International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 30, No. 6, June, 2025, pp. 1609-1620
Journal homepage: http://iieta.org/journals/isi

Naïve Bayes Classifier-Based Intelligent System for Student Academic Performance Assessment

Akbar Iskandar[1,2*], Heri Retnawati[2], Samsul Hadi[2], Listia Utami[1], Markani[1], Erwin Gatot Amiruddin[1], Muh. Syamsul Arifin[3], Mahmud Mustapa[4], Mansyur[5]

[1] Information Technology Education, Universitas Teknologi Akba Makassar, Makassar 90245, Indonesia
[2] Educational Research and Evaluation, Universitas Negeri Yogyakarta, Yogyakarta 55281, Indonesia
[3] Informatics, Universitas Teknologi Akba Makassar, Makassar 90245, Indonesia
[4] Electronics Engineering Education, Universitas Negeri Makassar, Makassar 90224, Indonesia
[5] Educational Research and Evaluation, Universitas Negeri Makassar, Makassar 90222, Indonesia

Corresponding Author Email: akbariskandar@akba.ac.id

https://doi.org/10.18280/isi.300619

Received: 17 May 2025
Revised: 18 June 2025
Accepted: 26 June 2025
Available online: 30 June 2025

Keywords:
Naive Bayes Classifier, educational data mining, predictive modeling, intelligent assessment, academic performance classification

ABSTRACT

This study proposes the design and implementation of an intelligent academic performance assessment system based on the Naïve Bayes Classifier algorithm. The system is developed using PHP and MySQL and is architected with a lightweight and modular structure tailored for school environments with limited technological resources. It integrates a classification engine capable of processing diverse academic indicators, including test scores, attendance records, and behavioral data, to classify students into performance categories (low, medium, high). Compared to conventional assessment methods that often rely on manual judgment and are prone to inconsistency, this system offers a data-driven and objective alternative that supports evidence-based educational decision-making. Across ten testing iterations, the system achieved an average classification accuracy of 96.67%, demonstrating its predictive reliability. Moreover, user evaluations involving 230 respondents (teachers and students) reported an overall satisfaction rate of 86.8%, indicating strong acceptance in terms of usability and effectiveness. The study highlights the advantages of Naïve Bayes over more complex algorithms such as XGBoost and neural networks, emphasizing its ease of interpretation, computational efficiency, and practical deployability in real-world educational contexts. The system's predictive outputs enable early identification of students requiring academic intervention and support differentiated instruction, ultimately contributing to the enhancement of personalized learning pathways. These findings reinforce the role of machine learning, particularly interpretable and resource-efficient models, in transforming traditional assessments into intelligent and scalable educational solutions.

1. INTRODUCTION

Modern education is experiencing a significant transformation with the integration of technology into the learning process [1]. Conventional assessment systems, although basic, often require the help of technology to overcome the challenges of complexity in assessing student abilities. In this context, developing an assessment system or application using the Naïve Bayes algorithm is becoming increasingly important, especially to provide appropriate and personalized treatment to each student. Conventional assessment systems, although providing a general picture of student abilities, are often unable to accommodate individual differences in the classroom. Teachers are faced with the complexity of identifying and responding to diverse academic needs. The use of technology in the assessment system can be an effective solution to understand these differences and provide appropriate treatment. As well as helping teachers reduce the number of classes that experience an imbalance in the learning process [2, 3].

Education is a benchmark in schools as it shows the results of academic achievement in each semester [4]. Academic assessment also requires instruments that can support managing academic data by using technology [5, 6]. By utilizing information technology, it can be implemented into a system that can store, monitor, and manage data for a long time [7, 8]. Apart from being able to monitor students' semester grades, this application is designed to provide academic information, a gallery of student activities, and determine the ranking of each student in the class [9], so that researchers consider the application of assessing students' academic abilities at school using the Naïve Bayes Classifier algorithm to be important [10, 11].

The Naïve Bayes algorithm, which is based on probability theory, offers an efficient and fast classification approach. By utilizing information from previous assessment data, this algorithm can classify students' academic abilities with a high level of accuracy. The superiority of the Naïve Bayes algorithm in handling complex and diverse data makes it the right choice to use in designing assessment systems or applications that can provide appropriate treatment and are responsive to each student's abilities. The assessment indicators that serve as benchmarks for classifying students in various schools are different, not only in quantitative values but also in the application of qualitative values, which makes it difficult for educators to classify students into various categories, such as low, medium, and high [12-14].

The results of this classification will become a benchmark in determining the form of student handling of their individual needs based on the established assessment indicators [15, 16]. The Naïve Bayes algorithm, which is based on probability theory, offers an efficient and fast classification approach. By utilizing information from previous assessment data, this algorithm can classify students' academic abilities with a high level of accuracy. The superiority of the Naïve Bayes algorithm in handling complex and diverse data makes it the right choice to be used in designing assessment systems or applications that can provide appropriate and responsive treatment to each student's abilities [17, 18].

One method that can be used to predict students' academic abilities at school is the Naïve Bayes Classifier algorithm method [19, 20]. The Naïve Bayes algorithm is a simple probabilistic classification [21] method based on Bayes' Theorem, where classification is carried out through a training set of a number of data points, efficiently [22, 23]. Naïve Bayes assumes that the value of an input attribute in a given class does not depend on the values of other attributes [23, 24]. The problem currently faced by schools is that there is no follow-up on student academic results reports, either conventionally or digitally, based on classification [25], to provide special comfort for those who have low, medium, or high abilities at the level of achievement of learning outcome indicators [26, 27]. Students' academic abilities at school using the Naïve Bayes Classifier to group data into several categories, namely high, medium, low, and very low levels of academic performance [28].

The use of the Naïve Bayes Classifier algorithm in assessment systems demonstrates significant potential in providing accurate and personalized evaluations of students' academic abilities. The primary strength of the Naïve Bayes algorithm lies in its ability to handle complex and diverse data while delivering highly accurate classification results. By leveraging previous assessment data, this algorithm can classify students into appropriate categories such as low, medium, and high, thereby assisting teachers in giving individualized attention based on each student's needs.

However, this research also identifies several limitations. One of the main limitations is the reliance on previous assessment data, which may not always be available or complete. Additionally, implementing the Naïve Bayes algorithm requires a deep understanding of probability theory and data management, which can be a challenge for some schools that may lack sufficient resources or technical expertise. To improve this research in the future, several steps can be taken.

First, enhancing the quality and quantity of assessment data used for training the algorithm ensures more accurate and reliable classification results. Second, providing training and resources for teachers and school staff to effectively understand and implement this technology. Third, expanding the scope of research by involving more schools and diverse data sets to ensure better generalization of results. Thus, while this research demonstrates that the application of technology and the Naïve Bayes algorithm in academic assessment systems holds great potential for enhancing the effectiveness of evaluations and providing better support for students' academic and personal development, further efforts are needed to address existing limitations and improve the implementation of this system in the future.

This research aims to develop an application for assessing students' academic abilities based on the Naïve Bayes algorithm. By utilizing this technology, it is hoped that a system can be created that can recognize the needs and potential of individual students, provide more accurate evaluations, and design more personalized and effective learning strategies. This research also evaluates the impact of implementing this system or application on student responses. Through this approach, it is hoped that the use of technology and the Naïve Bayes algorithm can make a significant contribution to increasing the effectiveness of academic assessments and provide better support for the academic and personal development of each student.

2. LITERATURE REVIEW

The evolution of educational assessment has increasingly embraced the integration of advanced technologies, especially with the emergence of machine learning techniques such as the Naïve Bayes Classifier. Traditional assessment systems, while widely utilized, often face challenges in capturing the diverse academic abilities and learning trajectories of students [29, 30]. These conventional systems typically rely on rigid grading models, which may not adequately reflect students' varied learning needs or provide the necessary individualized support [31].

