Jurnal Mantik, 7 (2) (2023), ISSN 2685-4236 (Online)

Published by:Institute of Computer Science (IOCS)

Jurnal Mantik

Journal homepage: www.iocscience.org/ejournal/index.php/mantik

Implementation of artificial neural network and support vector machine algorithm on student graduation prediction model on time

Yennimar Yennimar[1], M. Rafi Faturrahman[2], Siwa Nesen[3], M. Anhar Guci[4], Samuel Rifaldi Pasaribu[5]

[1,2,3,4,5]Faculty of Science and Technology, Informatics Engineering, Universitas Prima Indonesia, Medan, Indonesia

A R T I C L E I N F O

Article history:

Accepted Jul 25, 2023
Revised Jul 27, 2023
Accepted Jul 31, 2023

Keywords:

Algorithm;
Accuracy;
Classification;
SVM;
ANN.

ABSTRACT

This research aims to evaluate how Artificial Neural Network (ANN) and Support Vector Machine (SVM) algorithms can be used to predict student graduation on time. This research uses student data from Universitas Prima Indonesia (UNPRI) Medan to build a prediction model. ANN and SVM methods have been applied and compared to see the performance of each model. The test results show that the SVM model is superior in terms of accuracy and computational speed compared to the ANN model. In addition, the test results also show that the SVM model can be used to predict student graduation on time with an accuracy of 96.34%. This result shows that the SVM model is more effective in predicting student graduation on time compared to the ANN model.

Corresponding Author:

Name of Corresponding Yennimar,
Faculty of Science and Technology / Informatics Engineering,
Prima Indonesia University,
Jl. Belanga No.1, Medan, Sumatera Utara, 20113, Indonesia.
Email: yennimar@unprimdn.ac.id

1. INTRODUCTION

The Implementation of artificial neural network and support vector machine algorithm on student graduation prediction model on time is one of the methods used to predict student graduation on time. This method is done by using ANN and SVM algorithms that make it possible to predict with high accuracy. ANN and SVM algorithms are methods used to solve classification problems using supervised learning techniques. The SVM algorithm uses a linear transformation to solve classification problems. Where the SVM algorithm turns the classification problem into a linear optimization problem. The SVM algorithm uses kernel values to transform the available data into a form that is easier to read and understand by the SVM algorithm.

Meanwhile, ANN is a supervised learning method that uses a network of neurons to solve classification problems. This neuron network serves as a link between input and output data. The ANN algorithm comprises the input layer, the hidden layer, and the output layer as its three primary parts.

The data used to test this method is prima indonesia university student data. This data contains information about students, such as IPS semester 1 to 5. This data will be used to build a prediction model for student graduation on time.

After the data has been obtained, the ANN and SVM algorithms will be used to build a prediction model for on-time student graduation. The ANN and SVM algorithms will process the data and produce a prediction model for on-time student graduation. This model will be used to predict student graduation on time with high accuracy.

After the on-time student graduation prediction model is successfully built, this model will be used to predict student graduation on time. This model will help universities to determine students who are eligible to be accepted in the study program they want. With this prediction model, students can know whether they will graduate on time or not before they register. This model can also be used to help students who want to graduate on time by giving them the right strategies to achieve that goal. Thus, based on the results of previous research reviews, the research title " Implementation of Artificial Neural Network and Support Vector Machine Algorithm to the Prediction Model of Timely Student Graduation" was raised.

This research aims to develop a prediction model capable help predict student graduation on time. Furthermore, this research will identify factors that can affect the delay in student graduation. This research will also explain how the Artificial Neural Network (ANN) and Support Vector Machine (SVM) algorithms can be used as a prediction model to predict student graduation on time. To evaluate the developed model, this research will assess how accurate the resulting prediction is. Finally, the results of this research are expected to help universities to increase the on-time graduation rate of students by improving the support provided to students who are at high risk of delayed graduation.

