Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12973
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Prediction of Student Graduation with the Neural Network Method Based on Particle Swarm Optimization

Hafis Nurdin[1)*], Sartini[2)], Sumarna[3)], Yana Iqbal Maulana[4)], Verry Riyanto[5)]
[1,2,3)] Universitas Nusa Mandiri, [4,5)] Universitas Bina Sarana Informatika, Indonesia
[1)] hafis.nnr@nusamandiri.ac.id, [2)] sartini.sar@nusamandiri.ac.id, [3)] sumarna.smn@nusamandiri.ac.id, [4)] yana.yim@bsi.ac.id, [5)] verry.vry@bsi.ac.id

Submitted : Sep 6, 2023 | **Accepted** : Sep 11, 2023 | **Published** : Oct 1, 2023

Abstract: In private universities in Indonesia, student graduation is something that is worth paying attention to, because it will be an aspect of the success of the university. Universities certainly have data on students who graduated, where student graduation data is very important to be taken into consideration by private universities, however with a lot of data it will make it difficult for private universities to find information from this data. Other researchers have previously carried out student graduation data with the same data by examining student graduation data using other methods. So we need a data mining algorithm that has never been tested on student graduation data. The method used is the neural network method with an optimization method, namely the particle swarm optimization method, to test the data, which will later produce information that is very useful for universities. After testing the student graduation data and getting accuracy results using the neural network method of 84.55% and after being optimized using the particle swarm optimization method, the accuracy results were optimal with a value of 86.94%. These results can be used by private universities to predict that students will graduate on time before they take their final semester so that the student graduation rate will be high.

Keywords: Student Graduation; Neural Networks; Particle Swarm Optimization; Prediction;

INTRODUCTION

Higher education is one aspect that has an important role in the formation of quality individuals and the progress of a country's development. One indicator of the success of higher education is the student graduation rate, which is a challenge in improving the quality of higher education. A high graduation rate indicates the effectiveness of the educational program, while a low graduation rate can indicate problems in the learning process or academic management and can have a negative impact on the effectiveness and efficiency of the educational process and hinder the achievement of the goals of the educational institution. The problem that occurs in educational institutions in achieving high graduation rates is that they do not yet have a way to predict student graduation. And we don't yet know how students at these universities graduate on time (Rohmawan, 2018). In fact, predicting students graduating on time can be predicted from the first semester using existing data mining methods that have been carried out by researchers, including predicting graduating students using Decision Tree (DT), Neural Network (NN), Support Vector Machine (SVM) (Riyanto, Hamid, & Ridwansyah, 2019). Optimized graduation from previous results with SVM and Particle Swarm Optimization (PSO) (Suhardjono, Wijaya, & Hamid, 2019). Different from SVM and PSO, other researchers also predict

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i4.12973 p-ISSN : 2541-044X

student graduation using SVM and Genetic Algorithm (GA) (Ridwansyah, Wijaya, & Purnama, 2020). Other researchers also tried to predict student graduation using DT and PSO (Hendra, Azis, & Suhardjono, 2020).

Among the several algorithms used, the current goal is to improve the efficiency of predicting student graduation better than before and is a very important thing to do. So a method is needed to develop other methods that can help predict student graduation with high accuracy. One method that can be used to predict student graduation is using neural network techniques. Because NN is a computational model that resembles the human brain with the structure and function of the neural network that the NN has. This method has proven successful in various research fields, including in predicting and analyzing educational data. However, in optimizing neural network performance, careful methods and approaches are needed by determining the right parameters and architecture. Methods that combine optimization techniques with neural networks have become increasingly popular recently. One promising method is the use of particle swarm optimization in optimizing neural networks to predict student graduation (Iqbal, Herliawan, Ridwansyah, & Gata, 2020). The PSO method can optimize other methods, because PSO works like the social behavior of a group of animals, especially the movement of a group of birds or fish (Ridwansyah & Purwaningsih, 2018). PSO can be used to optimize the parameters involved in forming and training a neural network, in this case, regarding research on predicting student graduation. By using PSO, we can look for a combination of parameters that produces a neural network model with the best performance (Indah Ariyati, Ridwansyah, & Suhardjono, 2018). This PSO algorithm has been widely used in various fields, including parameter optimization in predictive models. When applied in the context of predicting student graduation, PSO can help find the optimal combination of weights and parameters for a neural network. The application of the PSO-based neural network method in predicting student graduation has the potential to overcome several problems that may occur in the conventional prediction process. The use of this method is expected to produce a more accurate and stable predictive model, because PSO is able to find a solution that is close to the global optimum in a complex parameter space.

