JURNAL INFOTEL
VOL. 17, NO. 1, FEBRUARY 2025, PP. 136–149. DOI:10.20895/INFOTEL.V17I1.1258

RESEARCH ARTICLE

Optimization of Naive Bayes and Decision Tree Algorithms through the Application of Bagging and AdaBoost Techniques for Predicting Student Study Success

Endi Febriyanto[1] and Wasilah[2,*]

[1,2]Institut Informatika dan Bisnis Darmajaya, Bandar Lampung 35142, Indonesia

*Corresponding email: wasilah@darmajaya.ac.id

Received: October 15, 2024; Revised: February 01, 2025; Accepted: February 25, 2025.

Abstract: The percentage of student failure in learning is still relatively high. Many countries, including Ghana, Nigeria, Enugu, and Indonesia, experience this condition. Internal and external factors that vary significantly between students have the potential to be the cause of failure. This condition cannot be allowed to continue. A special analysis is needed on the factors that can help improve student grades. Predictions of student success are urgently needed. These predictions can anticipate negative impacts that occur, including increased risk of dropout, decreased student motivation to learn, and decreased individual potential. The Naive Bayes and Decision Tree algorithms have been used to predict student success. However, despite their advantages, these two algorithms still have several weaknesses. It can cause the algorithm's performance not to be as expected. Several methods in ensemble techniques can improve algorithm performance. Two methods that are often used are Bagging and AdaBoost. Bagging and AdaBoost can help improve the performance of classification algorithms. This study combines Bagging and AdaBoost into the Decision Tree and Naïve Bayes algorithms to optimize the results in predicting student success. The stages are data collection, pre-processing, data split, data processing, and evaluation model. The results show that the Bagging and AdaBoost techniques have been proven to be effective in improving accuracy, precision, recall, and F1-score performance. Combining the Naïve Bayes algorithm with AdaBoost significantly increases accuracy, precision, and F1-score by 1.95%, 28.98%, and 15.79%.

Keywords: AdaBoost, bagging, decision tree optimization, Naive Bayes optimization, student learning success

 ISSN: 2085-3688; e-ISSN: 2460-0997

1 Introduction

Government attention should focus on improving the quality of human resources. Formal schools are an optimally regulated education system expected to create human resources to advance the nation [1]. Predicting student success plays a vital role in improving the effectiveness of education. Educational institutions can provide appropriate interventions to needy students by understanding the factors that influence success. This can reduce dropout rates, increase student retention, and support students at risk of dropping out. Accurate predictions also allow for the development of individualized learning programs, helping students reach their academic potential and improving the overall quality of education.

The percentage of student failure in learning is fairly high. Many countries, including Ghana, Nigeria, Enugu, and Indonesia, experience this condition. WAEC reports consistently show unsatisfactory student performance in the Ashanti region of Ghana in biology [2, 3]. Stakeholders in Nigeria have expressed concern about the poor academic performance of students in all categories of schools in Nigeria. In addition, the examination of students' academic performance in Basic Science in Enugu State, as indicated by the results of the Basic Education Certificate Examination (BECE) for 2018 to 2022, shows an alarming pattern of below average achievement [4]. Many countries experience similar conditions, including Indonesia.

This condition cannot be allowed. A special analysis of the factors that can help improve student grades is needed. The inability to predict student study success can have a significant negative impact. Educational institutions cannot identify student abilities without accurate predictions, which results in the lack of appropriate intervention. This condition can increase the risk of droppingout, reduce student learning motivation, and cause loss of individual potential. Predicting student study success plays a vital role in improving the effectiveness of education. Many factors influence the success of a study. Traditional approaches are often not practical enough. The machine learning approach has many algorithms that can be used to solve prediction problems, including the Naive Bayes and Decision Tree algorithms. The machine learning approach can analyze complex data more deeply and accurately.

Naive Bayes and Decision Tree algorithms have been used to predict student study success [4, 5]. Decision Tree has the advantage of being flexible so that it can improve the quality of the decisions produced [6] Naive Bayes's simplicity of calculations allows for faster and more efficient processes. The Naïve Bayes method only requires relatively small training data to determine the parameters needed in the classification process [7]. However, both algorithms have weaknesses. Decision Trees are prone to overfitting and unstable decisions. At the same time, the Naïve Bayes method only supports attributes with discrete or discretion data types and does not support attributes with continuous (numeric) values, so all attributes become independent. In addition, these attributes can contribute to the predicted attributes [8].