Previous research on predicting student graduation has involved various machine learning methods such as classification using Naive Bayes, Decision Trees, and k-Nearest Neighbors (k-NN) algorithms, as well as the application of regression methods such as Linear Regression or Logistic Regression. In addition, some studies may have used other algorithms such as Random Forest and Gradient Boosting Machines. Research has also focused on specific aspects of graduation, including drop-out prediction, graduation rates per major, and factors affecting timely graduation. In the context of research addressing the application of Artificial Neural Network and Support Vector Machine Algorithms on Timely Student Graduation Prediction Models, the difference lies in the focus of applying two main algorithms, namely ANN and SVM, as well as the use of unique datasets and methodological approaches. Thus, this research is expected to provide new relevant contributions in the field of student graduation prediction.

According to I. T. Utami. (2022) "Comparison of Support Vector Machine (SVM) Classification Performance and Binary Logistic Regression in Classifying Student Graduation Timeliness. Get the results that the test conducted with the SVM algorithm on a dataset consisting of 100 student data has a training data percentage rate of 94.8%.

According to Emy Haryatmi1, Sheila Pramita Hervianti. (2021) in a study entitled Implementation of the Support Vector Machine Algorithm for the Prediction Model of Timely Student Graduation, got test 1 results using a percentage of 90% training data and 10% testing data, the training data accuracy value was 92.5%, the testing data accuracy value was 94.4%, the testing data precision value was 94.5%, the testing data recall value was 94.5% and the testing data F-measure value was 94.5%.

According to Musa Hendri Janto Rahanra 1), Kusrini 2), Emha Taufiq Luthfi (2022) in a study entitled Analysis of Timely Graduation of Informatics Engineering Students Using the Artificial Neural Network (ANN) Algorithm for student graduation prediction models From this research, researchers used sample data 100 samples divided by 80% training data and 20% testing data can produce a prediction of graduation in the

informatics engineering study program of 0.90% prediction using an algorithm using Artificial Neural Network (ANN).

2. RESEARCH METHOD

In the research used, quantitative research is a type of research that aims to apply existing knowledge and methods to real situations or specific contexts. In this case, the research focuses on the Implementation of artificial neural network and support vector machine algorithm on student graduation prediction model on time. This research conducted several stages of work procedures. The work procedures carried out are as follows:

2.1 Dataset

Datasets are collections of data that have been organized and are generally patent for each input. In this study, a dataset is needed in the form of some student data to be studied, namely ips students from semester 1 to 5.

2.2 Methods

a. Support Vector Machine (SVM)

Support Vector Machine (SVM) is an algorithm for machine learning that is used to solve classification and regression issues.. SVM functions by constructing a hyperplane or "decision boundary" that maximizes the separation between two different data classes. The data are divided into many classes using this hyperplane..

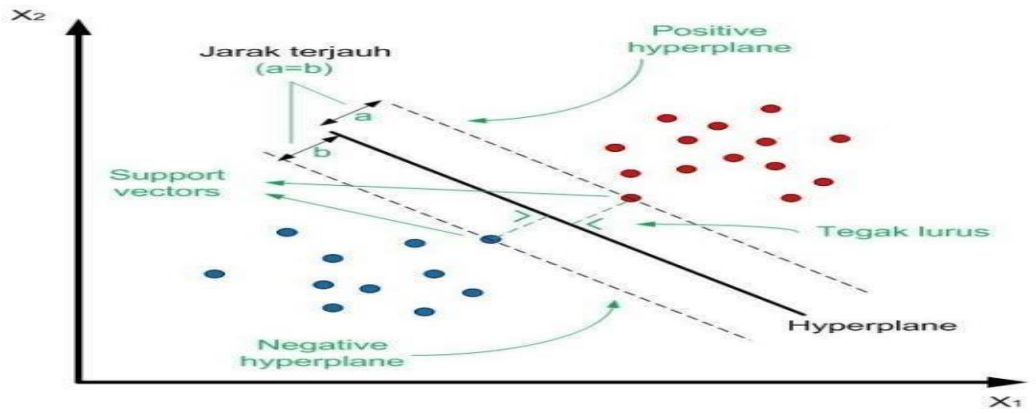

Figure 1. Support Vector Machine (SVM)

b. Artificial Neural Network (ANN)

ANN (Artificial Neural Network) algorithm is a mathematical model inspired by human biological neural networks. ANN is used in machine learning to model and learn complex patterns in data. The algorithm consists of several layers that are connected through adjustable weight connections.

