



Analysis of Student Graduation Classification Using the Decision Tree Method

Rahmadeni^{1*}, Muhammad Marizal², Rado Yendra³

^{1,2,3} Department of Mathematics, Faculty of Science & Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau

ARTICLE INFO	ABSTRACT
Published Online: 06 January 2026	Student graduation is one of the important indicators in assessing the success of the academic process in higher education. Low graduation rates can have an impact on the quality of institutions and the effectiveness of the education system. Therefore, an analysis method is needed that is able to accurately classify student graduation status based on available academic data. This study aims to analyze the graduation classification of students using the Decision Tree method with the help of the RapidMiner application. The data used is in the form of student academic data that has gone through the pre-processing stage, including data cleaning and transformation. The Decision Tree method was chosen because it produces a classification model that is easy to understand through clear decision rules. Model evaluation is carried out using a confusion matrix to measure classification performance. The results of the study show that the Decision Tree method is able to classify student graduation with a good level of accuracy and identify attributes that affect graduation. This research is expected to be the basis for decision-making for academics in an effort to increase student graduation rates.
Corresponding Author: Rahmadeni	
KEYWORDS: Student Graduation, Classification, Data Mining, Decision Tree, RapidMiner.	

1. INTRODUCTION

Universities have an important role in producing quality human resources through a structured educational process. One of the indicators of the success of the implementation of higher education is the graduation rate of students, especially on-time graduation. A low graduation rate can reflect a problem in the academic process that needs serious attention. Therefore, the use of student academic data is important to support objective and systematic graduation analysis [1].

The development of information technology encourages the use of large amounts of academic data to support decision-making in the field of education. Data mining is an approach that is able to explore patterns and important information from these data sets. One of the main techniques in data mining is classification, which aims to group data into specific categories based on its characteristics. This approach is considered effective in assisting academic evaluation and education policy planning [2].

Method Decision Tree is one of the classification algorithms that is widely used in the analysis of student academic data. This method is able to produce decision rules that are easy to understand and interpret. In addition, Decision Tree It can also identify the main attributes that affect student graduation. These advantages make Decision Tree suitable for use in the analysis of student graduation [3].

The use of supporting software plays an important role in the successful implementation of data mining methods. The RapidMiner app provides a visual interface that makes it easy to process and model data. RapidMiner also supports a variety of machine learning algorithms, including Decision Tree, making the model evaluation process easier. The use of RapidMiner is considered to be able to increase the effectiveness of academic data classification analysis [4].

In addition, the results of the classification of student graduation not only function as a prediction tool, but can also be used as a basis for academic evaluation. Information about the attributes that have the most influence on graduation can help universities in formulating strategies to improve the quality of education. Student academic assistance can be carried out in a more targeted manner based on the results of the classification. Therefore, this study aims to analyze the classification of student graduation using the Decision Tree with the help of the RapidMiner app [5].

2. THEORETICAL FOUNDATIONS

2.1 Data Mining

Data mining is the process of extracting hidden patterns and knowledge from large data sets using statistics, mathematics, and artificial intelligence techniques. In the field of education, data mining is known as Educational Data Mining (EDM),

which aims to analyze student academic data to support academic decision-making, such as graduation prediction and student performance evaluation. The use of data mining allows educational institutions to identify patterns of student learning behavior based on historical data, so that early academic intervention can be carried out [6]

The classification technique in data mining is used to group data into certain classes based on the attributes it has. In the context of student graduation, classification aims to predict whether a student will graduate on time or not based on academic variables such as GPA, exam scores, and the number of credits that have been taken [7].

2.2 Klasifikasi

Classification is one of the methods of Supervised Learning in data mining, where the model is built based on training data that already has a class label. The classification model is used to predict classes from new data whose category is not yet known. The classification process is very suitable for the problem of predicting student graduation because the graduation status has been clearly defined, such as graduating on time and not graduating on time.

