Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 2, April 2026, pp. 1669~1678
ISSN: 2302-9285, DOI: 10.11591/eei.v15i2.8029

Machine learning techniques for analyzing student's performance in Islamic Studies

Laili Ramadani[1], Eva Ardinal[2], Muhiddinur Kamal[3], Mahyudin Ritonga[4], Julhadi[4], Juliwis Kardi[1], Nuraiman[5]
[1]Islamic Education Study Program, Sekolah Tinggi Ilmu Tarbiyah Diniyyah Puteri Rahmah El Yunusiyyah, Padang Panjang, Indonesia
[2]Study Program Islamic Education, Institut Agama Islam Negeri Kerinci, Jambi, Indonesia
[3]Universitas Islam Negeri Sjech M. Djamil Djambek, Bukittinggi, Indonesia
[4]Department of Islamic Studies, Universitas Muhammadiyah Sumatera Barat, Padang, Indonesia
[5]Study Program Islamic Studies, Sekolah Tinggi Agama Islam Yayasan Dakwah Islam, Lubuk Sikaping, Indonesia

Article Info

Article history:

Received Dec 14, 2023
Revised Oct 28, 2025
Accepted Mar 5, 2026

Keywords:

Academic performance analysis
Classification
Feature selection
Forward selection
Fuzzy support vector machines
Islamic Studies
Machine learning

ABSTRACT

Many learning institutions and organizations are currently faced with the acute burden of trying to forecast the academic performance of their students. This paper reflects the application of machine learning tools to discuss the potential and performance of students in Islamic Studies. The framework suggested in this paper, will start with the acquisition of the historical data of the students in the input dataset. First, the forward selection wrapper method is used to select the most meaningful features thus eliminating the redundant qualities in the set of student data. Three types of classifiers are then used to create a classification model based on fuzzy support vector machines (SVM), K-nearest neighbors (K-NN), and Naive Bayes. In such a methodological approach, academic performance is predicted and results measured according to certain criteria. According to results of the experiment, it is noted that feature selection-fuzzy support vector machine (FS-fuzzy SVM) has an excellent accuracy of 99.9% with a sensitivity of 98.50% and a specificity of 98.50% and it is therefore seen to be more effective in predicting the academic performance of students in Islamic Studies.

Corresponding Author:

Mahyudin Ritonga
Department of Islamic Studies, Universitas Muhammadiyah Sumatera Barat
Padang, Indonesia
Email: mahyudinritonga@gmail.com

1. INTRODUCTION

When educational institutions adopt a more digital approach, it becomes simpler for teachers and other professionals working in the field of education to have access to information. The area of education has seen a significant transformation as a result of the incorporation of technology into numerous various aspects of the learning process. The amount of data that is stored in establishments is undergoing a transformation that was not anticipated as a result of the transition from offline to online mastering patterns and organization management. Because of the usage of learning management systems (LMSs), there has been an explosion of data concerning the behaviors, interactions, and performance of students [1]. Educational data mining (EDM) is becoming vital in resource optimization in educational institutions [2]. This enormous collection of student data presents a once-in-a-lifetime opportunity to get a deeper understanding of the ways in which students acquire knowledge and the ways in which educational outcomes might be improved. According to authors

[3], [4], early intervention that is based on predetermined expectations for student performance can be utilized even when instructional strategies are modified to fit a wide range of student backgrounds.

Universities play a crucial role in the process of creating, collecting, and disseminating knowledge in an environment that is highly competitive. In order to compensate for the limited resources, they have, they need to generate knowledge not just inside their own organization but also through collaborations with other organizations, such as enterprises, other educational institutions, and the community. Data mining is utilized in order to accomplish this goal. Over the course of the past few years, data mining has evolved as a topic of study that is at the forefront of its field. Data mining is a technique that encompasses a broad variety of methodologies that are utilized in fields as different as business and academics. Educational institutions are required to implement certain processes and techniques in order to get valuable insights from their historical and real-time data. This is because it involves a wide variety of activities and requires them to employ unique procedures and methods. As a result of the incorporation of information and communication technology into academic activities, universities are continuously modifying the manner in which they take in information and knowledge. It is imperative that educational institutions acknowledge and adjust to these developments if they wish to maintain their commitment to providing students with an education of the highest caliber [5].