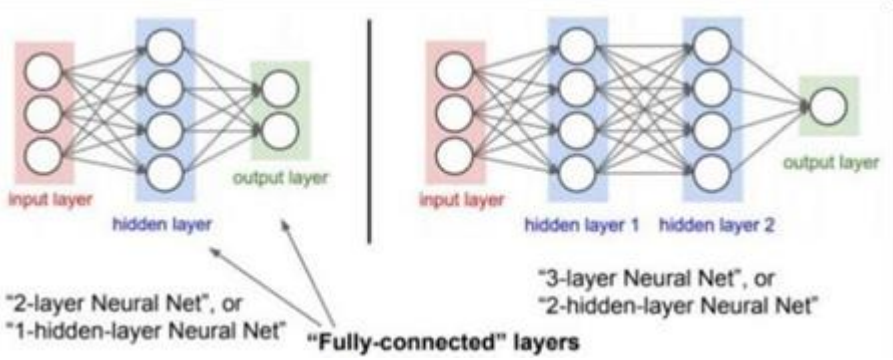

Figure 2. Artificial Neural Network (ANN)

The following are the stages of the framework or flowchart of the algorithm used to predict whether or not graduation is on time.

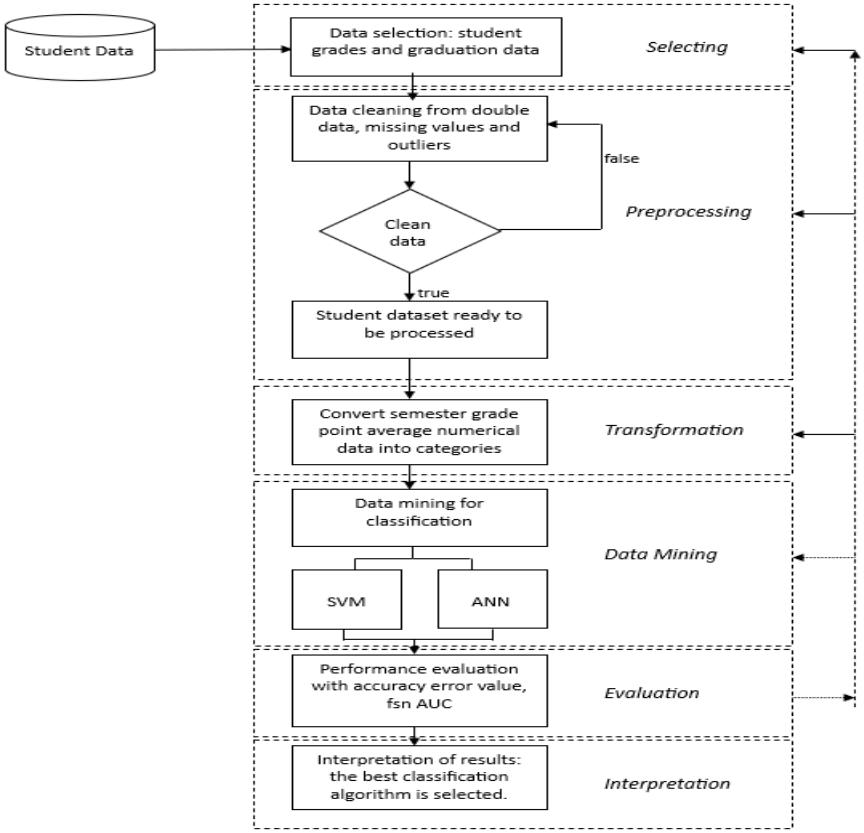

Figure 3. flowchart of svm and ann algorithm process

3. RESULTS AND DISCUSSIONS

The initial format of student data is CSV (Comma Separated Values), which is a common standard for storing datasets. However, data analysis cannot be done using this format. We converted the student data from CSV format to Jupyter Notebook Dataframe format as a result. Dataframes are matrices that describe tabular data in a way that makes it

simpler for us to visualize and understand the information. We can handle data in a variety of ways, including filtering and preprocessing, by using dataframes. Overall, the method of student data entry assisted us in getting the data ready for additional analysis.

```
[ ] total_data = len(data)
    print("Jumulh total data keseluruhannya : ",total_data)

    Jumulh total data keseluruhannya :  277

    data.head(10)
```