There are updates from several studies, methods and optimizations that have been carried out, so this research has an update where the neural network method that has been used by researchers with the same data above is tested again using the PSO optimization method that has been carried out on the SVM and DT methods that have been carried out. was carried out first, whereas the neural network method trial had never been carried out on the student graduation data.

LITERATURE REVIEW

Datamining is the process of processing data from databases on large data with the data needed to obtain knowledge information stored in that large data (Nasyuha et al., 2021). This process involves looking for patterns in the student graduation database to predict whether the student will graduate on time or not. In the data mining process there is a classification and there are categorical labels which become a class that has been determined in the data mining process as well as a set of input or predictor variables. Apart from the label, there is also each record which has information about the label. Data mining in this research was carried out by processing student graduation datasets.

Of the several studies on student graduation that have been described, they are studies with the same dataset, while there are several researchers who have conducted research on student graduation with different datasets and have different results. The following is previous research which can be seen in table 1.

Table 1. Research on student graduation with different datasets

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12973
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Writer	Algorithm	Accuracy Result
(Rohmawan, 2018)	Decision Tree	74,51%
	Neural Network	79,74%
(Pratama, Wihandika, & Ratnawati, 2018)	SVM	80,55%
(Mustafa, Ramadhan, & Thenata, 2018)	C.45	87,50%
(Yalidhan, 2018)	NN	98,97%
(Widaningsih, 2019)	C4.5	75,96%
	NAÏVE BAYES	76,79%
	KNN	68,05%
	SVM	74,04%
(Syarli & Muin, 2016)	NAÏVE BAYES	94,00%
(Salmu & Solichin, 2017)	NAÏVE BAYES	80,72%

In the student graduation research table with different data, there are several methods that have been carried out by DT, NN, SVM, C.45, Naïve Bayes (NB). So of the several methods, the highest accuracy is NN with an accuracy of 98.97%. However, research on student graduation with the same data has the following accuracy values

Table 2. Research on student graduation with the same dataset

Writer	Algorithm	Accuracy Result
(Riyanto et al., 2019)	Decision Tree	84,96%
	SVM	84,68%
	Neural Network	86,57%
(Suhardjono et al., 2019)	SVM	85,84%
	SVM+PSO	86,57%
(Ridwansyah et al., 2020)	SVM	85,81%
	SVM+GA	86,43%
(Hendra et al., 2020)	DT	86,55%
	DT+PSO	87,56%

From the research table, the three methods DT, NN, SVM only NN which was not optimized, therefore a trial was carried out to optimize the NN method with the PSO method. This method will be used to predict student graduation by carrying out a process to systematically acquire knowledge based on physical evidence. In predicting the neural network method is used because it has the advantage of working effectively, where in data preprocessing by converting data into a format that is easier to process and effective (Satapathy, Chittineni, Krishna, Murthy, & Reddy, 2012). Coupled with the PSO Algorithm which has the advantage that the algorithm can be adjusted in the selection of parameters and is very easy to implement (Widiastuti, Santoso, & Supriyanto, 2014). PSO can also provide good classification and increase the accuracy of values (I. Ariyati et al., 2020). PSO performance by updating from iteration to iteration using a population of particles in finding the most optimal solution for each particle (Ridwansyah, Ariyati, & Faizah, 2018).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i4.12973 p-ISSN : 2541-044X

METHOD

At the research stage, it uses a model designed with the classification optimization algorithm Particle Swarm Optimization or abbreviated as PSO by optimizing the Neural Network (NN) method. In this test, to find out the best model designed between PSO optimization on NN or without PSO optimization on NN. The dataset being tested uses a dataset of private tertiary institutions in the country of Indonesia where data collection is directly sourced from primary data not from second parties or public data. Where in the processing of the NN and PSO models can be seen from the research flow by briefly explaining the NN and PSO methods which can be seen in Fig 1.