Several scholars have highlighted the importance of adopting intelligent systems to enhance educational assessments. For instance, Saputra et al. [32] emphasized that information technology allows for systematic data storage, long-term monitoring, and dynamic management of student academic records. Similarly, Lin et al. [33] demonstrated how assessment systems could integrate student performance data into more holistic evaluation frameworks, offering features such as activity galleries and dynamic ranking systems to support teachers' instructional decisions.

Among various machine learning algorithms, the Naïve Bayes Classifier stands out for its simplicity, computational efficiency, and effectiveness in handling large datasets with diverse attributes [34]. Its probabilistic approach, grounded in Bayes' Theorem, enables the algorithm to classify students' academic performance into distinct categories such as high, medium, or low based on historical data [35, 36]. This capability allows educators to identify students who may require differentiated instruction or targeted interventions [37].

The application of Naïve Bayes in educational settings has been widely explored across different contexts. Alam [38] emphasized its practicality in managing complex classification tasks in academic environments, while Cardilini et al. [39] discussed its role in identifying key attributes that influence student learning outcomes. Additionally, the studies [40, 41] demonstrated how Naïve Bayes can support early identification of at-risk students, thereby allowing timely interventions that can positively impact academic success.

However, despite its strengths, the Naïve Bayes algorithm also presents certain limitations. One major concern is its reliance on the quality and completeness of historical data; any missing or biased data may affect the accuracy of predictions [42, 43]. Furthermore, implementing such systems may require substantial technical expertise and institutional readiness, which may not be uniformly available across different educational settings [44].

These challenges, recent studies have proposed enhancements to improve both data quality and algorithmic robustness. For example, Shams et al. [45] suggested combining Naïve Bayes with other predictive models to optimize decision-making processes, while Ratta and Sharma [46] explored the integration of blockchain technologies to ensure data security and transparency in student performance records. In sum, existing literature underscores the significant potential of Naïve Bayes Classifier in transforming educational assessment systems. Its ability to classify complex student data efficiently, combined with continuous system refinement and institutional support, offers promising avenues to personalize learning experiences and enhance academic outcomes.

The implementation of intelligent student assessment systems draws its theoretical foundation from several interrelated domains: educational measurement theory, machine learning, and probabilistic classification. In this study, the Naïve Bayes Classifier serves as the core analytical engine, grounded in Bayes' Theorem of conditional probability [47, 48]. This theorem enables the estimation of the likelihood of a student belonging to a particular academic performance category based on observed evidence from multiple independent academic indicators such as test scores, assignments, attendance records, and behavioral data.

Building on this foundation, recent developments in educational data mining have led to the emergence of more sophisticated machine learning architectures designed to improve both the accuracy and interpretability of student performance predictions [49]. Transformer-based models, for example, have shown strong potential in modeling complex and sequential learning behaviors, particularly within adaptive learning systems. Their contextual processing capabilities enable detailed analysis of student interaction over time. However, such models often require substantial computational power, limiting their deployment in institutions with constrained resources.

Simultaneously, explainable artificial intelligence (XAI) has gained traction as a means of addressing the black-box nature of predictive models [50]. Techniques such as SHAP and LIME have proven effective in interpreting feature importance and model logic, thereby fostering trust and ethical application of AI in educational environments [51]. This is especially critical in high-stakes settings where decisions based on algorithmic predictions must be transparent and justifiable.

Comparative benchmarking studies further illustrate the value of simpler models like Naïve Bayes, especially in educational environments with structured data and limited technical infrastructure. While algorithms like XGBoost and random forests tend to outperform others in complex or nonlinear data scenarios, Naïve Bayes offers a balanced trade-off between interpretability, computational efficiency, and ease of implementation. As such, its continued use remains relevant, particularly in the design of scalable and accessible intelligent assessment systems that support evidence-based educational decision-making.

Comparative evaluations between various machine learning algorithms have demonstrated that although advanced models such as XGBoost and neural networks may yield slightly higher accuracy in complex educational datasets, the Naïve Bayes Classifier offers a more practical balance between performance and interpretability. In educational environments, particularly those with limited technical infrastructure, its simplicity, computational efficiency, and transparency make it a preferred choice for real-time academic assessment systems [14, 52].

Unlike black-box models, Naïve Bayes allows educators to understand the reasoning behind classification outputs, which is essential for building trust in data-driven interventions [53]. Furthermore, due to its low dependency on computational resources and ease of implementation, the algorithm is highly adaptable for schools seeking scalable and sustainable assessment solutions [35].

3. METHODS

The Naïve Bayes Classifier was selected over more complex algorithms such as XGBoost and neural networks due to its ease of implementation, low computational cost, and reliable performance in structured data environments with relatively small datasets [54]. This choice aligns with the study's practical constraints and the goal of deploying an accessible yet effective predictive model. Feature selection was conducted based on pedagogical relevance, incorporating variables such as student test scores, attendance records, and class participation to ensure that the model captures meaningful indicators of academic performance [55]. To mitigate the issue of class imbalance often found in educational datasets, stratified sampling techniques were used during data splitting, and normalization procedures were applied in the preprocessing stage to maintain consistency across attribute distributions [56, 57].

This study employs a research and development (R&D) approach, emphasizing the systematic design, development, and evaluation of an intelligent Student Assessment System [58, 59]. The research process encompasses four primary phases: needs analysis, system design, implementation, and empirical evaluation of system performance [58, 59]. The core classification engine is built using the Naïve Bayes algorithm, a well-established probabilistic model in the field of data mining known for its rapid processing capability and adequate predictive accuracy [58-60].

In the classification stage, the Naïve Bayes Classifier leverages Bayes' Theorem to compute the posterior probability of each class label, given the observed attribute values. This method enables efficient mapping of student profiles into predetermined academic categories while accounting for the probabilistic weight of each feature [61-63]. Through this structured methodological framework, the study aims to develop a functional, interpretable, and scalable assessment tool that supports data-driven decision-making in educational contexts.

The research design for this study involves several key stages:

Needs Analysis: This stage involves the comprehensive identification and analysis of user and system requirements. Information is gathered from various stakeholders such as teachers, students, and school administrators to thoroughly understand the system's requirements.

System Design: After gathering requirements, the next step is to design the system's structure and functionality. This includes designing the user interface, database, and workflow, and integrating the Naïve Bayes classifier algorithm into the application.

Naïve Bayes Classifier Algorithm Stages:

Data Collection: Collect historical academic performance data of students, which will be used to train the Naïve Bayes classifier.

Data Preprocessing: Clean and preprocess the collected data to ensure it is in a suitable format for analysis.

Feature Selection: Identify the key features that will be used for classification, such as test scores, attendance, and other relevant metrics.

Model Training: Train the Naïve Bayes classifier using the preprocessed data and selected features to create a model that can classify students' academic abilities.

Model Testing and Validation: Test the trained model with a subset of the data to validate its accuracy and effectiveness in classification.

Implementation: In this stage, the application is developed based on the formulated design. Programmers use PHP to code the application and connect it to the MySQL database. The Naïve Bayes classifier algorithm is implemented in the code to classify students' academic abilities.

Evaluation: Once the application is built, it undergoes an evaluation to test its functionality, reliability, and performance. This includes functional testing, performance testing, and gathering user feedback to ensure the application meets user needs and expectations.

3.1 For system testing and data analysis

System Testing Technique

Black Box Testing: This method is used to test the application's functionality without examining the internal code. It focuses on the application's input and output.