	No	NIM	NAMA	IPS (semester 1)	IPS (semester 2)	IPS (semester 3)	IPS (semester 4)	IPS (semester 5)	kelulusan
0	1.0	193303030001	Sallyana Yek	4.0	3.74	3.55	3.90	4.00	lulus
1	2.0	193303030002	Anthony	4.0	3.74	4.0	3.85	3.95	tidak lulus
2	3.0	193303030003	William Chandra	3.45	3.61	3.27	3.79	3.64	tidak lulus
3	4.0	193303030004	Valentino Damar	4.0	3.61	3.89	3.65	3.80	tidak lulus
4	5.0	193303030005	Nicholas Livando	4.0	3.74	3.94	3.80	3.77	tidak lulus
5	6.0	193303030006	Alvin Rasid	4.0	3.61	4.0	3.90	4.00	tidak lulus
6	7.0	193303030009	Matthew Evan Phanie	4.0	3.61	3.89	3.90	4.00	tidak lulus
7	8.0	193303030011	Michael	3.75	3.43	3.83	3.47	3.91	tidak lulus
8	9.0	193303030012	Windy Candra	3.5	3.74	3.9	3.70	4.00	tidak lulus
9	10.0	193303030014	Steven Lie	3.69	3.26	3.89	4.00	3.59	tidak lulus

Figure 4. dataset of ips grades of science and technology faculty students

3.1 Data Preprocessing

Data preprocessing is an important stage in research, when processing and analysis preparation of the gathered data is required. On the student data in this instance, the researcher carried out a label encoding procedure. To make the process of data analysis easier, categorical data (such pass and fail) are converted into integers through the label encoding procedure. Pass, for instance, can be represented as 1, whereas not pass, can be encoded as 0. This makes it easier for the research's Machine Learning analysis method to comprehend and analyse the data. In order to prepare student data for further analysis, researchers can use the label encoding procedure, which results in Dataframes that are ready for processing as illustrated in the following figure.

```
data_clean.head(10)
```

	IPS (semester 1)	IPS (semester 2)	IPS (semester 3)	IPS (semester 4)	IPS (semester 5)	kelulusan
0	4.00	3.74	NaN	3.90	4.00	1
1	4.00	3.74	NaN	3.85	3.95	0
2	3.45	3.61	3.27	3.79	3.64	0
3	4.00	3.61	3.89	3.65	3.80	0
4	4.00	3.74	3.94	3.80	3.77	0
5	4.00	3.61	NaN	3.90	4.00	0
6	NaN	3.61	3.89	3.90	4.00	0
7	3.75	3.43	3.83	3.47	3.91	0
8	3.50	3.74	3.90	3.70	4.00	0
9	3.69	3.26	3.89	4.00	3.59	0

Figure 5. Dataset of Ips Grades of Students of the Faculty of Science and Technology in CSV Form

3.2 Data Split

In this research, student data must be separated into two significant pieces after being entered and preprocessed. The second 30% is used as test data, also known as test data, while the first 70% is used as training data, also known as train data. The goal of this is to guarantee that the algorithm can function with data that has never been seen before.

```
# Membagi dataset menjadi data latih dan data uji => data latih 70% data testing 30%
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3)

data_latih = len(X_train)
data_latih_y = len(y_train)
print("Jumlah data latih X_train = ",data_latih)
print("Jumlah data latih y_train = ",data_latih_y)
```

```
Jumlah data latih X_train =  191
Jumlah data latih y_train =  191
```

```
data_testing = len(X_test)
data_testing_y = len(y_test)
print("Jumlah data testing X_test = ",data_testing)
print("Jumlah data testing y_test = ",data_testing_y)
```

```
Jumlah data testing X_test =  82
Jumlah data testing y_test =  82
```