Two methods are often used in ensemble techniques, namely bagging and boosting [9, 10]. Bagging and boosting can support the unstable classification algorithm [11]. This study's formulation of the problem is: How is the application of Bagging and AdaBoost (adaptive boosting) techniques on the Decision Tree and Naïve Bayes algorithms for predicting student success? What are the results of analyzing the best algorithm performance in predicting student success?

Several studies have been conducted on bagging and AdaBoost methods to improve the performance of Decision Tree and Naïve Bayes algorithms. Research using this method has been conducted to predict the success of studies, including the success of students in universities [2,12], and predicting success in distance education [3]. In this study, what will be done is to compare the performance of the Decision Tree and Naïve Bayes classification algorithms with the application of ensemble techniques, namely the Bagging and AdaBoost techniques, to predict student learning success. Furthermore, a comparative analysis of the results of the application of the algorithm will be carried out to determine the performance of the Decision Tree algorithm, Naïve Bayes combined with ensemble techniques in making predictions.

Ensemble techniques have varying performances depending on the characteristics of the dataset and the algorithm used. Several studies have compared the performance of Bagging and AdaBoost using various machine learning algorithms [13]. Ensemble techniques such as Bagging and AdaBoost have been proven to be effective in improving the performance of classification models. Studies have shown that by combining several classification models (such as Naïve Bayes and Decision Tree) using these ensemble techniques, more accurate predictions can be produced [14]. There is research using this method in universities [12, 15], but very few studies are still focused on the classification of student learning outcomes [16]. Generally, research focuses on distance education [3], classifying learning types in education, learning styles [17], and student performance [18]. Research on predicting academic success has been conducted using various data, including those related to learning styles, gender, age, personality, educational background, and others. However, it has not been focused on academic grades in detail. At the same time, academic grades are a factor that is very close to students and has a high potential to have a strong influence. In this study, an ensemble technique will be combined with the Naïve Bayes algorithm and decision tree, using data that focuses on the details of students' academic grades and extracurricular activities. This is a novelty that is expected to improve algorithm performance.

2 Literature Review

Decision trees are a popular and powerful tool in data science and machine learning for classification, regression, and other predictive modeling tasks. They are valued for their simplicity, interpretability, and ability to handle categorical and continuous data. Decision trees work by recursively partitioning the data into subsets based on the value of input features, creating a tree-like model of decisions. The process starts with a root node and splits the data at each node based on specific criteria until reaching leaf nodes, which represent the final decision or classification [19]. The construction of decision trees involves selecting the best attribute to split the data at each node, which can be done using various measures such as Gain Ratio, Gini Index, or other node splitting measures [20].

Naïve Bayes is a popular algorithm used for text classification and sentiment analysis due to its simplicity and efficiency. It operates under the assumption of attribute independence, which can sometimes be violated in real-world data. This analysis explores how well Naïve Bayes models reflect people's views by examining their performance in various applications [21].

Ensemble methods, such as boosting and bagging, are powerful techniques used to improve the accuracy of classifiers by combining multiple models. AdaBoost is particularly well-known for its effectiveness in binary classification tasks. Boosting: This method sequentially trains classifiers, with each new classifier focusing on the errors made by the previous ones. AdaBoost is a popular boosting algorithm that adjusts the weights of misclassified instances, making it highly effective for binary classification tasks. Bagging, or Bootstrap Aggregating, involves training multiple classifiers independently on different subsets of the training data, created through bootstrapping. The final prediction is made by averaging all classifiers' predictions, which helps reduce variance and prevent overfitting [9].

Research on the classification of student graduation using data mining techniques has been conducted extensively. Research [18] using the Naïve Bayes algorithm showed quite good performance: 87% accuracy, 91% precision for the First Class class, 78% recall, and 81% F1-score. The features used in this study include department, level, weekly study time, satisfaction with the learning system, engagement in group discussions, engagement in school policy, attendance rate, engagement in curricular activities, learning method, accommodation type, gender, and age.

Research [15] The Stacking Ensemble method was used to predict student graduation, resulting in 95% accuracy, precision, recall, and F1-score in the range of 91%-93%. Meanwhile, other ensemble methods, such as Bagging and Boosting, generally have accuracy, precision, recall, and F1-score in the range of 68%-73%. The features used in this study were assignments, quizzes, mid-tests, and final tests. Research [22] combined the Random Forest (RF) and eXtreme Gradient Boosting (XGBoost) algorithms and showed a high level of accuracy in predicting students at risk of failure. This model achieved excellent evaluation metrics, with accuracy, precision, recall, and F1-score exceeding 90%. The features used included Lecture Notes, Materials, Video, Live Attendance, and Live Activities. In addition, research [23] comparing Decision Tree (DT) and Artificial Neural Network (ANN) showed similar performance, with an accuracy of 62.3%, precision of 36.7%, recall of 91.7%, and F1-score of 52.4%. The features used in this study include semester level, gender, nationality, birthday place, resources visit, announcement response, and extra discussion.