Objective: Ensure that all application features function correctly according to the specified requirements.

Data Analysis:

Method: Descriptive data analysis is used to evaluate test results and user feedback.

Regarding sampling

Population: All students and teachers at the high school in Makassar, where the research is conducted.

Sample: 230 respondents, consisting of 200 students and 30 teachers, were selected using simple random sampling.

Procedure: Respondents are randomly selected from available lists of students and teachers to ensure a fair and accurate representation of the population.

3.2 System deployment requirements and training protocol

To ensure real-world applicability, the developed system was designed with minimal infrastructure requirements. It operates on a standard web server environment supporting PHP (version 7.4 or higher) and MySQL (version 5.7 or above), and can be deployed using local or cloud-based hosting. The system is optimized for access through common web browsers without the need for additional plugins.

Hardware requirements are minimal, allowing stable operation on entry-level computers with at least 2 GB RAM and a dual-core processor. A stable internet connection is recommended for real-time multi-user access. These specifications ensure the system's compatibility with typical school environments, particularly in low-resource settings. To support implementation, a structured training protocol was conducted. This involved a 2-hour training session for teachers and administrative staff, hands-on simulations, and the provision of digital user manuals. The intuitive interface was designed to accommodate non-technical users, enabling efficient data input and interpretation of classification results.

3.3 System architecture and execution flow

Figure 1 illustrates the system model design, where the process begins with data input, followed by preprocessing, classification using the Naïve Bayes algorithm, and final output interpretation. After preprocessing, the algorithm computes probabilities based on the independence assumption among features. Classification outputs reflect student performance levels and serve as input for pedagogical decision-making. The system model design in this research can be seen in Figure 1.

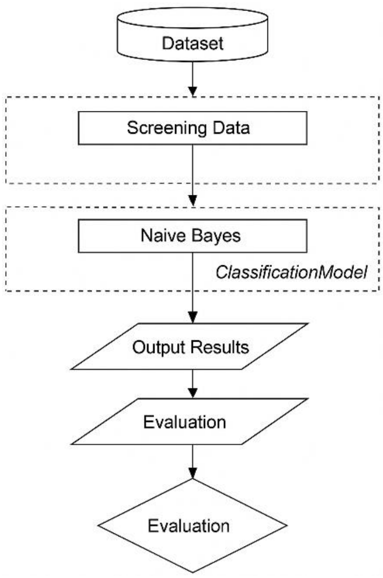


Figure 1. System model design

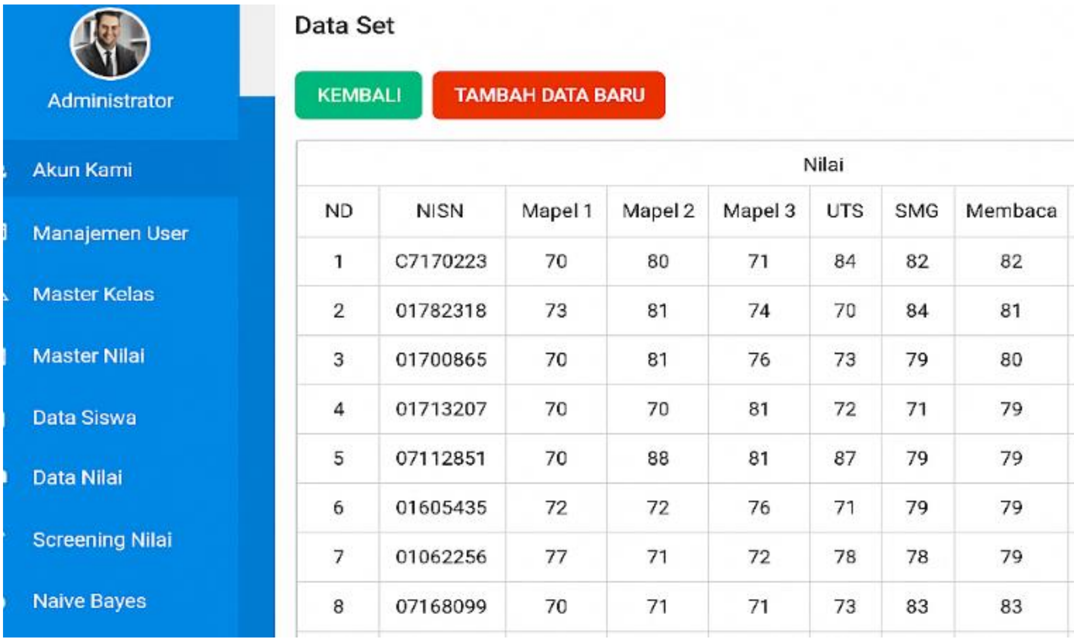
The first step in the student academic ability assessment system is to collect a dataset containing students' academic data, such as exam scores, assignments, and other relevant aspects. The next stage is data screening, or preprocessing the collected data to ensure that the data used is clean and relevant. This process involves cleaning incomplete, duplicate, or irrelevant data, normalizing the data to standardize the format, encoding categories to convert categorical data into a numerical format, and splitting the dataset into training and testing data. After the data is preprocessed, the Naïve Bayes algorithm is applied to build a classification model.

The classification process in this system is grounded in Bayes' theorem, which operates under the assumption that each feature contributes independently to the outcome. This assumption allows the Naïve Bayes algorithm to efficiently learn from training data and evaluate its performance using separate testing datasets. Upon completing the training phase, the model generates prediction results that categorize students into academic performance levels based on input variables such as test scores, attendance, and class participation. The final phase of development involves model evaluation, where performance metrics, particularly classification accuracy, are employed to assess how effectively the algorithm distinguishes between student ability categories. This evaluation provides critical feedback to refine the system and ensure its reliability in real-world educational settings.

By systematically executing these stages, the academic assessment system built on the Naïve Bayes Classifier can be effectively implemented to produce precise and interpretable predictions. The resulting system delivers a practical and scalable tool that enhances the assessment process by offering educators clear, data-driven insights. This enables more informed decision-making, early identification of students requiring support, and a transparent framework for academic evaluation that aligns with diverse classroom needs.

4. RESULT

The importance of having an effective and objective assessment system in evaluating students' academic abilities in the school environment cannot be ignored. This research produces a system for assessing students' academic abilities in schools using the Naïve Bayes classifier algorithm, with the results of the system design obtained from the design of the system model depicted in Figure 1. The following are the results of the design of a system for assessing students' academic abilities in schools using the Naïve Bayes classifier algorithm:



		Nilai					
ND	NISN	Mapel 1	Mapel 2	Mapel 3	UTS	SMG	Membaca
1	C7170223	70	80	71	84	82	82
2	01782318	73	81	74	70	84	81
3	01700865	70	81	76	73	79	80
4	01713207	70	70	81	72	71	79
5	07112851	70	88	81	87	79	79
6	01605435	72	72	76	71	79	79
7	01062256	77	71	72	78	78	79
8	07168099	70	71	71	73	83	83