The most recent technology advancements have resulted in a substantial increase in the amount of data that is being collected across all sectors of the economy. Because of this, there is a growing interest in mining this mountain of data for insights that have not been seen before. A broad variety of techniques are made available by data mining, and these techniques have a number of practical applications in a variety of sectors, including biology, telecommunications, education, retail, finance, and sales, among others. These techniques are helpful in the process of extracting usable data, which is necessary in order to satisfy the requirements of the industry. In this age of vast data storage, it is vital to have powerful tools for analyzing and interpreting data that is kept in repositories such as files and databases in order to extract relevant information for decision-making purposes. We have the ability to discover knowledge that is both helpful and pertinent by utilizing data mining. Workers who have completed their education are in high demand in today's society, which is undergoing fast change. It is a challenging endeavor to enhance the teaching methods utilized by educational institutions such as colleges and universities. Data mining is the most effective technique for assisting educators, students, alumni, managers, and other staff members in gaining additional insights from the massive volumes of data that are currently available to them. Graduate education, which is a vital component of the educational system in the country, serves as the basis for both creativity and competitiveness. One of the primary objectives of higher education systems is to continuously evaluate the quality of education in order to provide a better learning experience.

The utilization of data mining has the potential to fill in information gaps that exist within these systems. With the help of data mining technologies, we are able to discover patterns, connections, and outliers that were not previously visible. This, in turn, allows us to increase the effectiveness, efficiency, and speed of our operations. These enhancements will help the higher education system work more smoothly, increase graduation and retention rates, make it easier for students to transfer between schools, improve educational outcomes, and reduce costs. The system stands to gain a great lot from these upgrades. To raise the bar for student learning, those in charge of higher education policy should be able to use a data mining system to gather the answers they need to determine how to improve the learning environment [6]. According to Malvandi and Farahi [7] the incorporation of information and communication technologies (ICTs) into educational processes ensures that educational institutions are able to adapt to the ever-changing environment.

The learning effects of students are a good indicator of their abilities and capabilities; hence, the behavioral changes that students go through as a result of their experiences may be used to define their overall performance in an educational environment [8]. The activities of learners may provide insight into the degree to which they have improved their thinking, physical fitness, or factual understanding as a result of acquisition of information. Findings from the technique for altering students' conduct outside of class. Students' capability to make decisions based on their cognitive and motor abilities, as well as their grasp of the material being taught, are all observable markers of how well they are functioning. Other indicators include their growth in data processing and their capacity to comprehend the material. Students' performance in the areas of knowledge, attitude, and skills can be evaluated based on the outcomes of a series of classes, which can offer the basis for the evaluation.

Students used to be expelled from school if they did not perform well in any aspect of their education, including academics. The grades that the pupils received left both the teachers and the parents bewildered. It is essential to make an effort to anticipate student achievement for a number of reasons, including the efficient assignment of classes and resources at different levels. It is possible, in part, to ensure that educational institutions make effective use of their resources and that students receive the assistance they

require in order to achieve their goals. It is possible that a data mining system might prove to be an invaluable tool throughout this process by providing decision-makers with insights that can be utilized to the advantage of students and to sensibly distribute resources. According to Namoun and Alshanqiti [9], a with the help of counts of students raised hands, trips to activity resources, views of announcements, involvement in discussions, absences, and other student actions, we are able to make reasonable judgments regarding the performance of the students.

Through the use of a single prediction, we are able to provide students with guidance on how to improve their performance [10]. If the institution wishes to improve its position on a national and international scale, it is essential to keep track of the development that its students are making. The majority of educational institutions and organizations are currently facing an urgent requirement to forecast the performance of their pupils. Everyone who is involved, including students, teachers, and mentors, may benefit from having access to a timely prediction. As a consequence of this, mentors may provide students with additional help in order to foster improvements in their academic performance. According to the findings [11], in today's world, every school is responsible for managing and processing enormous amounts of student data. These data may be used to construct machine learning models that can predict the marks that students will receive. Vijayalakshmi and Venkatachalapathy [12] studied the decision tree, linear support vector classifier (SVC), random forest, stochastic gradient descent (SGD), and linear discriminant analysis is all examples of classification methods that are utilized for the purpose of prediction.

Machine learning is also important in making the analysis and the monitoring of the student performance in the field of Islamic Studies more useful as it helps to detect the illicit patterns and relation in educational data. By using machine learning models, data-driven insights can be used to analyze academic history, behavior patterns, attendance, and engagement of students to forecast the level of learning outcomes with high accuracy. Using fuzzy support vector machines (SVM), K-nearest neighbors (K-NN), and Naive Bayes, the institutions can automatically categorize the students according to their performance levels as well as learning potential. This predictive ability helps educators to deliver timely interventions, learning materials that are personalized, and help poorly performing students. Besides, feature selection tools, such as the forward selection wrapper, are useful in filtering down the dataset to make sure that the final data used to make the prediction includes only useful attributes. Accordingly, both the proper prediction of performance and the reinforcement of tactics of decision-making and regular evaluation can be vitalized through machine learning by increasing the adaptability, fairness, and efficacy of the educational process in Islamic Studies.