According to research [8] Classification of student academic data is able to provide fairly accurate prediction results if the attributes used are relevant and the data pre-processing process is carried out properly. Therefore, the selection of the right classification algorithm is an important factor in the success of student graduation analysis.

2.3 Metode Decision Tree

A Decision Tree is a classification method that forms a prediction model in a tree structure consisting of root nodes, branches, and leaf nodes. Each node represents an attribute, while each branch indicates the value of that attribute, and the leaf node represents the result of classification. The Decision Tree method is widely used because it is able to generate decision rules that are easy for users to understand and interpret.

Decision Tree algorithms like C4.5 work by selecting the best attributes based on the value of information gain or gain ratio to divide the data into the most optimal classes. In student graduation research, the Decision Tree is effective in processing numerical and categorical attributes such as GPA, social studies, and course scores to determine students' graduation status [9].

2.4 RapidMiner

RapidMiner is a data mining software that is widely used in academic research to carry out the classification process visually and in a structured manner, including in predicting student graduation using the Decision Tree algorithm. In research [10] RapidMiner is used to implement the C4.5 algorithm to build a student graduation prediction model based on complete academic data, such as GPA and Social Studies, thus simplifying the process of analyzing and evaluating the model.

2.5 Student Graduation Classification

Student graduation classification is one of the applications supervised by learning in educational data mining, which aims to predict student graduation status based on historical academic data, such as Semester Achievement Index (IPS), Cumulative Achievement Index (GPA), length of study, and other relevant variables. This technique is used to assist universities in conducting academic evaluations and planning learning strategies, and minimizing delays in the study period.

Research [11] indicates that the algorithm Decision Tree C4.5 is effective in predicting the admission of students of the Faculty of Information Technology, ISB Atma Luhur. The model is built through the pre-processing stages of data, model formation, Decision Tree, and evaluation using matrix confucius, producing high enough accuracy so that it can be used as a basis for academic decision-making.

In addition, the research [12] Apply the Decision Tree to classify the admission time of Ihsan Gorontalo University students. The results of the study show that the algorithm Decision Tree is able to predict student admissions with an accuracy of above 90%, so this method is suitable for use as a tool for evaluating the learning process and planning educational resources.

3. Research Methods

3.1 Types and Approaches to Research

This study uses a quantitative approach with a data mining method. The quantitative approach was chosen because the research focuses on processing and analyzing students' academi data in numerical form. The data mining method is used to explore patterns and information from students' academic data to classify graduation status. The classification technique used in this study is the Decision Tree algorithm.

3.2 Data Sources and Types

The data used in this study is secondary data, namely, student academic data obtained from university academic databases. The data includes several attributes relevant to student graduation, such as Cumulative Grade Point Average (GPA), the number of credits that have been taken, the length of the study period, and the status of the course. The target variable in this study is the graduation status of students, which is classified into graduates and non-graduates.

3.3 Research Stages

The research stages are carried out systematically, which includes data collection, pre-data processing, modeling, and evaluation of results. In the pre-processing stage, data cleansing is carried out to overcome lost, inconsistent, or duplicate data. Next, the data is transformed into a format that is appropriate for the modeling process. The modeling stage was carried out by applying the Decision Tree algorithm using the RapidMiner application.

3.4 Modeling Using Decision Trees

Classification modeling is done using the Decision Tree algorithm available in the RapidMiner application. This

algorithm works by forming a decision tree structure based on the attributes that have the highest information value. Each node on the tree represents an attribute, while a branch indicates the value of the attribute, and a leaf indicates the class of the classification result. The resulting model is used to classify the graduation status of students.

3.5 Model Evaluation

The evaluation of the classification model was carried out using a confusion matrix. A confusion matrix is used to measure model performance based on accuracy, precision, and recall values. The results of this evaluation are used to

assess the extent to which the Decision Tree model is able to classify student graduation appropriately. Models with high accuracy values are considered to have good classification performance.