The performance evaluation metrics include accuracy, precision, recall, and F1-score. Accuracy provides an overview of how well a model performs in classifying data. Percentage of correct predictions out of total predictions. Here is the formula for calculating the accuracy value:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{1}$$

Precision is the ratio between the data correctly classified as true positives divided by the total number of classified data. It is the Percentage of correct optimistic predictions out of the total optimistic predictions made by the model. Precision provides information about how reliable a model is in classifying data as positive. The higher the precision value, the fewer negative cases are incorrectly identified as positive, indicating that the model is more likely to produce accurate positive results. The formula used is:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

Recall, also known as sensitivity, is a measure that indicates how well a classification model can identify all actual positive cases in a dataset. In classification, recall is calculated

as the ratio of correctly predicted positives (true positives) divided by the total number of actual positive cases (true positives + false negatives). Percentage of correct optimistic predictions out of total actual positive cases in the data. In other words, recall measures the ability of a model to "remember" or "detect" all existing positive cases without missing any. Here is the recall formula:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

F1-score is an evaluation metric that measures the balance between precision and recall. Here is the formula of F1-score:

$$\text{Recall} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

3 Methodology

The identification process begins with a literature study. A literature study is conducted to identify attributes that have the potential to become criteria. At this stage, it is a systematic process of reviewing and analyzing literature, publications, scientific articles, books, and other relevant information sources. Criteria identification is based on the results of literature studies and expert confirmation. In this process, semi-structured interviews are conducted with experts. Interviews are based on questions that refer to the literature review results. The stages in this Study are data collection, data pre-processing, data split, data processing and evaluation model [15]. These stages are shown in Figure 1.

3.1 Data Collection

This research uses data sourced from Kaggle. The link address is https://www.kaggle.com/datasets/ mexwell/student-scores.

3.2 Data Pre-processing

In data mining, the preprocessing stage plays a very important role [24, 25]. The main stage is the cleaning stage [26]. The steps for cleaning data are deleting irrelevant columns, checking for missing data, checking for duplicate data, checking for outlier values , and checking for categorical data consistency. Data Preprocessing uses Google Collaboration.

3.3 Evaluation Model

In this study, the data processing techniques used are the Decision Tree and Naïve Bayes algorithms. The ensemble techniques used are Bagging and AdaBoost. This process involves dividing the classification into six different scenarios. The scenarios carried out are the implementation of Naive Bayes (NB) and Decision Tree (DT) algorithms, the application of the Bagging Technique on Naive Bayes (NB+BG) and Decision Tree (DT+ BG), the implementation of AdaBoost on Naive Bayes (NB+ADB) and Decision Tree (DT+ADB). Data processing using the Rapid Miner application.

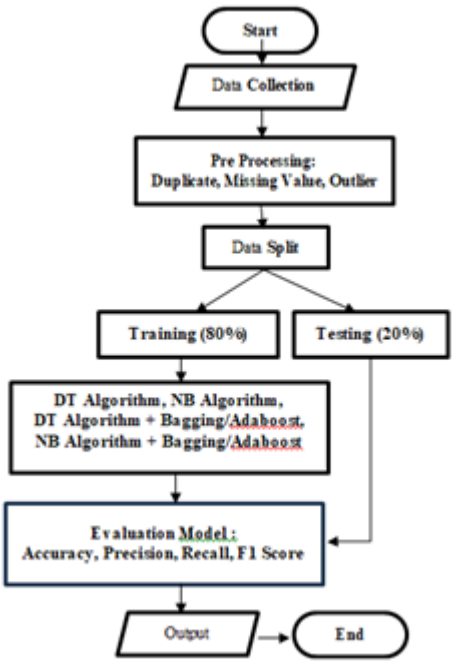

Figure 1: The stages of research.

The result analysis stage focuses on comparing results. This stage is carried out based on the test results in six scenarios. These six scenarios include testing the decision tree algorithm, Naïve Bayes algorithm, Decision Tree + Bagging algorithm, Decision Tree + AdaBoost algorithm, Naïve Bayes + Bagging algorithm, and Naïve Bayes + AdaBoost algorithm. By comparing these values, we can determine which method significantly improves classification performance.