Figure 6. Data Split

3.3 Model Training

Model traning is the method of improving a model's parameters for improved model performance. However, the SVM (Support Vector Machine) technique is used in this research's model to determine the findings of the diagnosis. The SVM model is trained with train data that has already been separated into 70% components. By providing training data and utilizing the technique to train, this model is then implemented. The model will be prepared to predict student data using test data once training is finished.

```
print("Berikut laporan metrics untuk data training \n",report_data_latih)
```

```
Berikut laporan metrics untuk data training
              precision    recall  f1-score   support

           0       0.89      0.61      0.72        28
           1       0.94      0.99      0.96       163

    accuracy                           0.93       191
   macro avg       0.92      0.80      0.84       191
weighted avg       0.93      0.93      0.93       191
```

Figure 7. SVM Training Model

After that, looking for the confusion matrix value, the results of the confusion matrix value of SVM can be seen in the figure below

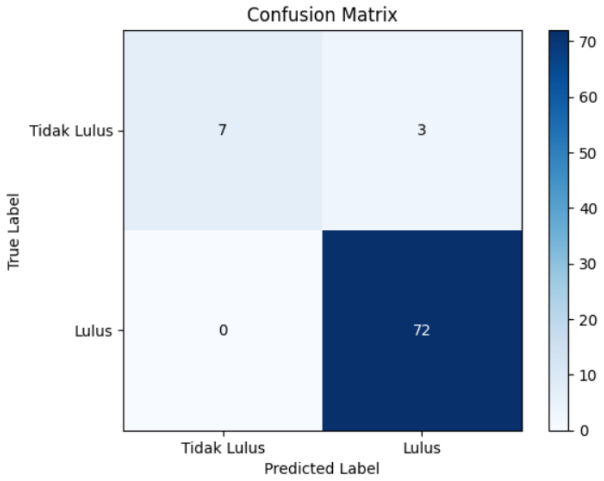

Figure 8. SVM confusion matrix

Model traning is the method of improving a model's parameters for improved model performance. However, in this research, the model in determining the diagnosis results uses the ANN (Artificial Neural Network) algorithm. Train data that has previously been divided into 30% parts is used to train the ANN model. By providing training data and utilizing the technique to train, this model is then implemented. The model will be prepared to predict student data using test data once training is finished.

Laporan metrics data latih dengan menggunakan model AAN

	precision	recall	f1-score	support
0	0.92	0.41	0.57	29
1	0.90	0.99	0.95	162
accuracy			0.91	191
macro avg	0.91	0.70	0.76	191
weighted avg	0.91	0.91	0.89	191

Figure 8. ANN Training Model

Next, the confusion matrix value is searched using ANN The results of this search can be seen in the figure below.

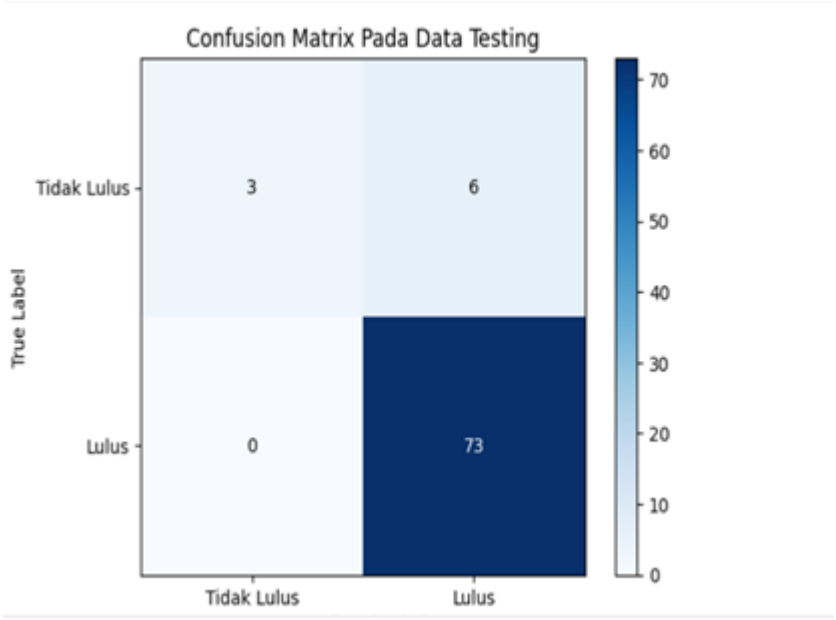

Figure 9. Confusion matrix ANN

In the research, the accuracy results from the Implementation of the Artificial Neural Network algorithm ranged from 92.68 (92%) and the accuracy results from the Implementation of the Support Vector Machine algorithm ranged from 96.34 (96%).