3.6 Tools and Software

This study uses the Rapid Miner application as a tool in the process of processing and analyzing data. RapidMiner is used to pre-process data, for Decision Tree modeling, and for the evaluation of classification results. The use of RapidMiner was chosen because of the ease of use and completeness of features that support the data mining process.

4. RESULTS AND DISCUSSION

4.1 Descriptive Data

Table 1. Actual Data

IPS 3	IPS 4	IPS 5	IPS 6	IPS 7	IPS 8	GPA	STATUS KELULUSAN
3,23	2,41	3	2,47	1,75	0	2,75	PRECISE
3,42	2,85	3,31	2,95	2,18		3,39	PRECISE
2,93	3,14	2,92	2,64	2,88	0,5	2,81	PRECISE
2,09	2,55	2,55	2,43	2,55	2,17	2,82	PRECISE
3,31	2,83	3,36	2,73	3,06	0	3,09	PRECISE
3,17	3,25	3,41	3,08	3,43	3	3,23	PRECISE
2,23	2,86	2,25	2,64	1,52	2,1	2,54	PRECISE
3,54	3,78	3,42	3,88	2,5	4	3,56	PRECISE
3,48	3,7	3,19	3,21	2,58	4	3,4	PRECISE
2,11	3,36	2,79	3,17	2,45	0	2,97	PRECISE
2,79	2,73	3,02	2,59	3,08	0	3,08	PRECISE
3,02	3,63	3,21	3,09	2,57	3	3,33	PRECISE
2,54	3,36	2,79	3,28	3,28	3	3,15	PRECISE
3,02	3,07	2,88	2,76	3	2,5	2,92	PRECISE
3,31	3,33	3,3	2,94	3,46	0	3,26	PRECISE
0,16	0,32	1,13	0,37	0,28	0,11	1,02	PRECISE
3,69	3,41	3,4	0	0	0	3,43	PRECISE
2,66	2,91	3,25	2,88	2,83	0	2,86	PRECISE
2,11	3	2,73	2,5	2,48	0	2,68	PRECISE
3,21	2,92	3,41	2,86	2,5	0	3,04	PRECISE
1,95	2,37	3,07	2,91	1,73	1,6	2,81	PRECISE
3,71	3,83	3,25	3,65	2,33	3,5	3,56	PRECISE
3,6	3,52	2,9	3,32	3,21	0	3,45	PRECISE
3,79	3,71	3,65	2,1	3,59	4	3,52	PRECISE
2,93	3,08	3,29	3,18	2,1	3	3,1	PRECISE
1,95	2,17	1,95	2,18	1,45	1,08	2,34	PRECISE
2,19	0,44	2,17	0,59	0,26	0	1,72	PRECISE
3,56	3,83	3,73	3,74	2,33	3,5	3,58	PRECISE
3,44	3,38	3,33	3,36	2,08	0	3,1	PRECISE
2,77	3,59	3,5	3,35	3,27	0	3,31	PRECISE
3,31	3,25	3,44	3,22	3,07	3,5	3,33	PRECISE
3,27	2,96	3,16	3,12	2,61	0	3,08	PRECISE
3,1	2,96	3,2	3,53	2,48	3	3,27	PRECISE

“Analysis of Student Graduation Classification Using the Decision Tree Method”