This article introduces computer learning techniques for analyzing the potential and performance of students in the field of Islamic Studies. The history data of students is obtained from the input dataset. Initially, the forward selection wrapper technique is employed to choose the features. The process of feature selection will effectively remove extraneous attributes from the dataset containing student information. A classification model is constructed utilizing 3 different classifiers: fuzzy SVM, K-NN, and Naive Bayes. Subsequently, academic performance is forecasted and outcomes are evaluated based on specific criteria.

2. LITERATURE SURVEY

If you are thinking about EDM in relation to mining, it is helpful to read Bansal *et al*. [13], which examines a number of different uses of EDM. It is possible that being familiar with all of these applications will make it simpler for you to carry out. Teachers would be able to readily evaluate their students' performance and concentrate on areas in which they need to improve if they took the learning profiles of their pupils and grouped them according to a similarity constraint.

In their research published in 2025, Bhatia *et al*. [14] assessed and investigated a number of different feature selection calculations. According to the findings obtained from the student dataset, the feature selection calculations that are accessible through the Weka tool are not significantly different from what was previously available. In combination with the random forest classifier, the vast majority of the currently available feature selection algorithms have demonstrated better performance in terms of their outcomes. According to the findings of this study, multilayer perceptron (MLP) classifiers performed far better than any other classifiers when it came to the collection of datasets from students. Based on the findings, it is evident that the various algorithms for determining elements require slight adjustments to their parameters in order to achieve higher performance.

As stated by Suyal and Mohod [15], the two most common approaches to doing sentiment analysis are the utilization of a voting group approach and the utilization of the Chi-Square methodology to reduce the number of features. The findings also demonstrated that selecting a trait that has a greater weight expectation is beneficial to the process of making a clear and accurate forecast.

It is the intention of Ade and Deshmukh [16] to bridge the gap between the present machine learning approaches of incremental learning and the exploratory classification of students for the purpose of their career choosing. Because the information capacity is not ideal for batch learning concepts, it is not

applicable in situations where the data is frequently created in a continuous manner. Consequently, the incremental ensemble learning algorithm that was described in the study is an effective method that may be utilized to provide the student with the most advantageous professional alternative that is now available.

It was recommended by Zhang and Qin [17] that the vast volumes of data that have opened up many doors for online education should be enhanced; new mining breakthroughs would surely bring improved ways and approaches for utilizing EDM in online education. Research by Doko and Bexheti [18] present the findings of a research project that aimed to enhance learning intelligence in educational technology through the utilization of videography. In addition to this, it is essential to provide more explanation that Doko and Bexheti [18] has offered on the exam questions and the future research addresses. The results of the investigations were supported by a number of EDM and learning analytics (L.A.) models and analyses on many occasions [18].

A classification strategy was proposed by [19] as a means of predicting the division of students based on previous data. Considering that the Naive Hypothesis is one of the methods that territorial units utilized for the purpose of data classification, we have decided to apply it here. The historical data of the students, which included things like gathering activities, class exams, workshops, and task marks, was utilized in order to make a prediction regarding how well they would perform at the conclusion of the previous semester.

A growing number of educational institutions, including colleges and universities, are employing data mining techniques to examine vast quantities of historical data pertaining to the academic achievement of students. The motivation behind this is the aspiration to enhance the quality of education while simultaneously addressing the ever-growing expenses associated with higher education by increasing the proportion of students who graduate within four years.

It is the purpose of these approaches to aid schools in providing better assistance to both students and teachers. This is accomplished by identifying which students are more likely to fail a certain class and then adjusting their support in accordance with that information. When it comes to addressing this issue, there are two major groups of solutions that make use of supervised learning methodology.

A single regression model is estimated in the first one, and it makes use of features that are associated to the student's prior performance in classes and their interactions with LMSs online. In light of these factors, our model ought to be able to provide an accurate prediction regarding the student's final grade. Utilizing factorization models, which were first developed for recommender systems, the second one makes predictions about the grades that students will receive for a variety of actions inside the course. In order to accurately anticipate the performance of students, multi-relational models were utilized. These models were used to uncover latent components that satisfy student-task and task-skill links. For the purpose of modeling the students' knowledge acquisition and competence growth over time, tensor factorization methodologies were applied. This was done in order to take temporal factors into consideration. In terms of prediction accuracy, factorization models perform better than single-regression models. This is due to the fact that they are able to customize their models to the specific requirements of each individual learner.