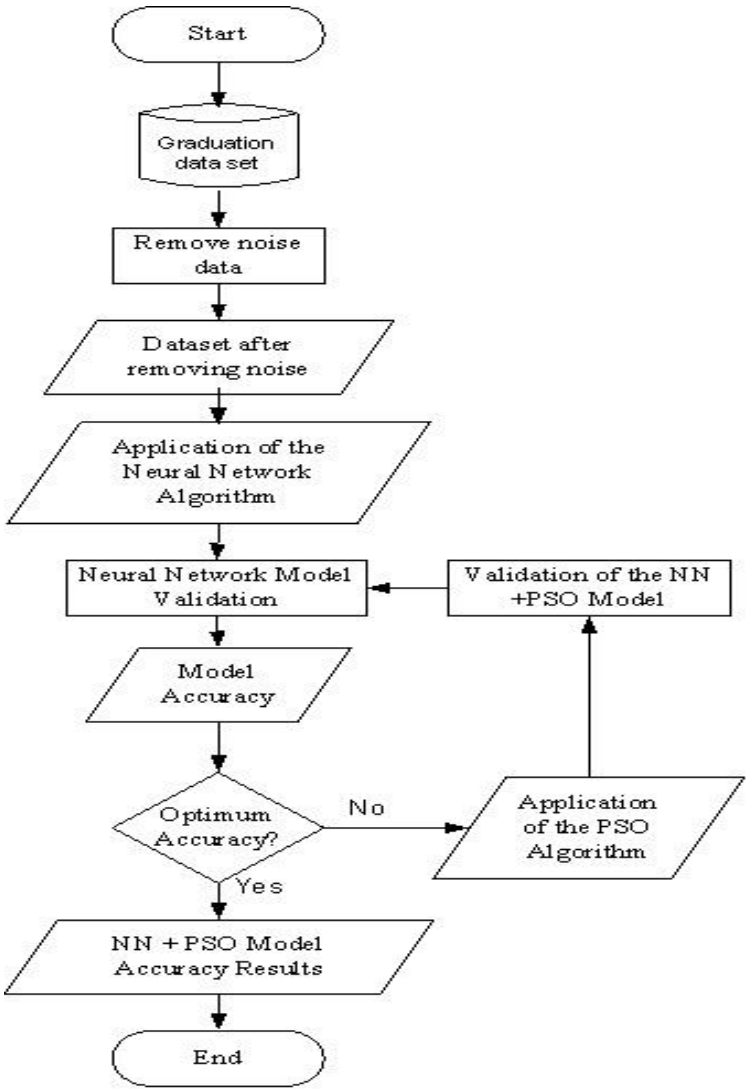

Fig. 1 Student graduation research flow diagram

In the process carried out using the neural network method, firstly data is collected based on primary data taken from student graduation data in 2013-2018. The number of graduation data parameters obtained was 796 and had 9 attributes as predictors and 1 attribute as the result of the predictor. Where the results have a label as to whether they passed on time or not. The data has been processed to remove noisy data, so that the processed data which will later be tested using the PSO-based NN and NN algorithms has no noise data. The NN algorithm is tested first to find out the results of the algorithm, if the results of the algorithm are not optimal for accuracy, it will be tested again with PSO. And if testing with PSO gets optimal results, you will get validation results for the NN algorithm optimized with PSO.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12973
e-ISSN : 2541-2019
p-ISSN : 2541-044X

RESULT

In the results of the research flowchart in Figure 1, the data is collected and obtained from primary data. Where this data is data that will later be tested using the NN and PSO methods after going through the noise data removal process, so that the data is clean from unused data after removing the noise data. Where is the data after removing the noise data that we can see in table 1 for the sample data obtained.