3,42	3,29	3,33	2,75	3	0	3,13	PRECISE
3,96	3,91	3,75	3,94	2,58	4	3,85	PRECISE
3,2	3	3,21	2,47	2,55	2,75	3,1	PRECISE
2,59	2,77	3,09	2,83	3,11	2,9	3,08	PRECISE
3,39	3,27	2,98	2,85	2,12	0	3,21	PRECISE
3,73	3,42	3,5	3,47	2,58	3,5	3,52	PRECISE
3,73	3,41	3,46	2,94	2,65	3	3,47	PRECISE
3,41	3,43	3,37	3,25	3,16	0	3,22	PRECISE
3,5	3,39	3,67	3,15	1,5	0	3,25	PRECISE
2,95	3,14	2,94	3,14	3,34	2,75	3,02	PRECISE
2,94	3,27	3,29	2,75	3,43	0	3,08	PRECISE
3,81	3,61	3,35	3,43	3,46	4	3,49	PRECISE
2,57	2,82	2,64	2,33	2,34	1,8	2,86	PRECISE
2,27	2,68	2,59	2,36	2,61	3	2,72	PRECISE
2,36	2,95	2,84	2,68	2,77	2,92	2,96	PRECISE
2,61	2,95	2,86	2,68	3,16	2,5	2,92	PRECISE
2,2	2,77	3,16	2,53	2,95	2,9	3,07	PRECISE
2,59	3,05	3,02	2,95	2,86	3	3,11	PRECISE
3	3,05	2,96	3,42	2,25	3	3,09	PRECISE
3,34	3,54	3,19	3,81	2,58	3	3,35	PRECISE
2,59	3,09	2,79	3,63	2,4	3,5	2,99	PRECISE
2,3	2,59	2,68	2,33	1,28	1,5	2,62	PRECISE
2,93	2,86	3,16	2,3	2,65	0,67	2,79	PRECISE
3,55	2,82	3,14	2,73	3,3	3	3,17	PRECISE
2,8	3,09	3,11	2,98	3,22	3	3,19	PRECISE
1,82	2,5	1,95	2,56	1,64	0,71	2,46	PRECISE
2,8	2,95	3,18	3,1	2,5	3	3,03	PRECISE
2,52	2,86	2,2	2,83	2,73	0	2,62	PRECISE
2,07	2,27	2,52	2,38	2,41	2,6	2,73	PRECISE
2,55	2,59	2,82	2,65	3	2,25	2,85	PRECISE
2,11	2,32	2,48	2,33	1,43	1,5	2,53	PRECISE
3,57	2,93	3,16	2,9	3,3	3	3,24	PRECISE
2,3	2,82	2,64	2,77	3,02	2,5	2,93	PRECISE
2,18	2,36	2,32	2,45	2,23	2,1	2,73	PRECISE
2,07	2,68	2,64	2,68	2,64	2,75	2,83	PRECISE
3,13	3,17	3,19	3,4	2,23	3	3,28	PRECISE
2,3	2,77	2,55	2,48	2	4	2,52	PRECISE
2,55	3	2,68	2,58	3,02	3	2,98	PRECISE
3	3,41	3,15	3,31	3,32	0	3,09	PRECISE
2,16	1,8	2,11	1	1,19	0,79	2,12	PRECISE
2,37	2,09	2,41	2,36	1,26	1,6	2,41	PRECISE
2,39	2,82	2,7	3,06	2,7	3,5	2,9	PRECISE
2,27	2,55	2,45	2,39	2,55	0,75	2,7	PRECISE
2,64	2,68	2,68	2,68	2,67	3	2,9	PRECISE
2,7	2,77	2,75	3	2,93	2,5	3,09	PRECISE
2,95	2,86	2,7	3,06	2,92	3	3	PRECISE
3,48	3,26	3,15	3,33	3,68	0	3,32	PRECISE

“Analysis of Student Graduation Classification Using the Decision Tree Method”