3. METHOD

Machine learning techniques for analyzing students' potential and performance in Islamic Studies are presented in this part. The framework depicted in Figure 1 is presented. The history data of students is obtained from the input dataset. Initially, the forward selection wrapper technique is employed to choose the features. The process of feature selection will effectively remove extraneous attributes from the dataset containing student information. A classification model is constructed utilizing 3 different classifiers: fuzzy SVM, K-NN, and Naive Bayes. Subsequently, academic performance is forecasted and outcomes are evaluated based on specific criteria.

Features are selected using forward selection technique. A blank slate of attributes is used as the starting point for the iterative process of forward selection. The technique consists of adding features in an iterative manner and analyzing performance in order to identify whether or not improvements are being made. This method is performed again and again until the performance of the model is not improved by the addition of more variables or features [20].

SVMs and maintenance vector networks (VNNs) are two examples of the numerous supervised learning models that are used in machine learning. These models explore data for the objectives of regression and classification by utilizing associated learning methods. A non-probabilistic binary linear classifier, a SVM training approach takes a set of training examples that are labeled as belonging to one of two categories and uses them to develop a model that assigns new instances to one of those classes. This technique is also known as a SVM training technique. The instances are represented as points in space by a SVM model, with the criteria for the two divisions being separated by a space that is both immediately understood and as big as

is realistically achievable. Within the same area, it is anticipated that new models will be able to fit into a category according to the portion of the gap that they fall into being used [21].

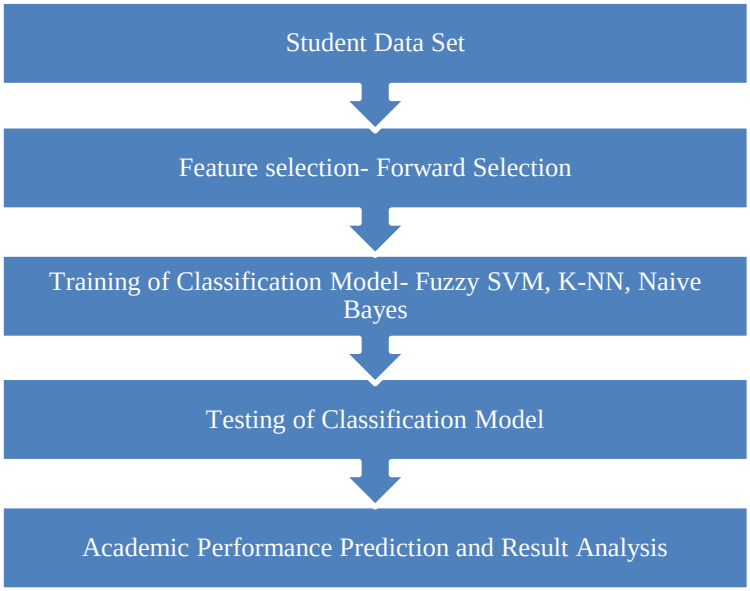

Figure 1. Machine learning techniques for analysis of student's potential and performance in Islamic Studies

The kernel approach enables SVMs to successfully map their inputs into high-dimensional feature spaces, which enables them to do both linear and non-linear classification. This is made possible by the kernel technique. In situations when there are unlabeled records, supervised learning is not likely to occur. In these circumstances, an unsupervised learning technique is necessary, which aims to uncover the underlying grouping of the data into groups before projecting further data to these created clusters. The use of SVMs is extremely advantageous for the categorization of text and hypertext. This is due to the fact that the utilization of SVMs significantly decreases the necessity for labeled preparation examples in both standard inductive and transformative contexts. A fuzzy vector machine technique is utilized to determine fuzzy membership for each sample in order to effectively remove the impact that noise has on SVMs. There are a few different kinds of fuzzy membership function estimate methods that were created, and they are all based on FSVM. This was done in order to improve the performance of SVM. Fuzzy logic is used to perform membership calculations in the feature space, and kernels are used as representations. More accurate generalizations and classifications were achieved as a result [22].

Algorithms such as K-NN are used to categorize things in the feature space according to how close they are to neighboring training examples. On the other hand, K-NN is an example of lazy learning, which is a type of learning that is based on examples and just approximates the function in close proximity to its true position. With K-NN, the complete computation is delayed so that it may be prepared for classification. The most significant problem with the K-NN technique is that it is extremely prone to having its accuracy entirely ruined by the inclusion of characteristics that are either too strong or too inappropriate. Furthermore, if the feature balance is unreliable in relation to its significance, then the accuracy of the feature balance will be bad. When applied to the student evaluation dataset, the closest neighbor technique demonstrates superior performance compared to the other four classifiers implemented in this research. These classifiers include Naive Bayes, decision tree, neural network, SVM, and K-NN, all of which are conducted in incremental order [23].