Table 1. Sample data of students who graduated

JK	JRS_SLTA	ORIGIN SLTA	IPK1	IPK2	IPK3	IPK4	IPK5	IPK6	ON TIME
1	4	4	2,55	2,61	2,64	2,74	2,97	3	YES
1	9	3	3,23	3,15	3,07	3,17	3,25	3,17	YES
2	1	3	2,73	2,73	2,62	2,64	2,89	3	YES
1	4	3	2,73	3,1	2,97	3,16	3,3	3,33	YES
1	5	2	2,18	2,44	2,11	2,36	2,66	2,57	NO
2	9	3	2,32	2,39	2,43	2,74	2,67	2,95	NO
2	6	5	3,64	3,56	3,59	3,53	3,53	3,51	YES
1	7	3	2,64	2,22	2,3	2,11	2,49	2,74	YES
1	7	5	2,5	2,71	2,57	2,58	2,75	2,94	NO
2	4	4	2,91	3,07	2,84	2,95	3,16	2,98	NO
1	9	5	3,09	2,88	2,89	3,02	3,1	3,15	YES
1	7	3	2,86	2,88	3,1	3,11	3,14	3,16	YES
1	8	5	3,64	3,71	3,8	3,83	3,77	3,77	YES
1	5	3	2,91	3,05	2,74	2,7	2,65	2,69	NO
2	5	2	2,36	2,17	2,08	2,46	2,89	2,91	YES
1	4	3	2,73	2,83	2,9	2,88	2,88	2,91	YES
1	9	5	3,45	3,37	3,46	3,49	3,49	3,56	YES
2	4	3	2,68	2,8	2,95	3,09	3,13	3,14	YES
2	4	3	3,32	3,49	3,56	3,56	3,63	3,61	YES
2	4	2	3,09	3,12	2,98	3,05	3,35	3,4	YES
1	4	3	3,41	3,29	3,13	3,1	3,1	3,18	YES

From the data table 1 the data will be tested and validated with a neural network model using rapid miner software with 10 cross validation steps. And you will get the confusion matrix results which can be seen in table 2.

Table 2. Confusion Matrix NN

	true YES	true NO	class precision
Pred YES	594	83	87.74%
pred. NO	40	79	66.39%
class recall	93.69%	48.77%	

From the confusion matrix table it can be concluded that 594 students who were predicted to graduate on time were in line with predictions, however 83 students who were predicted to graduate on time turned out to be failures. Likewise, 79 students were predicted not to graduate on time and it was in accordance with the predicted results, however 40 students were predicted not to graduate on time and it turned out that these students graduated on time. So the confusion matrix table produces an accuracy value of 84.55% and produces a bar chart of graduating students which can be seen in Fig 2.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12973
e-ISSN : 2541-2019
p-ISSN : 2541-044X

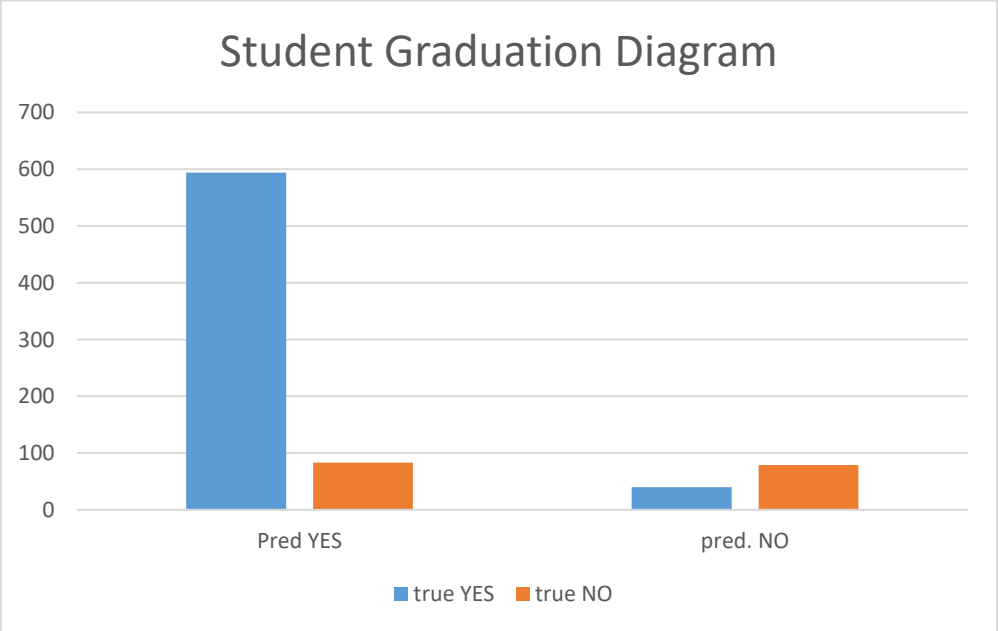

Fig. 2 Student prediction bar chart with NN

Fig. 2 explains that the blue diagram is the result of a yes prediction in the sense of passing on time, and the orange diagram depicts the result of a prediction of not passing on time. From table 2 it also produces the area under curve or abbreviated (AUC) which can be seen in Fig. 3