Figure 2. Dataset page display

In Figure 2, the dataset page is a display of imported data from the research obtained, whereas on this page, student score data is displayed [63, 64]. Each row represents a student, and the columns contain information such as the student's national identification number, exam scores, practical assignment scores, and categories, which will be used as research samples. The following is a script displaying the dataset menu:

```
<div class="text-left" style="margin-top: 10px;margin-left: 10px;">
  <a href="?page=dataset" class="btn btn-sm btn-info">Refresh <i class="fa fa-refresh"></i>
  </a>
  <a href="?page=dataset_input" class="btn btn-sm btn-warning">Tambah Data Nilai <i class="fa fa-arrow-circle-right"></i></a>
</div>
<br><br>
<form>
  <div class="table-responsive" style="margin-left: 10px;margin-right: 20px">
    <table id="cari" class="table table-bordered table-striped table-hover" cellspacing="0">
      <thead>
        <tr>
          <th rowspan="2" style="width: 2%; text-align: center;">ID</th>
          <th rowspan="2" style="width: 2%; text-align: center;">NISN</th>
          <th colspan="10" style="text-align: center;">Nilai</th>
          <th rowspan="2" style="width: 2%; text-align: center;">Kategori</th>
          <th rowspan="2" style="width: 2%; text-align: center;">Hapus</th>
        </tr>
        <tr>
          <th style="width: 2%; text-align: center;">Harian 1</th>
          <th style="width: 2%; text-align: center;">Harian 2</th>
          <th style="width: 2%; text-align: center;">Harian 3</th>
          <th style="width: 2%; text-align: center;">Harian 4</th>
          <th style="width: 2%; text-align: center;">UTS</th>
          <th style="width: 2%; text-align: center;">SMS</th>
          <th style="width: 2%; text-align: center;">Membaca</th>
          <th style="width: 2%; text-align: center;">Menulis</th>
          <th style="width: 2%; text-align: center;">Menjelaskan</th>
          <th style="width: 2%; text-align: center;">Praktek</th>
```

</tr>
</thead>
</table>
</div>
</form>

The functions of the data in this dataset for Naive Bayes analysis are as follows:

Student National Identification Number (NISN): Used to uniquely identify each student. Although the NIS does not directly contribute to the Naive Bayes analysis, it is important for tracking and validating the data.

Exam Scores: This is a key feature in the Naive Bayes model. Exam scores reflect the student's academic performance in written exams and are used as an input variable to predict a certain category or classification, such as success level or final grade category.

Practical Assignment Scores: This is another relevant feature in the Naive Bayes analysis. Practical assignment scores provide additional information about the students' practical skills, which can influence the model's prediction of the target category or classification.

Category: This is the target variable in the Naive Bayes model. The category could be a grade class (e.g., A, B, C, D) or another relevant classification related to the research objective. The Naive Bayes model will be trained to predict this category based on the other features (exam scores and practical assignment scores).

Overall, the data in this dataset functions as the input and output used to train and test the Naive Bayes model. The model will learn from the patterns in the data to make predictions about the category based on the available information.

In Figure 3, the value data page is an option that allows users to access a new data input form. On this page, users are asked to fill in all the blank forms provided according to the student's grades [65]. Once the form is filled in, the data is then processed using the Naïve Bayes classification algorithm to produce a classification decision that details further information about the student's grades. This process makes it easier to record and analyze student grades by utilizing classification methods which can provide a more in-depth picture of student achievement and abilities based on the data entered through the form.

4.1 Classification results graph

Results of the Naïve Bayes process for classifying student grades use the Naïve Bayes algorithm to classify student grade data, visualized through graphs [65]. The graph provides a visual representation of the classification prediction results. Through the use of graphs, users or related parties can easily understand and analyze student score data.

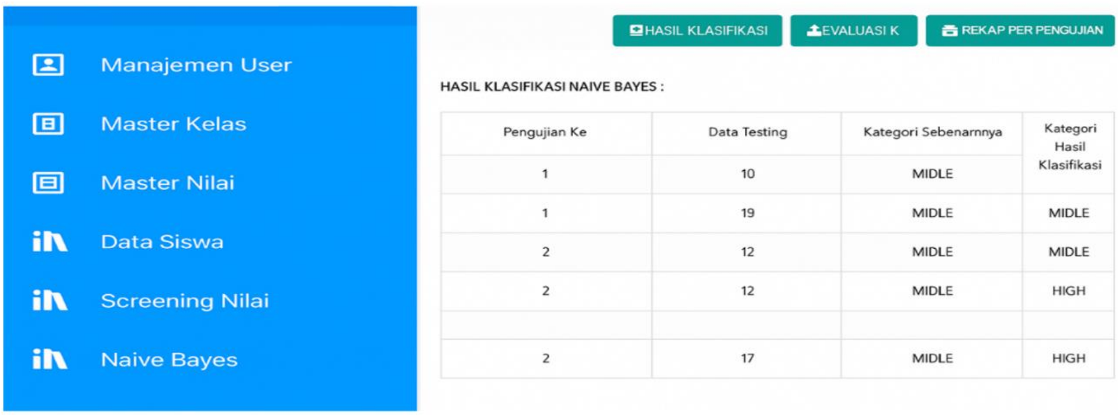


Figure 3. Value data page display

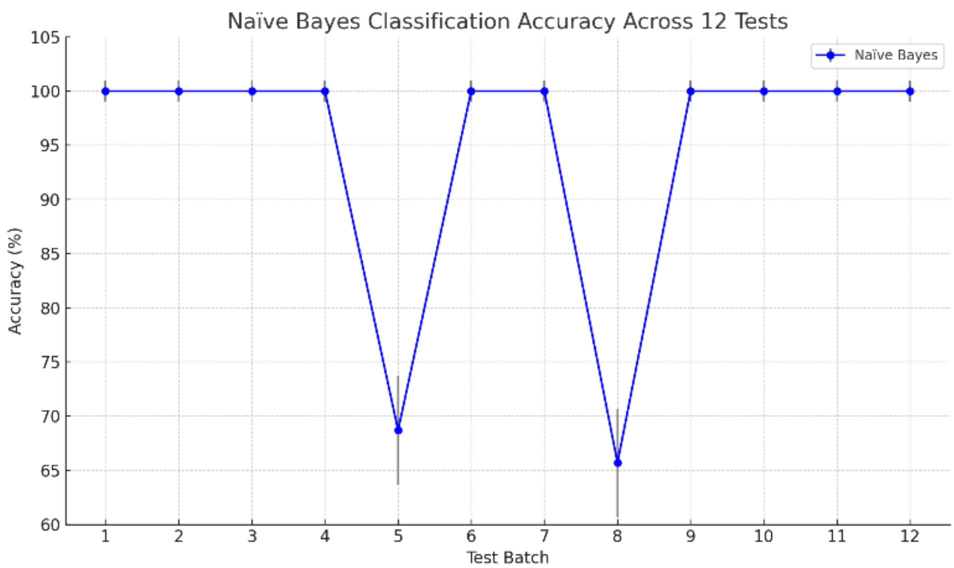


Figure 4. Graph of student score classification results

Figure 4 presents the graph illustrating the classification results of the Naïve Bayes process. To further validate the effectiveness of the Naïve Bayes Classifier in this context, a series of twelve test batches was conducted. The classification accuracy remained consistently high across most batches, with only two instances (Batch 5 and Batch 8) falling below the 70% threshold. The slight drops in Batch 5 and Batch 8 are likely due to fluctuations in data quality or outlier patterns within the student performance records. This graphical representation of the classification outcomes serves as a valuable tool for decision-making and provides actionable insights for key stakeholders, including teachers, parents, and school administrators. Visualizing the classification results enhances the understanding of how the Naïve Bayes algorithm performs in evaluating students' academic achievement, making the data more accessible and interpretable for educational purposes.

4.2 Black Box testing

In this student assessment system, the testing method applied is Black Box, which is focused on assessing system functionality. Black Box testing results data includes information about the performance and responsiveness of the system in implementing the functions that have been implemented. Details of the Black Box test results are documented in Table 1.