2,25	2,27	2,21	2,16	2,86	2,89	2,8	PRECISE
2,57	2,91	2,84	3,08	3,35	0	2,97	PRECISE
2,86	2,86	2,64	2,83	1,57	0	2,57	PRECISE
2,02	2,45	2,36	2,43	2,8	2,8	2,72	PRECISE
2,89	3,14	2,85	3,42	3,03	3,5	2,94	PRECISE
2,34	2,86	2,23	2,75	2,39	3,69	2,75	PRECISE
2,64	2,59	2,8	2,75	1,14	0	2,46	PRECISE
3,25	2,83	2,86	2,97	1,31	0	2,88	PRECISE
3,11	3,04	2,75	2,83	2,32	3	3,02	PRECISE
2,38	3,59	3,56	2,9	3,73	3	3,52	PRECISE
2,73	3,27	2,77	2,53	3,25	3	2,99	PRECISE
2,23	2,36	1,66	1,56	2,14	2,08	2,39	PRECISE
3,04	3,14	2,93	3	3,63	0	3,13	PRECISE
3,25	3,13	3,4	3,28	3,25	0	3,36	PRECISE
3,27	3,61	3,44	3,4	2,25	0	2,83	PRECISE
2,68	2,36	2,25	3,03	2,86	0	2,89	LATE
3,02	2,75	2,66	2,17	3,28	1	3,5	PRECISE
3,6	3,83	3,17	3,7	2,6	3,5	3,05	PRECISE
3,11	3,09	3,16	3,05	3,36	3	2,97	PRECISE
2,68	3	3,25	3,08	3,18	0		PRECISE
3,13	3,25	3,29	3,58	1,91	4	3,24	PRECISE
3,13	3,25	3,29	3,58	1,91	4	3,53	PRECISE
3,52	3,57	3,25	3,79	2,5	4		PRECISE
2,94	3,14	2,61	2,33	2	0	3,08	PRECISE
2,43	3,14	2,73	2,85	2,9	1,8	3,14	PRECISE
3,52	3,52	3,15	2,91	1,29	3,5	3,24	PRECISE
3,25	3,14	3,19	2,7	3,41	3,5	3,28	PRECISE
3,25	3,25	3,31	3,48	2,43	3	3,28	PRECISE
2,77	2,86	3,05	3	3,56	2,4	2,86	PRECISE
2,52	2,91	2,81	2,41	2,03	0,25	3,16	PRECISE
3,29	3,35	2,69	3,22	2,08	3	1,64	PRECISE
3,5	3,22	2,77	2,94	2,25	3	3,02	PRECISE
3,08	2,96	2,59	3,19	3,38	3		PRECISE
3,5	3,85	3,78	3,63	0	0	3,63	PRECISE
3,27	3,52	2,77	3,55	2,64	3,5	3,37	PRECISE
3,27	3,52	2,77	3,55	2,64	3,5	2,93	PRECISE
2,39	3,18	2,85	3,25	3,15	0	2,82	PRECISE
2,39	2,91	2,93	2,63	1,64	0	3,2	PRECISE
2,88	3,36	3,02	3,39	3,31	0	3,32	PRECISE
3,06	3,68	3,48	3,47	3,5	0	3,16	PRECISE
3,1	3,15	3	3,06	3,43	0	2,65	PRECISE
2,7	2,73	2,8	3,33	2,43	0	2,83	PRECISE
2,7	2,73	2,8	3,33	2,43	0	1,9	PRECISE
2,45	2,14	3,24	3,41	3,2	4	3,1	PRECISE
2,45	2,14	3,24	3,41	3,2	4	2,81	PRECISE
2,44	2,41	2,4	2,75	2,08	2,8	3,13	PRECISE
2,48	2,45	2,95	3,48	3,91	0	2,87	PRECISE

2,73	3	2,48	2,75	2,21	0	0,87	PRECISE
3,42	3,43	3,25	3,09	2,42	3,5	3,13	PRECISE
2,52	2,86	2,91	3,5	3,19	3	3,13	PRECISE
3,06	2,88	3,34	3,25	2,53	4	2,92	PRECISE
2,36	2,55	2,11	2,5	1,35	0	3,09	LATE
3,27	3,38	3,21	2,81	3,03	0	0,92	PRECISE
2,94	3,36	3,15	2,77	3,47	0	3,15	PRECISE
2,61	3,23	2,71	3,45	3,11	3	3,31	PRECISE
3,09	3,14	2,54	3,31	1,63	3,5	2,88	PRECISE
3,25	3,21	3,02	3,31	2,25	3	3,3	PRECISE
3,17	2,91	3,05	3,21	3,54	0	3,28	PRECISE
3,33	3,17	3,17	3,26	2,33	3,5	2,79	PRECISE
2,48	3,14	2,9	2,4	2,73	0	2,85	PRECISE
2,14	2,77	2,61	2,93	2,82	2,5	0,99	PRECISE
2,55	3	2,98	2,8	3,14	3	2,97	PRECISE
2,55	3	2,98	2,8	3,14	3	3,03	PRECISE
3,25	2,87	3	2,94	3,09	3	3,16	PRECISE
3	3,23	2,79	3	2,41	3	2,16	PRECISE