4 Results

This section discusses data collection, data pre-processing, confusion matrix, evaluation model, and feature importance.

4.1 Data Collection

The study's results show that several factors are related to student success. These factors are academic performance, demographic factors, school and environmental context, family and social environment, student engagement, and institutional policies. [27–29] The results are the basis for confirmation to the expert to validate the suitability of the attributes to be used. The experts involved in this Study were senior teachers and principals. The experts have master's qualifications in education and are experienced in managing senior high schools. Some of the points produced are academic and non-academic factors. Academic

factors are related to a subject score, and non-academic factors include extracurricular activities.

In this Study, the dataset used is the student scores dataset downloaded from Kaggle.com. The amount of initial data before the preprocessing process is 2000. This dataset consists of 19 attributes, 18 of which function as predictor variables, including ID first name, last name, gender, absence day, extracurricular activities, weekly self-study hours, career aspiration, math score, history score, physics score, chemistry score, biology score, and geography score. Meanwhile, academic success is another attribute that acts as a target variable or label. This target variable has two output values, namely, pass and fail.

4.2 Data Pre-processing

4.2.1 Remove irrelevant columns

This process aims to produce a cleaner, more consistent, and more appropriate dataset. It does this by deleting attribute columns that do not affect the modeling process. Some deleted columns or attributes are ID, first name, last name, and email. The remaining features are all numeric data types, except part-time jobs, extracurricular activities and career aspirations. The features involved are divided into independent variables and dependent variables (labels).

4.2.2 Changing category columns to numeric columns

At this stage, the academic success column is changed to a numeric column to see the relationship between the category and target variables. The numeric values are 1 and 0. Value 1 = 'pass' and value 0 = 'fail'.

4.2.3 Checking for missing values and duplicate data

The data-cleaning process is carried out on the student score dataset with inconsistent N/A and missing values. Duplicate data checking reduces redundancy, improves model accuracy, and improves data processing efficiency. The results of the missing value process show that no data was found missing, and no duplicate data was found.

4.2.4 Checking for outlier values

This stage is checking for outliers. Outlier checking is performed on the attributes involved. Outliers exist in several variables, including math scores, biology scores, and average values. In this study, outliers are handled in truncation. Values below the lower bound are set to become the lower limit values themselves, and values above the upper bound are changed to upper limit values. The instruction for handling outlier values:

4.3 Confusion Matrix

Confusion matrix testing was conducted on six research scenarios. The Confusion matrix results are shown in Table 1.

Table 1: Confusion matrix

	DT	NB	ADB + DT	ADB + NB	BG + DT	BG + NB
TP	1524	1494	1716	1718	1719	1685
FP	4	5	1	7	3	6
FN	6	36	6	4	3	37
TN	66	65	77	71	75	72

Table 2: Evaluation of Decision Tree and Naïve Bayes algorithm

Validation	Decision Tree Algorithm	Naïve Bayes Algorithm
Accuracy	99,38%	97,44%
Precision	92,50%	66,16%
Recall	94,29%	92,86%
F1-score	93,36%	77,27%

The confusion matrix results show that Bagging + Decision Tree produces the highest number of correct predictions (1719), while Naïve Bayes has the lowest results (1524). The highest false positives are in AdaBoost + Naïve Bayes (7), which indicates that this model is more accurate in predicting students who graduate. The highest false negatives are in Bagging + Naïve Bayes $(FN = 37)$, which shows that this model is often wrong in classifying students who graduate. The highest true negatives are in AdaBoost + Decision Tree.

4.4 Evaluation Model

Validation and Testing of the model on student graduation data is carried out to evaluate the model's performance in predicting the possibility of students graduating and failing. Testing is carried out based on attributes and preprocessing data. Testing is carried out on six scenarios: Decision Tree algorithm, Naive Bayes algorithm, combining Bagging Techniques on Naive Bayes and Decision Tree algorithms, and combining AdaBoost Techniques on Naive Bayes and Decision Tree algorithms.

K-fold Cross-Validation was performed using 10 folds. Table 2 shows the Decision Tree algorithm and Naïve Bayes Algorithm test results.

The next scenario is the implementation of the Bagging Technique on Decision Tree (DT+BG) and AdaBoost on Decision Tree (DT+ADB). Table 3 shows the test results.

