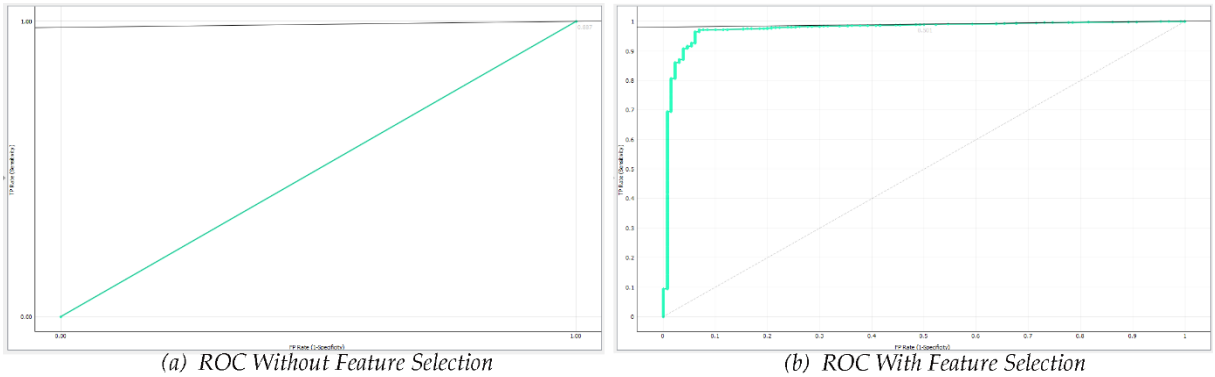
Table 5. Performance Comparison

Model	Train Time (s)	Test Time (s)	AUC	CA	F1	Precision	Recall
Neural Network	48,414	0,663	0,500	0,976	0,965	0,953	0,976
Information Gain + Neural Network	197,872	0,702	0,974	0,977	0,977	0,977	0,977

AUC values are used to represent probabilities. The accuracy level of the data classification can also be seen from the AUC value. Assessment of the level of accuracy of the AUC can be categorized into the following levels [21]:0.90 - 1.00 = excellent classification;

0.80 - 0.90 = good classification;
0.70 - 0.80 = fair classification;
0.60 - 0.70 = poor classification;
0.50 - 0.60 = failure.

Performance analysis is also described by using the ROC Curve. ROC curves are used to visualize, organize, and select classifiers based on their performance. Technically, the ROC curve is a two-dimensional graph having True Positive Rate (Sensitivity) and False Positive Rate (1-Specificity) results. The vertical line (Y) is the True Positive Rate axis, while the horizontal line (X) is the False Positive Rate axis. In this research, the analysis result using the ROC curve is shown in Figure 5. In Figure 5, the analysis is carried out by determining that the target class is not suspended. Figure 5.a has a level of accuracy with the failure category, while **Figure 5**.b has an accuracy level of excellent classification. In conclusion, because the diagonal line on the ROC curve is closer to perpendicular, the AUC value will be higher, meaning that if the AUC value is close to 1 or 100%, it belongs to the excellent classification category.



(a) ROC Without Feature Selection *(b) ROC With Feature Selection*

Figure 5. ROC Analysis

The Neural Network Algorithm produces a performance value included in the excellent classification category. This is influenced by the parameter model structure implemented in the algorithm. The parameter model applied to the Neural Network Algorithm was 100 neurons in hidden layers, and the maximum number of iterations is 200 times. **Figure 6** shows the report of the parameter model applied in the research.

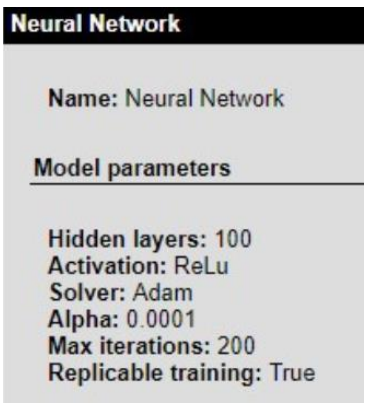


Figure 6. Neural Network Parameter Model

4. Conclusion

Classification using the Neural Network Algorithm with a dataset of social assistance recipients in East Kotawaringin district can produce performance scores on AUC, CA, F1, Precision, and Recall. It is known that the performance results on the data without using the feature selection algorithm produce values of AUC = 0.500, CA = 0.976, F1 = 0.965, Precision = 0.953, and Recall = 0.976. Then if the data is selected first using the Information Gain Algorithm, the results of the performance of the Neural Network Algorithm in classifying the data are AUC = 0.974, CA = 0.977, F1 = 0.977, Precision = 0.977, and Recall = 0.977. So, it can be concluded that the Neural Network Algorithm application by conducting feature selection with the Information Gain Algorithm first improved the performance in classifying datasets for social assistance recipients in East Kotawaringin Regency. The recommended Neural Network parameter models are Hidden Layers 100, Activation ReLu, Solver Adam, Alpha 0.0001, Max Iterations 200, and Replicable Training True.

References

[1] BPS, "Profil Kemiskinan di Indonesia September 2019," Jakarta, Jan. 2020. Accessed: Oct. 20, 2020. [Online]. Available: https://www.bps.go.id/pressrelease/2020/01/15/1743/persentase-penduduk-miskin-september-2019-turun-menjadi-9-22-persen.html.