ANN

	precision	recall	f1-score	support
0	1.00	0.33	0.50	9
1	0.92	1.00	0.96	73
accuracy			0.93	82
macro avg	0.96	0.67	0.73	82
weighted avg	0.93	0.93	0.91	82

SVM

	precision	recall	f1-score	support
0	1.00	0.70	0.82	10
1	0.96	1.00	0.98	72
accuracy			0.96	82
macro avg	0.98	0.85	0.90	82
weighted avg	0.96	0.96	0.96	82

Figure 10. accuracy of ANN and SVM

4. CONCLUSION

In this research, based on a recent literature study, it was found that SVM algorithm has a higher accuracy rate than ANN algorithm in predicting students' on-time graduation. However, many studies still use SVM because ANN has a strong ability to model complex

relationships between input features and output targets. This research proves that ANN can also be used to predict on-time or off-time graduation, achieving an accuracy of 93%. However, when using SVM, the accuracy increased to 96%, demonstrating the superiority of the SVM algorithm in achieving higher accuracy rates in this context. With this research, educational institutions can provide timely support to students who need help, and the results can inspire further research on the application of machine learning technology in education. In addition, the findings of this research can be adopted by other educational institutions as a model for implementing machine learning techniques to improve students' timely graduation, and increase the efficiency and effectiveness of efforts in managing students' academic progress.

REFERENCES

Anwar, K., Hanafiah, H., & Ebun, A. (n.d.). *Predicting Student Graduation Using Artificial Neural Network: A Preliminary study of Diploma In Accountancy Program at UiTM Sabah*.

Baashar, Y., Alkawsi, G., Mustafa, A., Alkahtani, A. A., Alsariera, Y. A., Ali, A. Q., Hashim, W., & Tiong, S. K. (2022). Toward predicting student's academic performance using artificial neural networks (ANNs). *Applied Sciences*, *12*(3), 1289.

Bangsa, R. A. P.-S. A. (2016). Seleksi Atribut Pada Metode Support Vector Machine Untuk Menentukan Kelulusan Mahasiswa E-Learning. *EVOLUSI: Jurnal Sains Dan Manajemen*, *4*(1).

Bangun, O., Mawengkang, H., & Efendi, S. (2022). Metode Algoritma Support Vector Machine (SVM) Linier Dalam Memprediksi Kelulusan Mahasiswa. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, *6*(4), 2006–2013.

Darmawan, A., Yudhisari, I., Anwari, A., & Makruf, M. (2023). Pola Prediksi Kelulusan Siswa Madrasah Aliyah Swasta dengan Support Vector Machine dan Random Forest. *Jurnal Minfo Polgan*, *12*(2), 387–400.

Haryatmi, E., & Hervianti, S. P. (2021). Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(2), 386–392.

Haryatmi, E., & Pramita Hervianti, S. (2021). Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(2), 386–392. https://doi.org/10.29207/resti.v5i2.3007

Isro'Mukti, Y. (n.d.). *Sistem Prediksi Lulus Tepat Waktu Tugas Akhir Mahasiswa Menggunakan Support Vector Machine (SVM)*.

Kurniawan, A. S. (2018). Implementasi Metode Artificial Neural Network Dalam Memprediksi Hasil Ujian Kompetensi Kebidanan (Studi Kasus Di Akademi Kebidanan Dehasen Bengkulu). *Pseudocode*, *5*(1), 37–44. https://doi.org/10.33369/pseudocode.5.1.37-44

Mukti, Y. I. (2020). Sistem Prediksi Lulus Tepat Waktu Tugas Akhir Mahasiswa Menggunakan Support Vector Machine (Svm). *JUTIM (Jurnal Teknik Informatika Musirawas)*, *5*(2), 110–115.

Musa Hendri Janto Rahanra, Kusrini, E. L. (2022). Analisa Kelulusan Mahasiswa Teknik Informatika Tepat Waktu Menggunakan Algoritma Artificial Neural Network (ANN). *JURNAL FATEKSA: Jurnal Teknologi Dan Rekayasa*, 1–11.