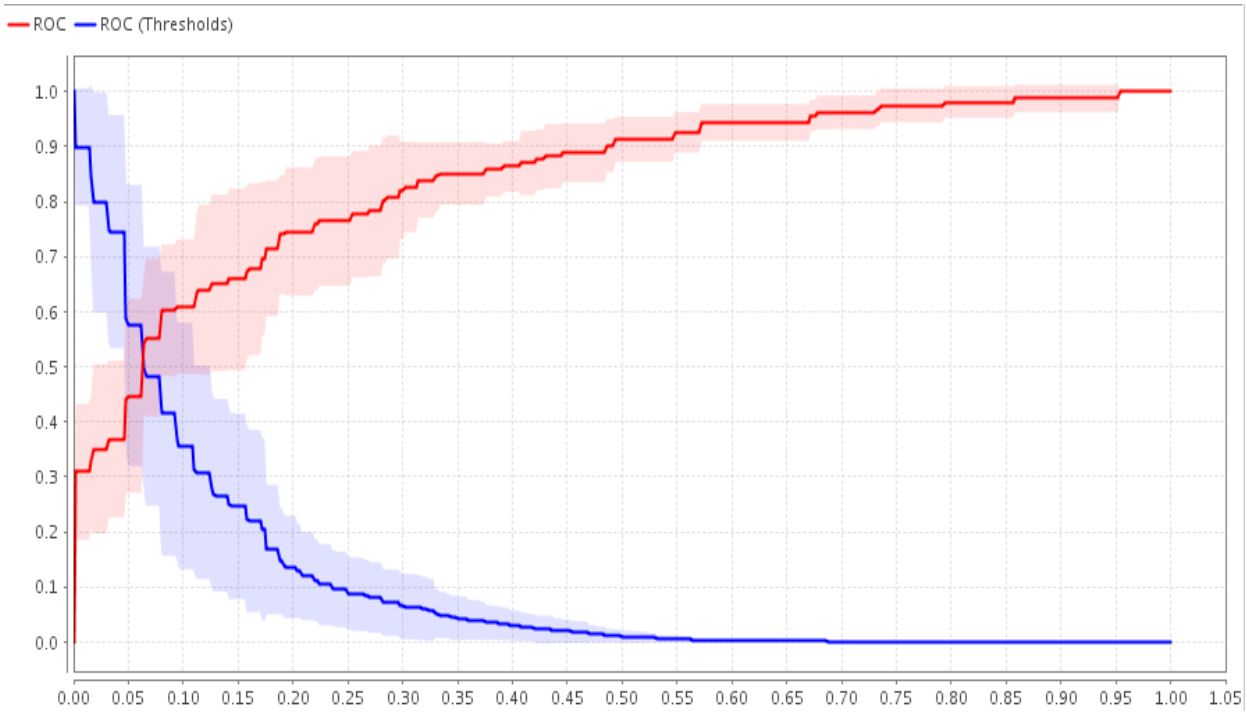

Fig. 3 Area Under Curve (AUC) NN

In Figure 3, it can produce an AUC of 0.845% or it can be interpreted that the AUC obtained very good results. And by testing student graduation data using the NN method, the neural network model architecture seen in Figure 4 will be produced.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12973
e-ISSN : 2541-2019
p-ISSN : 2541-044X

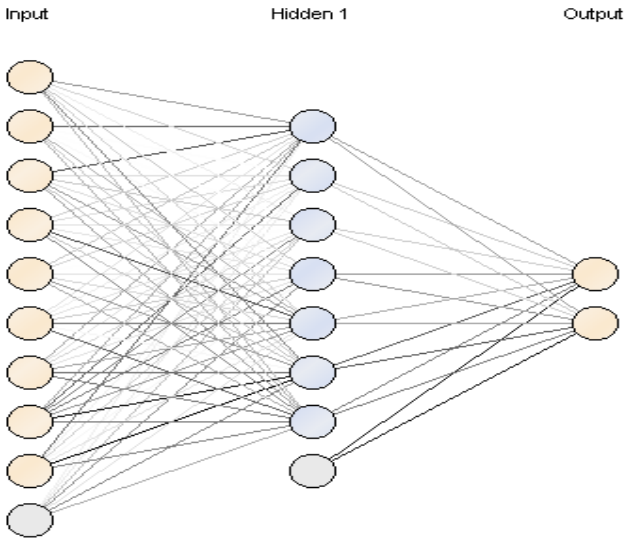

Fig. 4 Architecture of student graduation with NN method

In Figure 4, it can be seen that in the input layer there are 10 indicators, namely gender, high school major, high school origin, 1st semester GPA to 6th semester GPA and 1 threshold. The hidden layer uses 1 hidden layer with the weight value of each attribute in the input layer, as well as 2 outputs, namely passing on time or not passing on time.