Table 3: Evaluation of Decision Tree + BG and Decision Tree + ADB algorithm

Validation	Decision Tree + BG Algorithm	Decision Tree + ADB Algorithm
Accuracy	99,67%	99,61%
Precision	96,39%	93,28%
Recall	96,07%	98,75%
F1-score	96,23%	97,55%

The next scenario are implementation of the Bagging Technique on Naive Bayes (NB + BG) and AdaBoost on Naive Bayes (NB + ADB). Table 4 shows the test results.

Table 4: Evaluation of Naïve Bayes + BG and Naïve Bayes + ADB Algorithm

Validation	Naïve Bayes + BG Algorithm	Naïve Bayes + ADB Algorithm
Accuracy	97,61%	99,39%
Precision	68,17%	95,14%
Recall	92,32%	91,07%
F1-score	78,35%	93,06%

4.5 Feature Importance

Feature importance testing is conducted to see the influence between attributes. Testing is conducted on six research scenarios. The test results on each decision tree algorithm, Naïve Bayes and decision tree, Naïve Bayes with the addition of AdaBoost and bagging are shown in Figure 2.

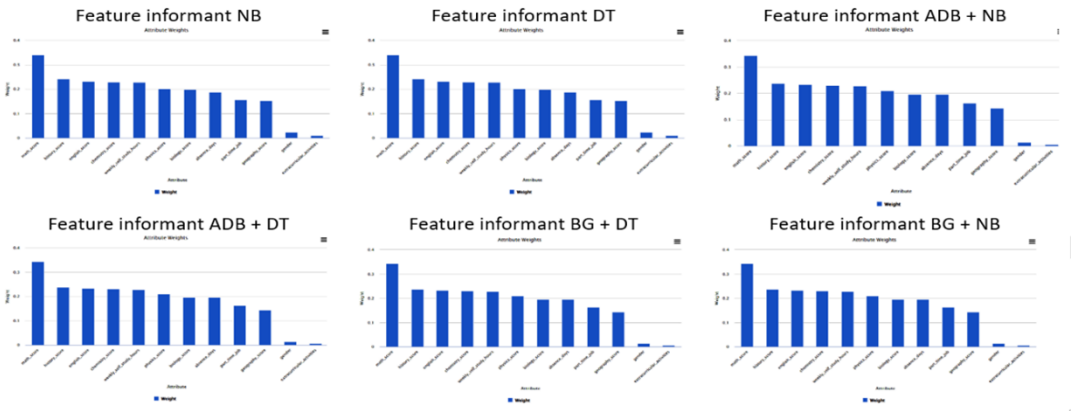

Figure 2: Feature importance.

The test results show that the most influential features in sequence are the three highest math_score, history_score, English_score. While the feature with the least influence is extracurricular_activities. In all six scenarios, the level of importance in model testing is the same.

5 Discussions

Based on the table above, the results of the comparative Testing on the student graduation grade dataset using the decision tree algorithm get accuracy value, precision, recall and F1-scores of 99.38%, 92.50%, 94.29% and 93.36%. The Naïve Bayes algorithm produces a

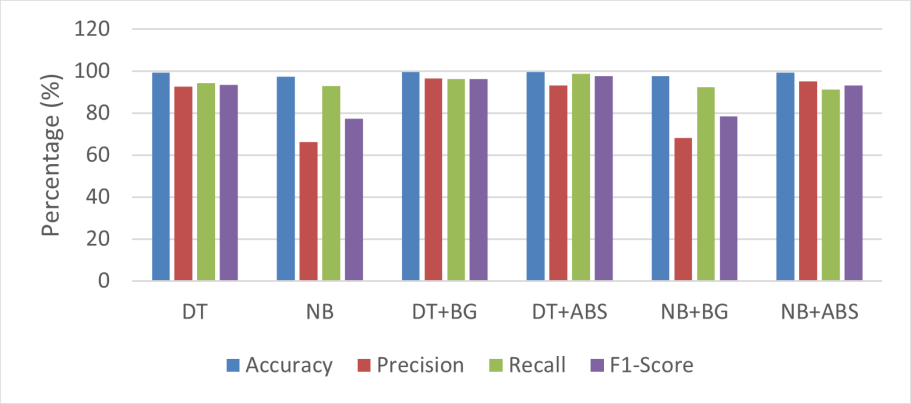

Figure 3: The performance for each evaluation.