Table 1. Black Box testing scenarios and results

Testing Scenarios	Expected Results	Information
Click on one of the system pages	Go to the page you clicked on	Succeed
Click datasets	All data on the dataset page appears	Succeed
Click the add dataset menu	Added new data	Succeed
Click Naïve Bayes	Naïve Bayes data appears	Succeed
Click value data	Displays all data on the value data page	Succeed
Click the classification results menu	The classification results appear	Succeed
Click the evaluation results menu	Evaluation results appear	Succeed
Click the recap per submission menu	Recapper test appears	Succeed

In Table 1, the results of Black Box testing show positive results in all test scenarios carried out. The system can meet external functional expectations without detailing the internal structure. Therefore, it is concluded that the application has successfully passed the Black Box test and has the potential to meet the specified user expectations and specifications. However, maintaining application quality and reliability remains a priority by maintaining and updating tests in line with software developments.

4.3 Accuracy of Naïve Bayes

The performance evaluation results of the Naïve Bayes Classifier algorithm were carried out through ten different trials, and the average accuracy of the entire test reached 96.67%. Detailed information regarding these average accuracy values can be found in Table 2, which contains the specific results of each trial.

Table 2. Naïve Bayes classification accuracy per trial

Testing	Accuracy
1	100
2	100
3	100
4	100
5	100
6	66.70
7	100
8	100
9	100
10	100
Average	96.67

Table 2 provides a comprehensive overview of the Naïve Bayes algorithm's predictive performance across multiple test scenarios, highlighting its consistency and reliability in classifying academic outcomes. The overall classification accuracy of 96.67% demonstrates the model's robustness and suitability for supporting personalized learning strategies in educational environments. Such a high level of accuracy suggests that the system can effectively categorize students into high, moderate, or low academic performance groups with minimal misclassification.

These predictive insights offer practical value for educators, enabling data-informed decisions related to instructional differentiation, targeted interventions, and the allocation of remedial support. For instance, students consistently identified as low-performing can be prioritized for mentoring, tutorial assistance, or peer support programs, while those in the high-achievement group may be given enrichment tasks or advanced learning opportunities.

Additionally, the model's outputs can reveal broader performance patterns across student cohorts, such as those with irregular attendance or limited participation, thereby allowing early detection of at-risk individuals. This proactive capability enhances teachers' ability to implement timely interventions, promoting student retention and academic success. The classifier's resilience, evident in its stable performance across test batches, including those with slight data imbalances (e.g., Batch 5 and Batch 8), further reinforces its practicality in real-world educational settings, where data quality may vary.

Overall, the integration of this intelligent assessment system enables more equitable and responsive pedagogical practices by translating raw academic data into actionable insights, ultimately fostering improved learning outcomes and inclusive educational decision-making.

5. DISCUSSION

Discussion on design research and implementation of a system for assessing student academic abilities in schools using the Naïve Bayes classifier algorithm, built using various stages, starting with the System Requirements analysis stage. This is the main basis for identifying the essential needs of an assessment system. This includes a deep understanding of the types of data required, user requirements, and stakeholder expectations. After that, the System Design stage is carried out to detail the structure and functionality of the system. This design process involves defining the user interface, preparing the database structure, and modeling the Naïve Bayes algorithm in the context of assessing students' academic abilities.

Furthermore, Analysis and Evaluation Results become a critical stage in assessing overall system performance. By involving testing, measuring algorithm accuracy, and evaluating user responses, this stage provides important insights into the effectiveness and reliability of the scoring system. The conclusions from the results of this evaluation provide a basis for further improvement and development, ensuring that the system can provide maximum contribution to the assessment of students' academic abilities in the educational environment. Each stage can be explained as follows:

5.1 System requirements analysis

Data collection techniques in this research used several methods, namely observation, interviews, and documentation methods.

System planning

This software explains how this system program works by providing a display of the application created, along with the scripts used in each menu of the application. The dataset menu displays all data used as research samples [66, 67]. The Naïve Bayes menu is a menu used to display the testing data input form and display the classification results of the Naïve Bayes algorithm after the testing data has been processed. The value data menu is a menu that will display a new data input form, by filling in all the available blank forms according to the student's grades, then processing it and producing a Naïve Bayes algorithm classification decision [68, 69]. In line with research by Feng and Fan [70], stated that to find out the results of the classification of student achievement at the school, the data mining method of Classification with the Naïve Bayes Algorithm was used [71, 72]. The next opinion [73], states that Naïve Bayes is a method of probabilistic reasoning, and the Naïve Bayes algorithm aims to classify data into certain classes.

Analysis and Evaluation Results

The results of Naïve Bayes calculations to classify student score data visualized through graphs were obtained in tests = 1 to 5, then tests 7-10 were 100%, and the lowest accuracy in test = 6 was 66.7%. The average accuracy value of the Naïve Bayes classifier algorithm from the total of all tests = 10, with an average of 96.67%. In line with the results of research [74], it was stated that using the Naïve Bayes Classifier, apart from being able to monitor students' semester grades, this application can provide academic information, a gallery of student activities, and can determine each student's ranking in class with a probability of accuracy of 66.94%.

Further research by Parvati and Belavgi [75], stated that in this research the Naïve Bayes algorithm and the accuracy obtained was 100%, indicating that the decision support system that had been created was successful and ran according to the expected results and obtained a percentage result of 90% so from these results it could be concluded that the decision support system used an algorithm Naïve Bayes works and can be used. Research results [76], show that by using the Naïve Bayes Classifier, you can monitor and determine class rankings so that it becomes a benchmark in schools by showing the results of achieving academic grades in each semester. Apart from monitoring students' semester grades, the application can provide academic information, a gallery of student activities, and determine the ranking of each student. The results of the assessment system testing analysis involving 30 teachers and 200 students as respondents, the following is the test results data are shown in Table 3.

Table 3. Teacher and student responses

Testing Aspect	Percentage of Teacher Satisfaction	Percentage of Student Satisfaction
User Interface	90%	85%
Readability and Navigation	88%	82%
Assessment Functionality	92%	88%
Use of Advanced Technology	85%	80%
User Ease of Learning	91%	87%
Total Satisfaction Percentage	89.2%	84.4%

Number of teacher respondents: 30
Number of student respondents: 200
Total respondents: 230 (30 teachers + 200 students)
Overall satisfaction percentage: (89.2+84.4)/2=86.8%.

In Table 3, the average percentage of teacher and student satisfaction reached 86.8%, indicating an overall positive level of satisfaction. Although the results were positive, it should be noted that student feedback provided a slightly lower percentage of satisfaction compared to teachers. Therefore, it is recommended to consider further feedback from students to improve the overall user experience. In addition, it is important to carry out continuous maintenance and software updates to maintain the quality and relevance of this assessment system.

5.2 Limitations and solutions

Based on the results of previous research and discussion, we recommend several avenues for further exploration based on the findings and discussions presented in this study. Firstly, researchers should delve deeper into methodological limitations that may arise, particularly concerning data collection processes, feature selection, and assumptions underlying the Naïve Bayes algorithm. Secondly, expanding the scope of research to yield more detailed and accurate results by integrating more advanced data collection techniques is crucial.

Thirdly, it is essential to consider the implementation of the assessment system in diverse school environments and its adaptation to various user needs. Lastly, to enhance user satisfaction, researchers are advised to solicit further feedback from students and conduct continuous maintenance and software updates. By addressing these aspects, future researchers can ensure the relevance and quality of the proposed assessment system in enhancing students' learning experiences and academic performance in educational settings.