Ofori, F., Maina, E., & Gitonga, R. (2020). Using machine learning algorithms to predict students' performance and improve learning outcome: A literature based review. *Journal of Information and Technology*, *4*(1), 33–55.

Qisthiano, M. R. (2022). Klasifikasi Terhadap Prediksi Kelulusan Mahasiswa Dengan Menggunakan Metode Support Vector Machine (SVM). *Seminar Nasional Teknologi Dan Multidisiplin Ilmu (SEMNASTEKMU)*, *2*(1), 204–208.

Ridwan, R., Lubis, H., & Kustanto, P. (2020). Implementasi algoritma neural network dalam memprediksi tingkat kelulusan mahasiswa. *Jurnal Media Informatika Budidarma*, *4*(2), 286–293.

Rohmawan, E. P. (2018). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Desicion Tree Dan Artificial Neural Network. *Jurnal Ilmiah MATRIK*, *20*(1), 21–30.

Romadhona, A., Suprapedi, & Himawan, H. (2017a). Prediksi Kelulusan Mahasiswa Tepat Waktu Berdasarkan Usia, Jenis Kelamin, dan Indeks Prestasi Menggunakan Algoritma Decision Tree. *Jurnal Teknologi Informasi*, *13*(1), 69–83. http://research.

Romadhona, A., Suprapedi, S., & Himawan, H. (2017b). Prediksi Kelulusan Mahasiswa Tepat Waktu Berdasarkan Usia, Jenis Kelamin, Dan Indeks Prestasi Menggunakan Algoritma

Decision Tree. *Jurnal Cyberku*, *13*(1), 8.
Romlah, U. H., & Solichin, A. (2022). Prediksi Kelulusan Mahasiswa Dengan Metode Naive Bayes dan Artificial Neural Network: Studi Kasus Fakultas Teknik UNIS Tangerang. *Faktor Exacta*, *15*(1).
Rudianto, R., Kania, R., & Solihati, T. I. (2022). Prediksi Kelulusan Mahasiswa Teknik Informatika Universitas Banten Jaya Menggunakan Algoritma *Neural Network*. *Jurnal Sistem Informasi Dan Informatika (Simika)*, *5*(2), 193–200.
Setiyani, L., Syamsudin, A., Gintings, A., & Arifin, D. (2020). The Analysis of Functional Needs on Undergraduate Thesis Information System Management. *International Journal of Advances in Data and Information Systems*, *1*(2), 50–59. https://doi.org/10.25008/ijadis.v1i2.184
Suhaimi, N. M., Abdul-Rahman, S., Mutalib, S., Hamid, N. H. A., & Hamid, A. (2019). Review on predicting students' graduation time using machine learning algorithms. *International Journal of Modern Education and Computer Science*, *11*(7), 1–13.
Suhardjono, Ganda, W., & Abdul, H. (2019). Prediksi Kellusan Menggunakan Svm Berbasis Pso. *Bianglala Informatika*, *7*(2), 97–101.
Supriyadi, E., & Sensuse, D. I. (2015). Optimasi Algoritma Support Vector Machine Dengan Particle Swarm Optimization Dalam Mendeteksi Ketepatan Waktu Kelulusan Mahasiswa : *Snit 2015*, 163–174. http://seminar.bsi.ac.id/snit/index.php/snit-2015/article/view/131
Thaniket, R., Kusrini, & Luthf, E. T. (2019). Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Algoritma Support Vector Machine. *JURNAL FATEKSA: Jurnal Teknologi Dan Rekayasa*, *13*(2), 69–83.
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.
Yandi Saputra, A., & Primadasa, Y. (2018). Penerapan Teknik Klasifikasi Untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbour Implementation of Classification Method to Predict Student Graduation Using K-Nearest Neighbor Algorithm. *Techno.Com*, *17*(4), 9.
Zulfa, M. I., Fadli, A., & Ramadhani, Y. (2019). Classification model for graduation on time study using data mining techniques with SVM algorithm. *AIP Conference Proceedings*, *2094*(1).