In Fig. 4, it can be seen that in the input layer there are 10 indicators, namely gender, high school major, high school origin, 1st semester GPA to 6th semester GPA and 1 threshold. The hidden layer uses 1 hidden layer with the weight value of each attribute in the input layer, as well as 2 outputs, namely passing on time or not passing on time.

Table 3. Atribute Weights (Optimize Weight)

attribute	weight
JK	0
JRS_SLTA	1
ASAL_SLTA	0.962
IPK1	0
IPK2	0
IPK3	1
IPK4	1
IPK5	1
IPK6	0.563

From table 3 it can be concluded that the attributes that have no effect are the attributes that have a value of 0 where there are three attributes that have a value of 0, namely JK (Gender), IPK1 (1st semester GPA) and IPK2 (2nd semester GPA). From the results of NN with PSO, you will get the confusion matrix results which can be seen in table 4.

Table 4. Confusion Matrix NN PSO Optimization

	true YES	true NO	class precision
Pred YES	605	75	88.97%
pred. NO	29	87	75.00%
class recall	95.43%	53.70%	

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 4, October 2023
DOI : https://doi.org/10.33395/sinkron.v8i4.12973
e-ISSN : 2541-2019
p-ISSN : 2541-044X

From the confusion matrix table, it can be concluded that 605 students were predicted to graduate on time according to predictions, however 73 students who were predicted to graduate on time turned out to be failures. Likewise, 87 students were predicted not to graduate on time and it was in accordance with the predicted results, however 29 students were predicted not to graduate on time and it turned out that these students graduated on time. So the confusion matrix table produces an accuracy value of 86.94% and produces a bar chart of graduating students which can be seen in Fig. 5.

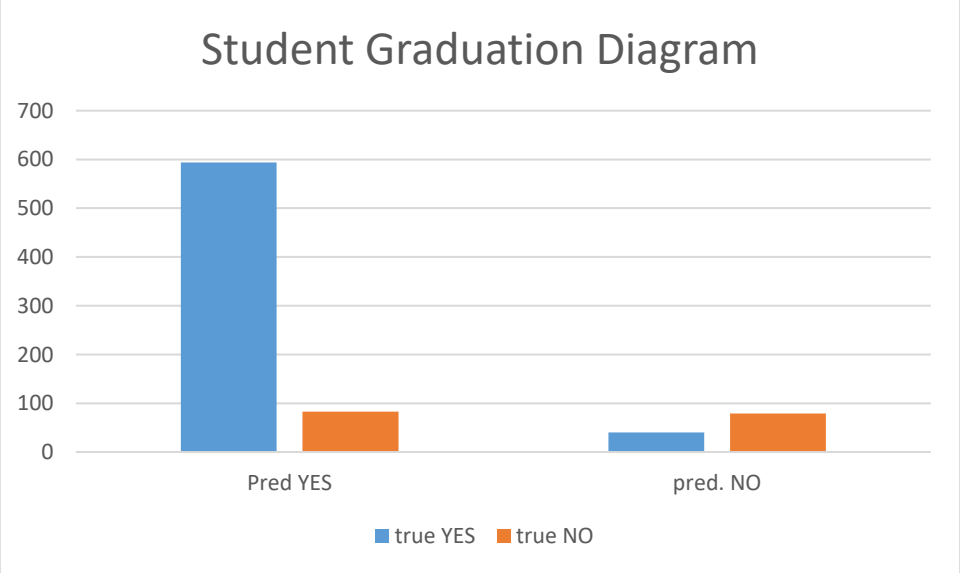

Fig. 5 Bar chart predicting student graduation with NN Optimal PSO

Figure 2 explains that the blue diagram is the result of a yes prediction in the sense of passing on time, and the orange diagram depicts the result of a prediction of not passing on time. From table 2 it also produces the area under curve or abbreviated (AUC) which can be seen in Fig. 6.