The processed student data is used to build a student classification model. Each attribute plays a determining role in predicting graduation status, while the student status attribute is determined as a class label in the classification process.

4.2 Modeling

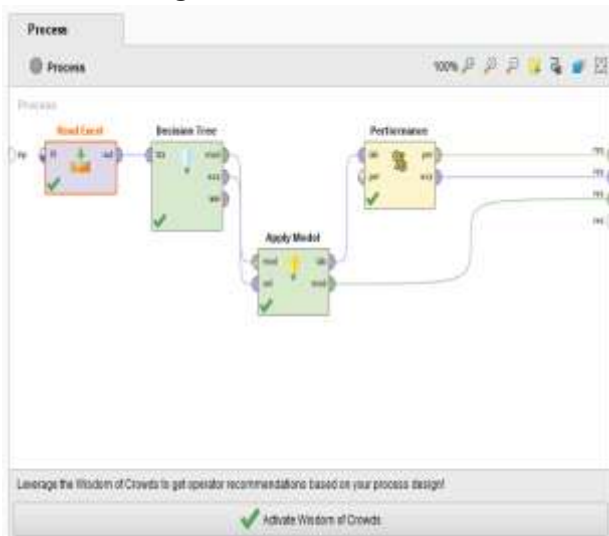


Figure 1. Modeling from Decision Tree Preprocessing

The image above shows the flow of the student admissions classification process using the RapidMiner application. The process begins by reading student data using the read Excel operator. Furthermore, the data is processed using a decision tree algorithm to build a classification model based on available academic attributes. The model that has been formed is then applied to the data using the apply model operator to obtain predicted results of student permit status. The final stage is an evaluation of the performance of the

model using a performance operator that produces accuracy, precision, and recall values based on the confusion matrix.

4.3 Evaluation

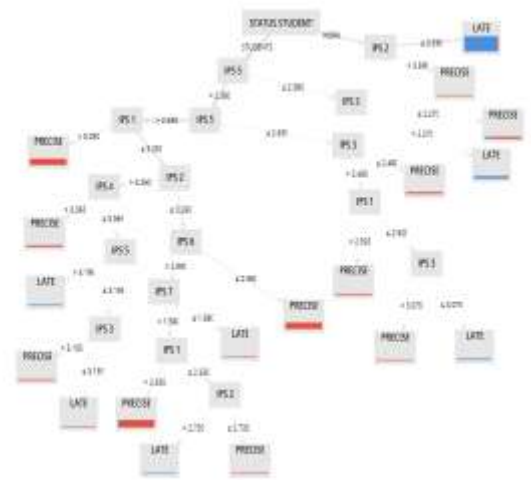


Figure 2. Decision Tree Structure

The figure above illustrates the structure of the decision tree used to classify student status based on the Semester Achievement Index (IPS) score. In this model, the main attribute used as a separator is the social studies score in a given semester, which is chosen because it has a significant influence on the student's admission status. The classification process begins with the grouping of students based on their social studies scores. Students who have social studies scores above a certain threshold tend to be directed to the branch to graduate on time, while those who are below the threshold are directed to the branch that graduates late.