performance of 97.44%, 66.16%, 92.86% and 77.27%. In the third scenario, the merger decision tree algorithm and bagging obtained 99.67%, 96.39%, 96.07% and 96.23%. It shows an increase in inaccuracy, precision, recall and F1-score, namely 0.29%, 3.89%, 1.78% and 2.87%. In the fourth scenario, the combination of decision and AdaBoost produces an accuracy value, precision, recall and F1-score of 99.61%, 93.28%, 98.75%, and 97.55%. It shows an increase in accuracy, recall, precision and F1-score: 0.23%, 0.78%, 4.46% and 4.19%. In the fifth scenario, between the Naïve Bayes algorithm and the bagging technique, the accuracy value precision, recall, and F1-score are 97.61%, 68.17%, 92.32%, and 78.35%. It shows that the amount is 0.17%, 2.01%, and 1.08%. While recall, there was a decrease of 0.54%. In the sixth scenario, namely the combination of the Naïve Bayes algorithm with AdaBoost, the accuracy value is precision, recall and F1-score of 99.39%, 95.14%, 91.07% and 93.06%. It shows an increase of 1.95%, 28.98%, and 15.79%. Meanwhile, recall decreased by 1.79%. The performance for each evaluation is shown in graphical form in the Figure 3.

Figure 3 shows that, for the first: the Decision Tree algorithm achieved an impressive accuracy rate of 99.38%, indicating exceptionally high performance. The Naive Bayes algorithm demonstrated a lower accuracy of 97.44%. The application of Bagging enhanced the accuracy of the Decision Tree to 99.67%, while the use of AdaBoost increased it to 99.61%. For Naive Bayes, the Bagging technique marginally improved accuracy to 97.61%, and with the addition of AdaBoost, it reached 99.39%. Overall, both Bagging and AdaBoost significantly improved the accuracy of the Decision Tree and Naive Bayes algorithms, with the Decision Tree consistently maintaining superior performance.

The second, the Decision Tree algorithm, without ensemble techniques, exhibited a precision rate of 92.50%. In contrast, Naive Bayes displayed a considerably lower precision of 66.16%. However, Naive Bayes showed significant improvement when combined with Bagging and AdaBoost, achieving precision rates of 68.17% and 95.14%, respectively. When combined with Bagging, the Decision Tree attained the highest precision rate of 96.39%, while the combination with AdaBoost yielded a slightly lower precision of 93.28%.

The third, the Decision Tree algorithm, demonstrated a recall of 94.29%, reflecting good sensitivity, whereas Naive Bayes had a slightly lower recall rate of 92.86%. The application of AdaBoost to the Decision Tree significantly increased the recall to 98.75%. In con-

trast, combining Bagging and AdaBoost with Naive Bayes resulted in a recall that remained lower than the Decision Tree.

The last, F1-score for the Decision Tree was 93.36%, while Naive Bayes lagged with a much lower score of 77.27%. The application of Bagging improved the Decision Tree's F1-score to 96.23%, and the implementation of AdaBoost further enhanced it to 97.55%. For Naive Bayes, the Bagging method yielded an F1-score of 78.35%, and AdaBoost succeeded in elevating it to 93.06%.

Some previous studies that have been conducted in predicting learning success include using the Stacking Ensemble, Random Forest (RF) + XGBoost and Decision Tree (DT) and Artificial Neural Network (ANN). The performance produced from previous studies and the performance of this study are shown in Table 5.

Table 5: Previous research performance and research results

Previous Studies	Accuracy	Precision	Recall	F1-score
Naïve Bayes [18]	87%	91%	78%	81%
Stacking Ensemble [15]	95%	91%-93%	91%-93%	91%-93%
Random Forest (RF) + XGBoost [3]	90%	90%	90%	90%
Decision Tree (DT) and Artificial Neural Network (ANN) [23]	62,3%	36,7%	91,7%	52,4%
Decision Tree (DT)	99,38%	92,50%	94,29%	93,36%
Naïve Bayes (NB)	97,44%	66,16%	92,86%	77,27%
Decision Tree (DT) + BG	99,67%	**96,39%**	96,07%	96,23%
Naïve Bayes (NB) + BG	97,61%	68,17%	92,32%	78,35%
Decision Tree (DT) + ADB	99,61%	93,28%	**98,75%**	**97,55%**
Naïve Bayes (NB) + ADB	**99,39%**	95,14%	91,07%	93,06%

Performance comparison between previous research and the results of this study shows a significant increase. The highest result for accuracy is AdaBoost + Naïve Bayes, the highest precision is Bagging + Decision Tree and Recall and F1-score on Decision Tree (DT) + ADB. Likewise, the results of the confusion matrix test conducted showed that the selection of algorithms and ensemble techniques had a significant impact on the balance of precision and recall in predicting student study success. Bagging + Decision Tree is better at reducing false negatives, so it can help institutions identify students at risk of failing more accurately. In contrast, AdaBoost + Naïve Bayes is useful for ensuring that students who are predicted to graduate actually graduate. Institutions can utilize these findings in academic monitoring systems, for example, by adjusting tutoring interventions or academic support strategies based on the prediction error patterns of the models used.