6. CONCLUSION

This study confirms the effectiveness of a web-based academic assessment system built using the Naïve Bayes Classifier algorithm. By integrating key academic indicators such as test scores, attendance, and class participation, the system demonstrated high reliability in predicting student performance levels, with an average classification accuracy of 96.67%. The implementation of the system not only improved data management but also enhanced the objectivity and depth of the assessment process. The favorable acceptance rate of 86.8% among 230 respondents, including students and teachers, highlights the system's usability and relevance in practical educational environments. These findings underscore the model's potential to transform conventional assessment practices into more efficient, transparent, and data-driven approaches that support individualized learning pathways. However, one limitation of this study lies in the scope of the dataset, which was restricted to a single institution. This may limit the generalizability of the results across diverse educational settings.

Further studies are encouraged to explore the integration of more advanced models, such as ensemble techniques or deep learning algorithms, particularly when handling complex or large-scale educational data. Longitudinal studies are also recommended to assess the sustained impact of such systems on teaching practices and student achievement over time. Additionally, incorporating features like real-time analytics, multilingual interfaces, and adaptive feedback could further increase the system's scalability and usability. Collaborative research across institutions will be valuable to validate the model's applicability in various curricular and demographic contexts.

AUTHOR CONTRIBUTIONS

All authors have contributed to the results of this study. They start from research preparation, looking for references, system design, system testing, and manuscript preparation, until the results of the last revision.

FUNDING

This research received funding from the Government of Indonesia through the Ministry of Education, Culture, Research, and Technology (Kemdikbudristek/Diktisaintek) under Grant Number: 130/C3/DT.05.00/PL/2025.

ACKNOWLEDGMENT

The authors would like to express their profound gratitude to the Government of Indonesia for its generous financial support provided through the Ministry of Education, Culture, and Technology (Diktisaintek), which made this research possible. The authors further acknowledge the invaluable contributions of the school principals and teaching staff who facilitated data collection and field activities. Sincere appreciation is also extended to Universitas Teknologi Akba Makassar and Universitas Negeri Yogyakarta for their continuous institutional support throughout the execution and completion of this study.

REFERENCES

[1] Suyuti, H., Setyanto, A. (2023). The use of augmented reality technology in preserving cultural heritage: A case study of Old Jami Mosque of Palopo. Ceddi Journal of Information System and Technology (JST), 2(1): 28-37. https://doi.org/10.56134/jst.v2i1.35

[2] Hashim, A.S., Awadh, W.A., Hamoud, A.K. (2020). Student performance prediction model based on supervised machine learning algorithms. IOP Conference Series: Materials Science and Engineering, 928(3): 032019. https://doi.org/10.1088/1757-899X/928/3/032019

[3] Bujang, S.D.A., Selamat, A., Ibrahim, R., Krejcar, O., Herrera-Viedma, E., Fujita, H., Ghani, N.A.M. (2021). Multiclass prediction model for student grade prediction using machine learning. IEEE Access, 9: 95608-95621. https://doi.org/10.1109/ACCESS.2021.3093563

[4] Ida, N.L., Arifin, R. (2023). Android-Based Personal Hygiene Educational Game Application for Children. Ceddi Journal of Education, 2(2): 36-44. https://doi.org/10.56134/cje.v2i2.46

[5] Thant, K.S., Thu, E.T.T., Khaing, M.M., Myint, K.L., Tin, H.H.K. (2021). Evaluation of student academic performance using Naïve Bayes classifier. Advances in Computer and Communications, 1(1): 46-52. https://doi.org/10.26855/acc.2020.12.005

[6] Matraeva, A.D., Rybakova, M.V., Vinichenko, M.V., Oseev, A.A., Ljapunova, N.V. (2020). Development of creativity of students in higher educational institutions: Assessment of students and experts. Universal Journal of Educational Research, 8(1), 8-16. https://doi.org/10.13189/ujer.2020.080102

[7] Yanes, N., Mostafa, A.M., Ezz, M., Almuayqil, S.N. (2020). A machine learning-based recommender system for improving students learning experiences. IEEE Access, 8: 201218-201235. https://doi.org/10.1109/ACCESS.2020.3036336

[8] Pallathadka, H., Wenda, A., Ramirez-Asís, E., Asís-López, M., Flores-Albornoz, J., Phasinam, K. (2023). Classification and prediction of student performance data using various machine learning algorithms. Materials Today: Proceedings, 80: 3782-3785. https://doi.org/10.1016/j.matpr.2021.07.382

[9] Sathe, M.T., Adamuthe, A.C. (2021). Comparative study of supervised algorithms for prediction of students' performance. International Journal of Modern Education and Computer Science, 15(1): 1. https://doi.org/10.5815/ijmecs.2021.01.01

[10] Ali, Z.M., Hassoon, N.H., Ahmed, W.S., Abed, H.N. (2020). The application of data mining for predicting academic performance using k-means clustering and naïve bayes classification. International Journal of Psychosocial Rehabilitation, 24(3): 2143-2151. https://doi.org/10.37200/IJPR/V24I3/PR200962

[11] Perez, J.G., Perez, E.S. (2021). Predicting student program completion using Naïve Bayes classification algorithm. International Journal of Modern Education and Computer Science, 12(3): 57. https://doi.org/10.5815/ijmecs.2021.03.05

[12] Meiriza, A., Lestari, E., Putra, P., Monaputri, A., Lestari, D.A. (2020). Prediction graduate student use naive bayes classifier. In Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019) (pp. 370-375). Atlantis Press. https://doi.org/10.2991/aisr.k.200424.056

[13] Injadat, M., Moubayed, A., Nassif, A.B., Shami, A. (2020). Systematic ensemble model selection approach for educational data mining. Knowledge-Based Systems, 200: 105992.