The construction of a Bayesian classifier is accomplished by employing the idea that the position of a (natural) class is to conduct an analysis of the feature values for members of that class. The examples are arranged in categories according to the values that are shared by the elements. Natural types are a term that is occasionally used to refer to these kinds of courses. The objective feature is linked to an independent class that is not intrinsically binary in this instance. If an agent is able to comprehend the course, then it is able to deduce the principles that govern the other properties. This is the fundamental idea behind the Bayesian classifier methodology. In the event that it is unable to determine the quality, the level can be investigated by employing the Bayes rule and at least some of the feature values.

The learning agent is responsible for constructing a probabilistic model of the attributes in order for Bayesian classifiers to function properly. Following that, it makes use of this model to make a prediction regarding the classification of a new example. Naive Bayes classifiers are extremely scalable since they require linear parameters for learning problem variables (features/predictors). This makes them suitable for distributed learning. Classifiers may be constructed in a straightforward manner using the Naive Bayes approach. The assigning of labels to issue situations is accomplished by these models through the utilization of feature vectors, with the labels being selected from a limited pool of potential labels. In point of fact, it is more of a family of algorithms that are based on a generic concept than it is a particular method for training such classifiers: classification

The Naive Bayes classifiers all make the assumption that the value of a feature is fully independent of the cost of any other feature, provided that the class variable is provided. In spite of its immature design and assumptions that appear to be straightforward, Naive Bayes classifiers have demonstrated remarkable performance in a number of demanding real-world circumstances. Another advantage of the Naive Bayes algorithm is that it requires relatively minimal training data in order to estimate the parameters that are required for classification [24].

4. RESULTS AND DISCUSSION

Data pertaining to a sample of 500 pupils is gathered across 32 distinct categories. These students are studying in different higher educational institutions of Indonesia. The primary characteristics encompassed in this study were school, gender, age, residential address, family size, parental status, maternal and paternal education levels, maternal and paternal employment statuses, travel duration, study duration, academic setbacks, daily activities, leisure time, news consumption, health status, and absenteeism. The evaluation of classifiers' efficacy is a fundamental aspect of machine learning. Various measures can be employed to evaluate a prediction model [25]. Nevertheless, there exist specific measures that are unsuitable for data sets characterized by skewed classes. Therefore, it is imperative to carefully choose metrics that align with the nature of the data set. The forward selection wrapper approach is employed as the primary method for feature selection. By employing this feature selection technique, the student dataset will undergo a process of eliminating any attributes that lack relevance. Subsequently, a classification model is developed through the utilization of fuzzy SVM, K-NN, and Naive Bayes classifier. After completing this task, the student's academic performance is forecasted, and the results are evaluated using a predetermined set of criteria. Tables 1-3 and Figure 2 display the results. The FS-fuzzy SVM achieves a 99.9% accuracy. The FS-fuzzy SVM exhibits a sensitivity of 98.50%. The FS-fuzzy SVM has a specificity of 98.50%.

In order to determine the accuracy of the categorization, divide the total number of occurrences by the number of cases that were successfully separated into categories. The typical circumstances in which it is utilized are those in which the number of instances is distributed uniformly across all classes.

The area under the receiver operating characteristic curve (AUC-ROC) is one method for assessing performance in binary classification. The accuracy of a prediction model is evaluated based on how effectively it can differentiate between positive and negative events. An example of a probability plot is a ROC curve. Both specificity and sensitivity are considered to be the two primary components of ROC.

The sensitivity, which is often referred to as recall, is the proportion of right classifications to total classifications. It is determined by dividing the total number of correct positives by the total number of correct negatives. The percentage of negative situations that were properly expected is referred to as the specificity, which is often referred to as the genuine negative rate. The formula for calculating it is to divide the total number of cases that are negative by the sum of instances that are both negative and false positive.

Table 1. Accuracy comparison of fuzzy SVM, K-NN, and Naive Bayes algorithm for academic performance prediction

Algorithm	Classification accuracy (%)
FS-fuzzy SVM	99
FS-K-NN	96
FS-Naive Bayes	88.5

Table 2. Sensitivity comparison of fuzzy SVM, K-NN, and Naive Bayes algorithm for academic performance prediction

Algorithm	Sensitivity accuracy (%)
FS-fuzzy SVM	98.5
FS-K-NN	94.75
FS-Naive Bayes	87.50

Table 3. Specificity comparison of fuzzy SVM, K-NN, and Naive Bayes algorithm for academic performance prediction