These findings underscore that ensemble techniques, particularly Bagging in Naive Bayes, can effectively detect students at risk of academic difficulties, which is crucial for educational settings. The high sensitivity of these models allows early identification of students who may be struggling, enabling timely interventions. Educational institutions can leverage these optimized predictive models to design more precise intervention programs, such as targeted tutoring, counseling, or career guidance, based on identified risk factors like high or low test scores. Moreover, such models facilitate more efficient and cost-effective intervention planning by focusing on students who need the most support rather than broad and untargeted approaches.

However, this research still has limitations and needs to be developed by combining features more completely and using a larger data set or using appropriate primary data.

6 Conclusion

Decision Tree algorithm, the Bagging and AdaBoost techniques have demonstrated their effectiveness in enhancing machine learning models' performance, accuracy, precision, recall, and F1-score. The Decision Tree consistently outperformed the Naive Bayes algorithm, especially when combined with Bagging and AdaBoost, leading to superior results in both accuracy and precision. Although the Naive Bayes algorithm showed notable improvement when integrated with AdaBoost, its performance remained lower than that of the Decision Tree. Among the combinations tested, the Decision Tree coupled with AdaBoost emerged as the most optimal in achieving high accuracy, precision, recall, and F1-score. This finding suggests that even relatively simple algorithms can be highly effective in predicting students' academic success through ensemble methods.

Based on the results of the analysis and conclusions that have been described, the following are suggestions for further research:

1. Further research is recommended to explore other algorithms, such as Random Forest, Gradient Boosting, or SVM, that may provide better results in predicting student study success.
2. The use of more extensive and more varied datasets is also essential to increase the validity of the results.
3. In addition to using Bagging and AdaBoost, other ensemble techniques, such as Stacking or XGBoost, can be considered to improve model performance.
4. Implementing deeper cross-validation and hyperparameter tuning can ensure an optimal model.

Further research must also analyze the factors that influence prediction using feature importance or SHAP values.

Acknowledgments

We express our highest appreciation to KEMENRISTEK DIKTI, which has provided financial support for this research through the Master's Thesis Research Program. We also express our gratitude to the Darmajaya Information and Business Institute for their support so that this research could be completed.

References

[1] W. Wisroni and M. F. Rozi, "Educational alternatives to the empowerment process village community," *SPEKTRUM: Jurnal Pendidikan Luar Sekolah (PLS)*, vol. 10, no. 4, pp. 689–696, 2022.

[2] N. A. Butt, Z. Mahmood, K. Shakeel, S. Alfarhood, M. Safran, and I. Ashraf, "Performance prediction of students in higher education using multi-model ensemble approach," *IEEE Access*, vol. 11, pp. 136091–136108, 2023.

[3] H. Karalar, C. Kapucu, and H. Gürüler, "Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system," *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, p. 63, 2021.

[4] D. Pradana and E. Sugiharti, "Implementation data mining with naive bayes classifier method and laplace smoothing to predict students learning results," *Recursive Journal of Informatics*, vol. 1, no. 1, pp. 1–8, 2023.

[5] Y. A. Alsariera, Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, and N. Ali, "Assessment and evaluation of different machine learning algorithms for predicting student performance," *Computational intelligence and neuroscience*, vol. 2022, no. 1, p. 4151487, 2022.

[6] G. Nanfack, P. Temple, and B. Frénay, "Constraint enforcement on decision trees: A survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 10s, pp. 1–36, 2022.

[7] M. Ragab, A. M. Abdel Aal, A. O. Jifri, and N. F. Omran, "[retracted] enhancement of predicting students performance model using ensemble approaches and educational data mining techniques," *Wireless Communications and Mobile Computing*, vol. 2021, no. 1, p. 6241676, 2021.

[8] S. Chen, G. I. Webb, L. Liu, and X. Ma, "A novel selective naïve bayes algorithm," *Knowledge-Based Systems*, vol. 192, p. 105361, 2020.

[9] S. González, S. García, J. Del Ser, L. Rokach, and F. Herrera, "A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities," *Information Fusion*, vol. 64, pp. 205–237, 2020.

[10] I. D. Mienye and Y. Sun, "A survey of ensemble learning: Concepts, algorithms, applications, and prospects," *Ieee Access*, vol. 10, pp. 99129–99149, 2022.