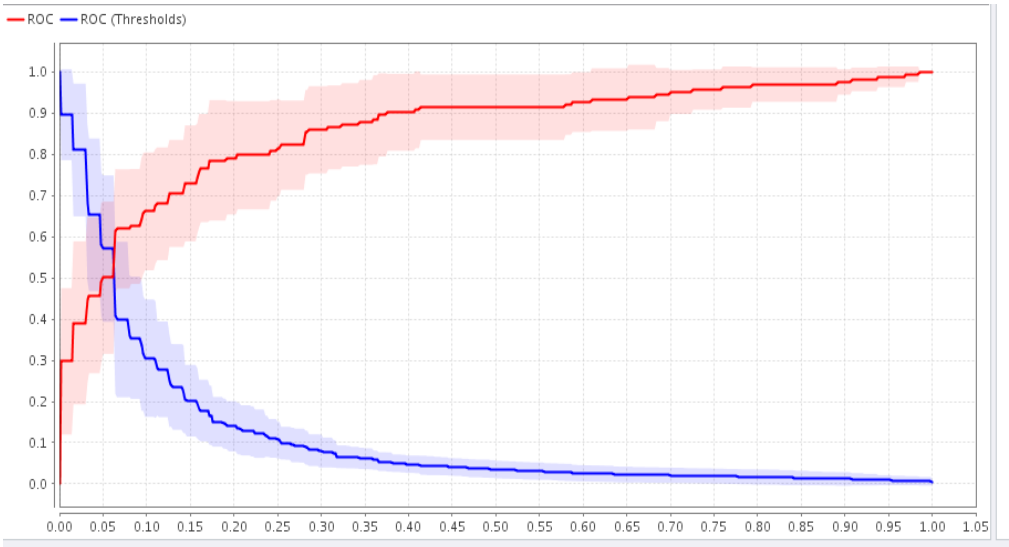

Fig. 3 Area Under Curve (AUC) NN PSO Optimization

In Fig. 3, it can produce an AUC of 0.860% or it can be interpreted that the AUC obtained very good results.

*name of corresponding author

DISCUSSIONS

Based on the results of the testing process carried out on the student graduation dataset in implementing the neural network method with particle swarm optimization optimization, it results that the proposed method is the most accurate and effective type of method technique used to implement the neural network model. The neural network produced an accuracy of 84.55% and an AUC of 0.845% after using the particle swarm optimization method, producing a higher accuracy value compared to the neural network method and these results were very optimal with an accuracy value of 86.94% and an AUC of 0.860%. So that this student graduation data can be used by universities to analyze current student data.

CONCLUSION

The process of testing student graduation data is carried out using neural network and particle swarm optimization methods. The process begins with collecting student graduation data in 2013-2018. From this data, 9 parameters were obtained as predictors and 1 as the predictor result. After that the data will be removed based on incomplete data or what is also called noise data. After the noise data is removed, the data will be tested using the neural network method. After testing the neural network method, the data will be optimized again using the particle swarm optimization method and produce optimal attribute weights so that non-optimal attributes are removed. And the evaluation results of the two methods by looking at the confusion matrix table show that the neural network method that has been optimized with particle swarm optimization gets good accuracy results.

REFERENCES

Ariyati, I., Rosyida, S., Ramanda, K., Riyanto, V., Faizah, S., & Ridwansyah. (2020). Optimization of the Decision Tree Algorithm Used Particle Swarm Optimization in the Selection of Digital Payments. *Journal of Physics: Conference Series*, *1641*(1). https://doi.org/10.1088/1742-6596/1641/1/012090

Ariyati, Indah, Ridwansyah, & Suhardjono. (2018). Implementasi Particle Swarm Optimization untuk Optimalisasi Data Mining Dalam Evaluasi Kinerja Asisten Dosen. *JIKO (Jurnal Informatika Dan Komputer) STMIK AKAKOM*, *3*(2), 70–75.