https://doi.org/10.1016/j.knosys.2020.105992
[14] Parkavi, R., Karthikeyan, P., Sheik Abdullah, A. (2024). Predicting academic performance of learners with the three domains of learning data using neuro-fuzzy model and machine learning algorithms. Journal of Engineering Research, 12(3): 397-411. https://doi.org/10.1016/j.jer.2023.09.006
[15] Norz, L.M., Dornauer, V., Hackl, W.O., Ammenwerth, E. (2023). Measuring social presence in online-based learning: An exploratory path analysis using log data and social network analysis. The internet and higher education, 56: 100894. https://doi.org/10.1016/j.iheduc.2022.100894
[16] Farhana, S. (2021). Classification of academic performance for university research evaluation by implementing modified Naive Bayes algorithm. Procedia Computer Science, 194: 224-228. https://doi.org/10.1016/j.procs.2021.10.077
[17] Sembiring, M.T., Tambunan, R.H. (2021). Analysis of graduation prediction on time based on student academic performance using the Naïve Bayes Algorithm with data mining implementation (Case study: Department of Industrial Engineering USU). IOP Conference Series: Materials Science and Engineering, 1122(1): 012069. https://doi.org/10.1088/1757-899X/1122/1/012069
[18] Hasan, R., Palaniappan, S., Mahmood, S., Sarker, K.U., Abbas, A. (2020). Modelling and predicting student's academic performance using classification data mining techniques. International Journal of Business Information Systems, 34(3): 403-422. https://doi.org/10.1504/IJBIS.2020.108649
[19] Dogan, A., Birant, D. (2021). Machine learning and data mining in manufacturing. Expert Systems with Applications, 166: 114060. https://doi.org/10.1016/j.eswa.2020.114060
[20] Issah, I., Appiah, O., Appiahene, P., Inusah, F. (2023). A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. Decision Analytics Journal, 7: 100204. https://doi.org/10.1016/j.dajour.2023.100204
[21] Balaji, V.R., Suganthi, S.T., Rajadevi, R., Kumar, V. K., Balaji, B.S., Pandiyan, S. (2020). Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier. Measurement, 163: 107922. https://doi.org/10.1016/j.measurement.2020.107922
[22] Vishwakarma, M., Kesswani, N. (2023). A new two-phase intrusion detection system with Naïve Bayes machine learning for data classification and elliptic envelop method for anomaly detection. Decision Analytics Journal, 7: 100233. https://doi.org/10.1016/j.dajour.2023.100233
[23] Gu, Y. (2023). Exploring the application of teaching evaluation models incorporating association rules and weighted naive Bayesian algorithms. Intelligent Systems with Applications, 20: 200297. https://doi.org/10.1016/j.iswa.2023.200297
[24] Sawhney, R., Malik, A., Sharma, S., Narayan, V. (2023). A comparative assessment of artificial intelligence models used for early prediction and evaluation of chronic kidney disease. Decision Analytics Journal, 6: 100169. https://doi.org/10.1016/j.dajour.2023.100169
[25] Quilez-Robres, A., González-Andrade, A., Ortega, Z., Santiago-Ramajo, S. (2021). Intelligence quotient, short-term memory and study habits as academic achievement predictors of elementary school: A follow-up study. Studies in Educational Evaluation, 70: 101020. https://doi.org/10.1016/j.stueduc.2021.101020
[26] Khajenezhad, A., Bashiri, M.A., Beigy, H. (2021). A distributed density estimation algorithm and its application to naive Bayes classification. Applied Soft Computing, 98: 106837. https://doi.org/10.1016/j.asoc.2020.106837
[27] Blanquero, R., Carrizosa, E., Ramírez-Cobo, P., Sillero-Denamiel, M.R. (2021). Variable selection for Naïve Bayes classification. Computers & Operations Research, 135: 105456. https://doi.org/10.1016/j.cor.2021.105456
[28] Durak, A., Bulut, V. (2024). Classification and prediction-based machine learning algorithms to predict students' low and high programming performance. Computer Applications in Engineering Education, 32(1): e22679. https://doi.org/10.1002/cae.22679
[29] Schultz, M., Young, K., K. Gunning, T., Harvey, M.L. (2022). Defining and measuring authentic assessment: A case study in the context of tertiary science. Assessment & Evaluation in Higher Education, 47(1): 77-94. https://doi.org/10.1080/02602938.2021.1887811
[30] Velmurugan, P.R., Swadhi, R., Varshney, K.R., Regins, J.C., Gayathri, K. (2025). Creating engaging and personalized learning experiences in distance education: AI and learning analytics. In AI and Learning Analytics in Distance Learning, pp. 103-126. IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-7195-4.ch005
[31] Meylani, R. (2024). A comparative analysis of traditional and modern approaches to assessment and evaluation in education. Batı Anadolu Eğitim Bilimleri Dergisi, 15(1): 520-555. https://doi.org/10.51460/baebd.1386737
[32] Saputra, N., Putera, R.E., Zetra, A., Azwar, Valentina, T.R., Mulia, R.A. (2024). Capacity building for organizational performance: A systematic review, conceptual framework, and future research directions. Cogent Business & Management, 11(1): 2434966. https://doi.org/10.1080/23311975.2024.2434966
[33] Lin, L., Zhou, D., Wang, J., Wang, Y. (2024). A systematic review of big data driven education evaluation. Sage Open, 14(2): 21582440241242180. https://doi.org/10.1186/s40561-023-00262-w
[34] Ngartera, L., Issaka, M.A., Nadarajah, S. (2025). Hybrid Naïve bayes models for scam detection: Comparative insights from email and financial fraud. IEEE Access, 13: 85207-85216. https://doi.org/10.1109/ACCESS.2025.3569216
[35] Nakhipova, V., Kerimbekov, Y., Umarova, Z., Suleimenova, L., Botayeva, S., Ibashova, A., Zhumatayev, N. (2024). Use of the naive Bayes classifier algorithm in machine learning for student performance prediction. International Journal of Information and Education Technology, 14(1): 92-98. https://doi.org/10.18178/ijiet.2024.14.1.2028
[36] Raji, N.R., Kumar, R.M.S., Biji, C.L. (2024). Explainable machine learning prediction for the academic performance of deaf scholars. IEEE Access, 12: 23595-23612. https://doi.org/10.1109/ACCESS.2024.3363634
[37] Ziernwald, L., Hillmayr, D., Holzberger, D. (2022). Promoting high-achieving students through differentiated instruction in mixed-ability classrooms—

A systematic review. Journal of Advanced Academics, 33(4): 540-573. https://doi.org/10.1177/1932202X221112931
[38] Alam, A. (2022). Employing adaptive learning and intelligent tutoring robots for virtual classrooms and smart campuses: Reforming education in the age of artificial intelligence. In: Shaw, R.N., Das, S., Piuri, V., Bianchini, M. (eds) Advanced Computing and Intelligent Technologies. Lecture Notes in Electrical Engineering, vol 914. Springer, Singapore. https://doi.org/10.1007/978-981-19-2980-9_32
[39] Cardilini, A.P., Risely, A., Richardson, M.F. (2022). Supervising the PhD: Identifying common mismatches in expectations between candidate and supervisor to improve research training outcomes. Higher Education Research & Development, 41(3): 613-627. https://doi.org/10.1080/07294360.2021.187488
[40] Albreiki, B., Rustamov, Z., Rustamov, J., Zaki, N. (2024). Enhancing Predictive Performance in Identifying At-Risk Students: Integration of Topological Features, Node Embeddings in Machine Learning Models. In: Khine, M.S. (eds) Machine Learning in Educational Sciences. Springer, Singapore. https://doi.org/10.1007/978-981-99-9379-6_10
[41] Pesovski, I., Jolakoski, P., Trajkovik, V., Kubincova, Z., Herzog, M.A. (2025). Predicting student achievement through peer network analysis for timely personalization via generative AI. Computers and Education: Artificial Intelligence, 8: 100430. https://doi.org/10.1016/j.caeai.2025.100430
[42] Jui, T.D., Rivas, P. (2024). Fairness issues, current approaches, and challenges in machine learning models. International Journal of Machine Learning and Cybernetics, 15(8): 3095-3125. https://doi.org/10.1007/s13042-023-02083-2
[43] Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B., Tabona, O. (2021). A survey on missing data in machine learning. Journal of Big Data, 8(1): 140. https://doi.org/10.1186/s40537-021-00516-9
[44] Lopatina, H., Tsybuliak, N., Popova, A., Hurenko, O., Suchikova, Y. (2024). Inclusive education in higher education institution: Are Ukrainian faculty members' ready for it? Research in Education, 118(1): 49-72. https://doi.org/10.1177/003452372312077
[45] Shams, M.Y., Gamel, S.A., Talaat, F.M. (2024). Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making. Neural Computing and Applications, 36(11): 5695-5714. https://doi.org/10.1007/s00521-023-09391-2
[46] Ratta, P., Sharma, S. (2024). A blockchain-machine learning ecosystem for IoT-Based remote health monitoring of diabetic patients. Healthcare Analytics, 5: 100338. https://doi.org/10.1016/j.health.2024.100338
[47] Shah, S.A.A., Zaman, Q. (2025). Optimal selection of a probabilistic machine learning model for predicting high run chase outcomes in T-20 international cricket. Journal of Sports Sciences, 1-19. https://doi.org/10.1080/02640414.2025.2488157
[48] Zhang, L., Haldorai, A., Naik, N. (2024). GAN-Augmented Naïve Bayes for identifying high-risk coronary artery disease patients using CT angiography data. Scientific Reports, 14(1): 23278. https://doi.org/10.1038/s41598-024-73176-3
[49] Trakunphutthirak, R., Lee, V.C. (2022). Application of educational data mining approach for student academic performance prediction using progressive temporal data. Journal of Educational Computing Research, 60(3): 742-776. https://doi.org/10.1177/07356331211048777
[50] Arya, S., Aggarwal, S., Soni, N., Nishant, N., Ansar, S.A. (2025). Explainable Artificial Intelligence (XAI) in Critical Decision-Making Processes. In: Hassanien, A.E., Anand, S., Jaiswal, A., Kumar, P. (eds) Innovative Computing and Communications. ICICC 2024. Lecture Notes in Networks and Systems, vol 1039. Springer, Singapore. https://doi.org/10.1007/978-981-97-4152-6_32
[51] Niu, T., Liu, T., Luo, Y.T., Pang, P.C.I., Huang, S., Xiang, A. (2025). Decoding student cognitive abilities: A comparative study of explainable AI algorithms in educational data mining. Scientific Reports, 15(1): 26862. https://doi.org/10.1038/s41598-025-12514-5
[52] Sahlaoui, H., Alaoui, E.A.A., Agoujil, S., Nayyar, A. (2024). An empirical assessment of smote variants techniques and interpretation methods in improving the accuracy and the interpretability of student performance models. Education and Information Technologies, 29(5): 5447-5483. https://doi.org/10.1007/s10639-023-12007-w
[53] Gaftandzhieva, S., Talukder, A., Gohain, N., Hussain, S., Theodorou, P., Salal, Y.K., Doneva, R. (2022). Exploring online activities to predict the final grade of student. Mathematics, 10(20): 3758. https://doi.org/10.3390/math10203758
[54] Gao, R., Merzdorf, H.E., Anwar, S., Hipwell, M.C., Srinivasa, A.R. (2024). Automatic assessment of text-based responses in post-secondary education: A systematic review. Computers and Education: Artificial Intelligence, 6: 100206. https://doi.org/10.1016/j.caeai.2024.100206
[55] Huang, A.Y., Lu, O.H., Huang, J.C., Yin, C.J., Yang, S.J. (2020). Predicting students' academic performance by using educational big data and learning analytics: Evaluation of classification methods and learning logs. Interactive Learning Environments, 28(2): 206-230. https://doi.org/10.1080/10494820.2019.1636086
[56] Guleria, P., Sood, M. (2023). Explainable AI and machine learning: Performance evaluation and explainability of classifiers on educational data mining inspired career counseling. Education and Information Technologies, 28(1): 1081-1116. https://doi.org/10.1007/s10639-022-11221-2
[57] Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. Smart Learning Environments, 9(1): 11. https://doi.org/10.1186/s40561-022-00192-z
[58] Babu, J.C., Suresh, Y., Kiran, A., Babu, A.R., Reddy, C.M.K. (2024). Google Appstore Data Classification Using ML Based Naïve's Bayes Algorithm: A Review. In: Gunjan, V.K., Zurada, J.M., Singh, N. (eds) Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough. Studies in Computational Intelligence, vol 1117. Springer, Cham. https://doi.org/10.1007/978-3-031-43009-1_14
[59] Demilie, W.B. (2024). Plant disease detection and classification techniques: a comparative study of the performances. Journal of Big Data, 11(1): 5. https://doi.org/10.1186/s40537-023-00863-9