Furthermore, the branching process is continued by considering the social studies score in the following semesters (social studies 2, social studies 3, and so on), thus adding

dimension to the analysis. Each leaf node on the decision tree indicates the final result of the classification, where the node labeled "PRECISE" represents students who have stable academic performance and tend to be high, while the "LATE" node indicates students who are at risk of falling behind.

Based on this analysis, it can be concluded that good academic performance, reflected in high and stable social studies scores, is essential to predict timely achievement. Conversely, low or fluctuating IPS values can increase the risk of delays. Therefore, this decision tree model not only serves as a prediction tool, but also as a useful means of analysis to understand the patterns of students' academic behavior. The implementation of this method can assist colleges in academic decision-making and planning, while also providing insight into the support that may be needed for students who are at risk of being late.

5. CLOSING

This research successfully applied the Decision Tree method to analyze and classify student permission status based on available academic data. The results of the evaluation show that Decision Tree can identify significant patterns in the data, as well as provide a good level of accuracy in predicting accuracy in a timely manner. By using the RapidMiner application, the analysis process becomes more efficient and easier to understand, thus supporting better decision-making in an academic environment.

The importance of this research lies in its ability to not only predict the status of applicants, but also provide insight into the factors that affect students' academic performance. By understanding the influential attributes, universities can take strategic steps to improve the quality of education and support students who are at risk of falling behind.

It is hoped that this research can be a basis for further study and development of other analytical methods in the field of education. In addition, support from the academic side is very important to implement research results into academic mentoring programs, so that the goal of improving the quality of education can be achieved optimally.

REFERENCES

1. J. South, "Implementation of Data Mining for Prediction of Timely Graduation of STIMIK ESQ Students Using Decision Tree C4. 5," vol. 5, no. 2, pp. 89–98, 2021.
2. A. Rohman and S. Mujiyono, "Prediction Modeling of Student Graduation Predicate Using Decision Tree C4. 5," vol. II, no. 2, pp. 1–5, 2021.
3. A. Setyawan, A. Fitriani, E. Rilvani, U. P. Bangsa, and K. Bekasi, "CLASSIFICATION OF POVERTY IN INDONESIA WITH," vol. 3, no. 7, 2025.
4. Z. Fatah et al., "CLASSIFICATION OF STUDENT GRADUATION USING THE DECISION TREE METHOD," vol. 1, no. 4, pp. 58–64, 2024.
5. P. A. Prayesy and I. Ruswita, "G-Tech : Journal of Applied Technology," vol. 7, no. 1, pp. 21–28, 2023.
6. Z. Fatah and S. Bella, "Classification of Student Academic Status Using Decision Tree," vol. 4, pp. 141–146, 2025.
7. Z. Hasanah, M., & Fatah, "The Application of Decision Trees in the Classification of Student Graduation," vol. 4, 2025.
8. K. A. Putri, D. Febriawan, and F. N. Hasan, "Implementation of Data Mining to Predict Student Study Period with Decision Tree Algorithm (C4 . 5)," vol. 13, pp. 39–47, 2024.
9. U. Suriani, "The Application of Data Mining to Predict Student Graduation Rate Using Algorithms," vol. 3, no. 2, pp. 55–66, 2023.
10. N. Khasanah, D. Uki, E. Saputri, T. Hidayat, and F. Aziz, "Prediction of Student Graduation Using the C4 Algorithm. 5 with RapidMiner : A Case Study of XYZ College Academic Data," vol. 4, no. 2, pp. 100–107, 2025.
11. I. Widyaningrum, M. Putri, R. Rusdah, L. Suryadi, and D. Anubhakti, "Prediction of Graduation of Students of the Faculty of Information Technology ISB Atma Luhur Using Algorithms," vol. 12, pp. 363–369, 2023.
12. S. Kasus, F. Ilmu, H. Universitas, and I. Gorontalo, "Classification of Student Graduation Time Using the Decision Tree Method," vol. 3, no. 1, pp. 31–36, 2024.