Algorithm	Specificity accuracy (%)
FS-fuzzy SVM	98.5
FS-K-NN	97.75
FS-Naive Bayes	93.5

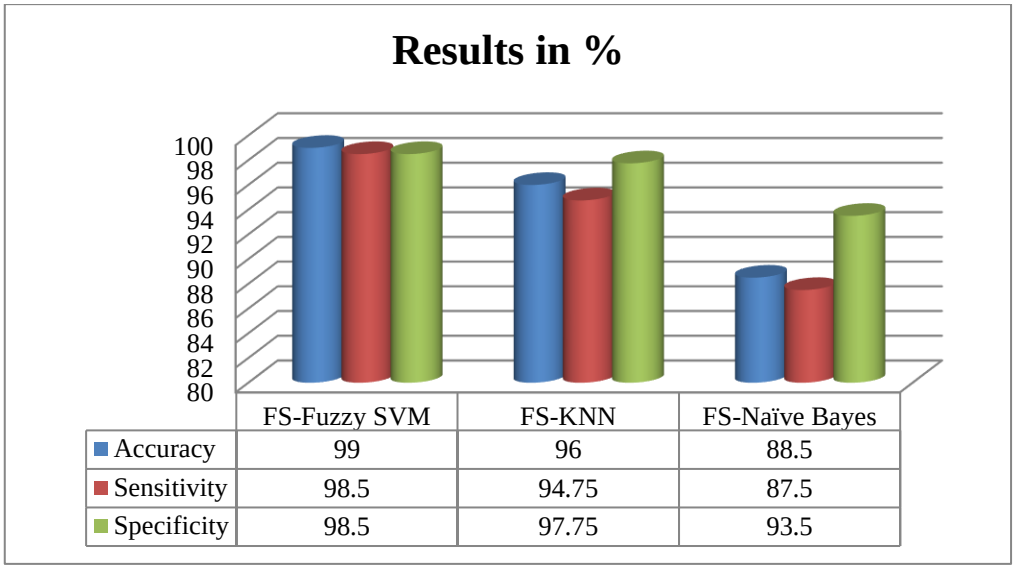

Figure 2. Performance comparison for student performance classification and prediction

5. CONCLUSION

An increasing quantity of educational establishments, encompassing both colleges and universities, are employing data mining methodologies to analyze extensive volumes of historical data pertaining to the academic advancement of students. The rationale behind this is the imperative to enhance the caliber of education while simultaneously tackling the prediction of academic achievement. This article introduces machine learning algorithms designed to evaluate the aptitude and achievement of students in the field of Islamic Studies. The forward selection wrapper technique is employed for the purpose of feature selection. This feature option will be used to remove any irrelevant attributes from the student data collection. In the subsequent step, a classification model is constructed utilizing fuzzy SVM, K-NN, and the Naive Bayes classifier. Subsequently, the student's scholastic achievements are expected, and the results are evaluated according to a predetermined set of criteria. The accuracy of the FS-fuzzy SVM is 99%. The sensitivity of the FS-fuzzy SVM is 98.50%. The specificity of the FS-fuzzy SVM is 98.50%. In further iterations, this proposed model has the potential to be expanded in order to recommend appropriate courses to individual students. In near future, this proposed model can also be used predict employability of Islamic students. Also, the proposed model can be enhanced by applying much in depth input data set from different regions and countries stored inside a cloud based centralized system.

This suggested model can be further developed in additional iterations to prescribe individual learning paths and the right course choices to particular students. Further, the next research can be aimed at predicting the employability and professional preparedness of Islamic Studies students, which is achieved through the incorporation of socio-economic and behavioral data. It is also possible to improve the framework and use a more detailed and multifaceted dataset generated in various regions and institutions and stored in a centralized system based on the cloud. Also, the use of deep learning and hybrid ensembles might additionally enhance the precision of prediction and offer more information about the determinants of academic performance.

ACKNOWLEDGMENTS

We are deeply thankful to our institutions for providing the necessary resources and facilities that enabled the successful completion of this research.

FUNDING INFORMATION

No funding received for this research work.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Laili Ramadani	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Eva Ardinal		✓				✓		✓	✓	✓	✓	✓		
Muhiddinur Kamal	✓	✓	✓	✓	✓	✓		✓		✓	✓		✓	
Mahyudin Ritonga	✓	✓		✓		✓		✓	✓	✓			✓	
Julhadi		✓				✓		✓	✓	✓	✓	✓		
Juliwis Kardi	✓	✓	✓	✓		✓	✓			✓	✓		✓	✓
Nuraiman	✓		✓	✓	✓	✓		✓	✓	✓			✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study will be made available on request

REFERENCES

[1] W. Ahmed, M. A. Wani, P. Plawiak, S. Meshoul, A. Mahmoud, and M. Hammad, "Machine learning-based academic performance prediction with explainability for enhanced decision-making in educational institutions," *Scientific Reports*, vol. 15, pp. 1-21, 2025, doi: 10.1038/s41598-025-12353-4.