[60] Yesilkaya, B., Sayilgan, E., Yuce, Y.K., Perc, M., Isler, Y. (2023). Principal component analysis and manifold learning techniques for the design of brain-computer interfaces based on steady-state visually evoked potentials. Journal of Computational Science, 68: 102000. https://doi.org/10.1016/j.jocs.2023.102000

[61] Vital, T.P., Sangeeta, K., Kumar, K.K. (2021). Student classification based on cognitive abilities and predicting learning performances using machine learning models. International Journal of Computing and Digital Systems, 10(1): 63-75. https://doi.org/10.12785/ijcds/100107

[62] Xu, W., Yang, L. Y., Liu, X., Jin, P.N. (2024). Examining the effects of different forms of teacher feedback intervention for learners' cognitive and emotional interaction in online collaborative discussion: A visualization method for process mining based on text automatic analysis. Education and Information Technologies, 29(6): 6525-6551. https://doi.org/10.1007/s10639-023-12097-6

[63] Kumar, S., Kumar, N., Dev, A., Naorem, S. (2023). Movie genre classification using binary relevance, label powerset, and machine learning classifiers. Multimedia Tools and Applications, 82(1): 945-968. https://doi.org/10.1007/s11042-022-13211-5

[64] Kukkar, A., Mohana, R., Sharma, A., Nayyar, A. (2024). A novel methodology using RNN+ LSTM+ ML for predicting student's academic performance. Education and Information Technologies, 29(11): 14365-14401. https://doi.org/10.1007/s10639-023-12394-0

[65] Trung, T., Hoang, A.D., Nguyen, T.T., Dinh, V.H., Nguyen, Y.C., Pham, H.H. (2020). Dataset of Vietnamese student's learning habits during COVID-19. Data in Brief, 30: 105682. https://doi.org/10.1016/j.dib.2020.105682

[66] Smys, S., Haoxiang, W. (2021). Naïve Bayes and entropy based analysis and classification of humans and chat bots. Journal of ISMAC, 3(1): 40-49. https://doi.org/10.36548/jismac.2021.1.004

[67] Phoenix, P., Sudaryono, R., Suhartono, D. (2021). Classifying promotion images using optical character recognition and Naïve Bayes classifier. Procedia Computer Science, 179: 498-506. https://doi.org/10.1016/j.procs.2021.01.033

[68] Lee, B.T., Barber, G.P., Benet-Pagès, A., Casper, J., et al. (2022). The UCSC genome browser database: 2022 update. Nucleic Acids Research, 50(D1): D1115-D1122. https://doi.org/10.1093/nar/gkab959

[69] Sadriddinovich, J.T. (2023). Capabilities of SPSS software in high volume data processing testing. American Journal of Public Diplomacy and International Studies (2993-2157), 1(9): 82-86.

[70] Feng, G., Fan, M. (2024). Research on learning behavior patterns from the perspective of educational data mining: Evaluation, prediction and visualization. Expert Systems with Applications, 237: 121555. https://doi.org/10.1016/j.eswa.2023.121555

[71] Vangara, V., Vangara, S.P., Thirupathur, K. (2020). Opinion mining classification using naive bayes algorithm. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(5): 495-498. https://doi.org/10.35940/ijitee.E2402.039520

[72] Li, Q.N., Li, T.H. (2020). Research on the application of Naive Bayes and Support Vector Machine algorithm on exercises Classification. Journal of Physics: Conference Series, 1437(1): 012071. https://doi.org/10.1088/1742-6596/1437/1/012071

[73] Muhammad, B.A., Qi, C., Wu, Z., Ahmad, H.K. (2024). An evolving learning style detection approach for online education using bipartite graph embedding. Applied Soft Computing, 152: 111230. https://doi.org/10.1016/j.asoc.2024.111230

[74] Battiston, P., Gamba, S., Santoro, A. (2024). Machine learning and the optimization of prediction-based policies. Technological Forecasting and Social Change, 199: 123080. https://doi.org/10.1016/j.techfore.2023.123080

[75] Parvati, V., Belavgi, A. (2024). Optimized deep network based students performance analysis for college admissions. Multimedia Tools and Applications, 83(24): 64369-64391. https://doi.org/10.1007/s11042-024-18135-w

[76] Kurup, P.S., Jeba, S.A. (2024). A secure and efficient student performance prediction system based on blockchain and data mining technologies. In Blockchain Intelligent Systems, pp. 209-227. CRC Press.