Hendra, Azis, M. A., & Suhardjono. (2020). ANALISIS PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN DECISSION TREE BERBASIS PARTICLE SWARM OPTIMIZATION. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, *9*(1), 102–107. https://doi.org/https://doi.org/10.32736/sisfokom.v9i1.756

Iqbal, M., Herliawan, I., Ridwansyah, & Gata, W. (2020). Implementation of Particle Swarm Optimization Based Machine Learning Algorithm for Student Performance Prediction. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, *6*(2), 195–204. https://doi.org/10.33480/jitk.v6i2.1695.IMPLEMENTATION

Mustafa, M. S., Ramadhan, M. R., & Thenata, A. P. (2018). Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier. *Creative Information Technology Journal*, *4*(2), 151. https://doi.org/10.24076/citec.2017v4i2.106

Nasyuha, A. H., Jama, J., Abdullah, R., Syahra, Y., Azhar, Z., Hutagalung, J., & Hasugian, B. S. (2021). Frequent pattern growth algorithm for maximizing display items. *Telkomnika (Telecommunication Computing Electronics and Control)*, *19*(2), 390–396. https://doi.org/10.12928/TELKOMNIKA.v19i2.16192

Pratama, A., Wihandika, R. C., & Ratnawati, D. E. (2018). Implementasi Algoritme Support Vector Machine (SVM) untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2*(March), 1704–1708.

Ridwansyah, Ariyati, I., & Faizah, S. (2018). Particle Swarm Optimization Berbasis Co-Evolusioner. *Jurnal SAINTEKOM*, *9*(2), 166–177. Retrieved from http://stmikplk.ac.id/jurnal/index.php/saintekom/article/view/96/62

Ridwansyah, & Purwaningsih, E. (2018). Particle Swarm Optimization Untuk Meningkatkan Akurasi Prediksi Pemasaran Bank. *Jurnal PILAR Nusa Mandiri*, *14*(1), 83–88.

*name of corresponding author

Ridwansyah, R., Wijaya, G., & Purnama, J. J. (2020). Hybrid Optimization Method Based on Genetic Algorithm for Graduates Students. *Jurnal Pilar Nusa Mandiri*, *16*(1), 53–58. https://doi.org/10.33480/pilar.v16i1.1180

Riyanto, V., Hamid, A., & Ridwansyah. (2019). Prediction of Student Graduation Time Using the Best Algorithm. *Indonesian Journal of Artificial Intelligence and Data Mining*, *2*(2), 1–9. https://doi.org/http://dx.doi.org/10.24014/ijaidm.v2i1.6424

Rohmawan, E. P. (2018). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Decision Tree dan Artificial Neural Network. *Jurnal Ilmiah MATRIK Vol.20 No.1, April 2018:21-30*, *20*(1), 21–30.

Salmu, S., & Solichin, A. (2017). Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naïve Bayes : Studi Kasus UIN Syarif Hidayatullah Jakarta. *Seminar Nasional Multidisiplin Ilmu (SENMI) 2017*, (April), 701–709.

Satapathy, S. C., Chittineni, S., Krishna, S. M., Murthy, J. V. R., & Reddy, d P. V. G. D. P. (2012). Kalman particle swarm optimized polynomials for data classification. *Appl. Math. Model*, *36*(1), 115–126.

Suhardjono, Wijaya, G., & Hamid, A. (2019). PREDIKSI WAKTU KELULUSAN MAHASISWA MENGGUNAKAN SVM BERBASIS PSO. *Bianglala Informatika*, *7*(2), 97–101.

Syarli, S., & Muin, A. (2016). Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). *Jurnal Ilmiah Ilmu Komputer*, *2*(1), 22–26.

Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, *13*(1), 16–25. https://doi.org/10.36787/jti.v13i1.78

Widiastuti, N. A., Santoso, S., & Supriyanto, C. (2014). ALGORITMA KLASIFIKASI DATA MINING NA¨IVE BAYES BERBASIS PARTICLE SWARM OPTIMIZATION UNTUK DETEKSI PENYAKIT JANTUNG PARTICLE SWARM OPTIMIZATION UNTUK DETEKSI PENYAKIT JANTUNG. *Pseudocode*, *1*(1).

Yalidhan, M. D. (2018). Implementasi Algoritma Backpropagation Untuk Memprediksi Kelulusan Mahasiswa. *Klik - Kumpulan Jurnal Ilmu Komputer*, *5*(2), 169. https://doi.org/10.20527/klik.v5i2.152

*name of corresponding author