[2] J. Gu, "Predicting student academic achievement using stacked ensemble learning with deep neural networks and fuzzy-based feature selection," *Scientific Reports*, vol. 15, pp. 1-20, 2025, doi: 10.1038/s41598-025-20779-z.

[3] M. Ba-Aoum, M. Alrezq, J. Datta, and K. P. Triantis, "Predicting student self-efficacy in Muslim societies using machine learning algorithms," *Frontiers in Big Data*, vol. 7, pp. 1-16, 2024, doi: 10.3389/fdata.2024.1449572.

[4] X. Wen and D. Pan, "Charting new territories: fuzzy systems in English language teaching and learning," *PeerJ Computer Science*, vol. 11, 2025, doi: 10.7717/peerj-cs.2887.

[5] K. Shukla, A. Shukla, and R. Singh, "Neural networks based face recognition system for biometric security," *Indian Journal of Engineering*, vol. 20, no. 53, pp. 1–9, May 2023, doi: 10.54905/disssi/v20i53/e16ije1640.

[6] R. Mahyudin, D. T. Sanchez, S. T. Mrudula, R. K. Veluri, and J. Malik, "Design of Automated Smart Attendance System Using Deep Learning Based Face Recognition," in *2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, Chennai, India, 2024, pp. 1-6, doi: 10.1109/ICONSTEM60960.2024.10568825.

[7] S. Malvandi and A. Farahi, "Provide a method for increasing the efficiency of learning management systems using educational data mining," *Indian Journal of Science and Technology*, vol. 8, no. 28, pp. 1-8, 2015, doi: 10.17485/ijst/2015/v8i28/82454.

[8] S S. Roy and A. Garg, "Analyzing performance of students by using data mining techniques a literature survey," in *2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON)*, Mathura, India, 2017, pp. 130-133, doi: 10.1109/UPCON.2017.8251035.

[9] Namoun and A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," *Applied Sciences*, vol. 11, no. 1, p. 237, 2020, doi: 10.3390/app11010237.

[10] A. Rahman, R. A. Mutiarawan, A. Darmawan, Y. Rianto, and M. Syafrullah, "Prediction of Students Academic Success Using Case Based Reasoning," in *2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, Bandung, Indonesia, 2019, pp. 171-176, doi: 10.23919/EECSI48112.2019.8977104.

[11] S. Alwarthan, N. Aslam, and I. U. Khan, "An Explainable Model for Identifying At-Risk Student at Higher Education," *IEEE Access*, vol. 10, pp. 107649-107668, 2022, doi: 10.1109/ACCESS.2022.3211070.

[12] V. Vijayalakshmi and K. Venkatachalapathy, "Comparison of predicting student's performance using machine learning algorithms," *International Journal of Intelligent Systems and Applications*, vol. 11, no. 12, pp. 34-45, 2019, doi: 10.5815/ijisa.2019.12.04.

[13] R. Bansal, A. Mishra, and S. N. Singh, "Mining of educational data for analysing students' overall performance," in *2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*, Noida, India, 2017, pp. 495-497, doi: 10.1109/CONFLUENCE.2017.7943202.