[11] A. Plaia, S. Buscemi, J. Fürnkranz, and E. L. Mencía, "Comparing boosting and bagging for decision trees of rankings," *Journal of Classification*, vol. 39, no. 1, pp. 78–99, 2022.

[12] O. W. Adejo and T. Connolly, "Predicting student academic performance using multi-model heterogeneous ensemble approach," *Journal of Applied Research in Higher Education*, vol. 10, no. 1, pp. 61–75, 2018.

[13] A. Z. Zakaria, A. Selamat, H. Fujita, and O. Krejcar, "The best ensemble learner of bagged tree algorithm for student performance prediction," in *Knowledge Innovation Through Intelligent Software Methodologies, Tools and Techniques*, pp. 55–64, IOS Press, 2020.

[14] C. Jalota, "An effectual model for early prediction of academic performance using ensemble classification," *J. Lang. Ling. Soc.(JLLS)*, vol. 3, no. 02, pp. 19–33, 2023.

[15] N. A. Butt, Z. Mahmood, K. Shakeel, S. Alfarhood, M. Safran, and I. Ashraf, "Performance prediction of students in higher education using multi-model ensemble approach," *IEEE Access*, vol. 11, pp. 136091–136108, 2023.

[16] W. Punlumjeak, S. Rugtanom, S. Jantarat, and N. Rachburee, "Improving classification of imbalanced student dataset using ensemble method of voting, bagging, and adaboost with under-sampling technique," in *IT Convergence and Security 2017: Volume 1*, pp. 27–34, Springer, 2017.

[17] L. Lisnawita, G. Guntoro, and M. Musfawati, "Implementation of naïve bayes for classification of learning types," *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, no. 1, pp. 44–54, 2022.

[18] O. B. Akanbi, "Application of naive bayes to students' performance classification," *Asian Journal of Probability and Statistics*, vol. 25, no. 1, pp. 35–47, 2023.

[19] T. Thomas, A. P. Vijayaraghavan, S. Emmanuel, T. Thomas, A. P. Vijayaraghavan, and S. Emmanuel, "Applications of decision trees," *Machine learning approaches in cyber security analytics*, pp. 157–184, 2020.

[20] N. E. I. Karabadji, I. Khelf, H. Seridi, S. Aridhi, D. Remond, and W. Dhifli, "A data sampling and attribute selection strategy for improving decision tree construction," *Expert Systems with Applications*, vol. 129, pp. 84–96, 2019.

[21] N. Normah, "Naïve bayes algorithm for sentiment analysis windows phone store application reviews," *Sinkron: jurnal dan penelitian teknik informatika*, vol. 3, no. 2, pp. 13–19, 2019.

[22] H. Karalar, C. Kapucu, and H. Gürüler, "Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system," *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, p. 63, 2021.

[23] S. Hussain, S. Rehman, S. Raza, A. Mahmood, and S. Kundi, "Significance of education data mining in student's academic performance prediction and analysis," *International Journal of Innovations in Science & Technology*, vol. 5, no. 3, pp. 215–231, 2023.

[24] H. Jamshed, S. A. Khan, M. Khurrum, S. Inayatullah, and S. Athar, "Data preprocessing: A preliminary step for web data mining," *3c Tecnología: glosas de innovación aplicadas a la pyme*, vol. 8, no. 1, pp. 206–221, 2019.

[25] Chithra, P. Kiran, and Manoj, "The novel method for data preprocessing CLI," *Advances in Intelligent Systems and Technologies*, pp. 117–120, Dec. 2022.

[26] E. J. Elvin Jafarov, "Data cleaning before uploading to storage," *ETM - Equipment, Technologies, Materials*, vol. 13, pp. 117–127, Feb. 2023.

[27] C. McKinley Yoder, M. A. Cantrell, and J. L. Hinkle, "Disparities in high school graduation by identity and disability using intermediate and long-term educational outcomes," *J. Sch. Nurs.*, vol. 40, pp. 266–274, June 2024.

[28] G. Zhang, T. J. Anderson, M. W. Ohland, and B. R. Thorndyke, "Identifying factors influencing engineering student graduation: A longitudinal and cross-institutional study," *Journal of Engineering education*, vol. 93, no. 4, pp. 313–320, 2004.

[29] P. K. L. M. Aubourg, "The non-financial factors that can explain the low graduation rate of black students in higher education," *International Journal of Community Development and Management Studies*, vol. 7, pp. 39–54, 2023.