[2] T. B. Kotim, "Garis Kemiskinan dan Penduduk Miskin 2017-2019," Badan Pusat Statistik Kabupaten Kotawaringin Timur, 2020. https://kotimkab.bps.go.id/indicator/23/760/1/garis-kemiskinan-dan-penduduk-miskin.html (accessed Oct. 25, 2020).

[3] Gugus Tugas Percepatan Penanganan COVID-19, "Peta Sebaran Kasus COVID-19 di Indonesia," covid19.go.id, 2020. https://covid19.go.id/peta-sebaran-covid19(accessed Oct. 23, 2020).

[4] A. Sinkov, G. Asyaev, A. Mursalimov, and K. Nikolskaya, "Neural networks in data mining," 2016, doi: 10.1109/ICIEAM.2016.7911596.

[5] N. Hadianto, H. B. Novitasari, and A. Rahmawati, "Klasifikasi Peminjaman Nasabah Bank Menggunakan Metode Neural Network," J. Pilar Nusa Mandiri, 2019, doi: 10.33480/pilar.v15i2.658.

[6] A. P. Windarto, "Implementasi JST Dalam Menentukan Kelayakan Nasabah Pinjaman KUR Pada Bank Mandiri Mikro Serbelawan Dengan Metode Backpropogation," J-SAKTI (Jurnal Sains Komput. dan Inform., vol. 1, no. 1, pp. 12–23, 2017, doi: 10.30645/j-sakti.v1i1.25.

[7] N. Nurahman and P. Prihandoko, "Perbandingan Hasil Analisis Teknik Data Mining 'Metode Decision Tree, Naive Bayes, Smo Dan Part' Untuk Mendiagnosa Penyakit Diabetes Mellitus," J. Inf., vol. 4, no. 1, pp. 39–44, 2019, doi: 10.25139/inform.v4i1.1403.

[8] T. Makmur, "Teknologi Informasi :Dampak dan Implikasi Bagi Perpustakaan,

Pustakawan Serta Pemustaka," INFO Bibl. J. Perpust. dan Ilmu Inf., vol. 1, no. 1, pp. 65–74, 2019, doi: 10.24036/ib.v1i1.12.

[9] W. Setiawan, "Era Digital dan Tantangannya," Semin. Nas. Pendidik. 2017, 2017, Accessed: Sep. 30, 2020. [Online]. Available: http://eprints.ummi.ac.id/151/2/1. Era Digital dan Tantangannya.pdf.

[10] Y. A. Prasetyo and A. Djauhar, "Mendorong Profesionalisme Pers Melalui Verifikasi Perusahaan Pers," J. DEWAN PERS, vol. 14, pp. 1–72, Jun. 2017, Accessed: Oct. 03, 2020. [Online]. Available: https://dewanpers.or.id/assets/ebook/jurnal/715-BUKU JURNAL DEWAN PERS 14 INDONESIA.pdf.

[11] Homeland Security, "Recommended Practice: Improving Industrial Control Systems Cybersecurity with Defense-In-Depth Strategies," Ics-Cert, 2016, Accessed: Oct. 27, 2020. [Online]. Available: https://us-cert.cisa.gov/sites/default/files/recommended_practices/NCCIC_ICS-CERT_Defense_in_Depth_2016_S508C.pdf.

[12] G. Kesavaraj and S. Sukumaran, "A study on classification techniques in data mining," Jul. 2013, doi: 10.1109/ICCCNT.2013.6726842.

[13] M. J. Zaki and W. Meira, Jr, Data Mining and Machine Learning. Cambridge University Press, 2020.

[14] R. Sowmya and K. R. Suneetha, "Data Mining with Big Data," in Proceedings of 2017 11th International Conference on Intelligent Systems and Control, ISCO 2017, Feb. 2017, pp. 246–250, doi: 10.1109/ISCO.2017.7855990.

[15] J. Suntoro, Data Mining Algoritma dan Implementasi dengan Pemrograman PHP, 1st ed. Jakarta: Elex Media Komputindo, 2019.

[16] A. Rohman, "Komparasi Metode Klasifikasi Data Mining Untuk Prediksi Penyakit Jantung," Neo Tek., vol. 2, no. 2, pp. 21–28, Dec. 2017, doi: 10.37760/neoteknika.v2i2.766.

[17] C. Cristina and A. Kurniawan, "Sejarah, Penerapan, dan Analisis Resiko dari Neural Network: Sebuah Tinjauan Pustaka," J. Inform. J. Pengemb. IT, vol. 3, no. 2, pp. 259–265, 2018, doi: 10.30591/jpit.v3i2.890.

[18] S. W. Ren, X. Qi, and Y. Q. Wang, "Serum amyloid A and pairing formyl peptide receptor 2 are expressed in corneas and involved in inflammation-mediated neovascularization," Int. J. Ophthalmol., vol. 7, no. 2, pp. 187–193, Apr. 2014, doi: 10.3980/j.issn.2222-3959.2014.02.01.

[19] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Elsevier, 2012.

[20] N. Saputra, K. Tania, and R. Heroza, "Penerapan Knowledge Management System (Kms) Menggunakan Teknik Knowledge Data Discovery (Kdd) Pada Pt Pln (Persero) Ws2Jb Rayon Kayu Agung," J. Sist. Inf., 2016.

[21] F. Gorunescu, Data mining: Concepts, models and techniques, vol. 12. Springer, Berlin, Heidelberg, 2011.