[14] R. Bhatia, S. Yadav, R. Sharma, Shubneet, A. R. Yadav, and N. S. Talwandi, "A Comparative Study of Feature Selection Techniques for Predicting Student Academic Performance Using Educational Data," in *Proceedings of Data Analytics and Management*, 2025, pp. 352–362, doi: 10.1007/978-3-032-04222-4_30.
[15] S. R. Suyal and M. M. Mohod, "Quality improvisation of student performance using data mining techniques," *International Journal of Scientific and Research Publications*, vol. 4, no. 4, pp. 1-4, 2014.
[16] R. Ade and P. R. Deshmukh, "An incremental ensemble of classifiers as a technique for prediction of student's career choice," in *2014 First International Conference on Networks & Soft Computing (ICNSC2014)*, Guntur, India, 2014, pp. 384-387, doi: 10.1109/CNSC.2014.6906655.
[17] W. Zhang and S. Qin, "A brief analysis of the key technologies and applications of educational data mining on online learning platform," in *2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA)*, Shanghai, China, 2018, pp. 83-86, doi: 10.1109/ICBDA.2018.8367655.
[18] E. Doko and L. A. Bexheti, "A systematic mapping study of educational technologies based on educational data mining and learning analytics," in *2018 7th Mediterranean Conference on Embedded Computing (MECO)*, Budva, Montenegro, 2018, pp. 1-4, doi: 10.1109/MECO.2018.8406052.
[19] M. M. Arcinas *et al.*, "An efficient course recommendation system for higher education students using machine learning techniques," *Bulletin of Electrical Engineering and Informatics*, vol. 14, no. 2, pp. 1468–1475, 2025, doi: 10.11591/eei.v14i2.7711.
[20] J. Xu, W. Lu, J. Li, and H. Yuan, "Dependency maximization forward feature selection algorithms based on normalized cross-covariance operator and its approximated form for high-dimensional data," *Information Sciences*, vol. 617, pp. 416–434, 2022, doi: 10.1016/j.ins.2022.10.093.
[21] C. Huang *et al.*, "A feature weighted support vector machine and artificial neural network algorithm for academic course performance prediction," *Neural Computing and Applications*, vol. 35, no. 16, pp. 11517–11529, 2023, doi: 10.1007/s00521-021-05962-3.
[22] R. Mehdi and M. Nachouki, "A neuro-fuzzy model for predicting and analyzing student graduation performance in computing programs," *Education and Information Technologies*, vol. 28, no. 3, pp. 2455–2484, 2023, doi: 10.1007/s10639-022-11205-2.
[23] S. Hussain and M. Q. Khan, "Student-Performulator: Predicting Students' Academic Performance at Secondary and Intermediate Level Using Machine Learning," *Annals of data science*, vol. 10, no. 3, pp. 637–655, 2023, doi: 10.1007/s40745-021-00341-0.
[24] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learning Environments*, vol. 9, no. 1, p. 11, 2022, doi: 10.1186/s40561-022-00192-z.
[25] M. I. H. Nayan *et al.*, "Comparison of the performance of machine learning-based algorithms for predicting depression and anxiety among university students in Bangladesh: A result of the first wave of the COVID-19 pandemic," *Asian Journal of Social Health and Behavior*, vol. 5, no. 2, p. 75, 2022, doi: 10.4103/shb.shb_38_22.

BIOGRAPHIES OF AUTHORS

Laili Ramadani is a lecture at the Sekolah Tinggi Ilmu Tarbiyah Diniyyah Puteri Rahmah El Yunusiyyah, Padang Panjang, Indonesia. She received the Doctoral degree from the Department of Islamic Education at Islamic State University Imam Bonjol Padang. She currently becomes a member of ADRI and ADPETIKISINDO. She works on Islamic education methodology, research methodology, and Islamic education. She can be contacted at email: lailiramadani86@gmail.com.

Eva Ardinal is a Lecture an Arabic Language and Education at the Faculty of Tarbiyah, Institute Agama Islam Negeri Kerinci. He received the Master degree from the Islamic Education of Postgraduate Islamic State University Imam Bonjol Padang. He works on Arabic Language education, research methodology, and Arabic Language and literature. He can be contacted at email: eardinal@gmail.com.

Muhiddinur Kamal is Associate Professor at Universitas Islam Negeri Sjech M. Djamil Djambek, Bukittinggi, West Sumatra, Indonesia. He received the Doctoral degree from Universitas Negeri Padang. He works on general education, education science, and philosophy of education. He can be contacted at email: muhiddinurkamal@gmail.com.

Mahyudin Ritonga is Professor an Arabic Language and Education at the Faculty of Islamic Studies, Muhammadiyah University of West Sumatra. He received the Doctoral degree from the Islamic Studies at Graduate School of Islamic State University Syarif Hidayatullah Jakarta. He can be contacted at email: mahyudinritonga@gmail.com.

Julhadi is Associate Professor an Islamic Education at the Postgraduate Program, Muhammadiyah University of West Sumatra. He received the Doctoral degree from the Islamic Education at Postgraduate Program of Islamic State University Imam Bonjol Padang. He works on Islamic education methodology, material science of Islamic education, and research methodology. He can be contacted at email: julhadi15@gmail.com.

Juliwis Kardi is Lecture an Islamic Education at the Sekolah Tinggi Ilmu Tarbiyah Diniyyah Puteri Rahmah El Yunusiyyah, Padang Panjang, Indonesia. He received the Doctoral degree from the Islamic Education at Postgraduate Program of Islamic State University Imam Bonjol Padang. He works on media of Islamic education, Islamic education science, and research methodology. He can be contacted at email: juliwiskardi86@gmail.com.

Nuraiman is a Lecture at Sekolah Tinggi Agama Islam Yayasan Dakwah Islam, Lubuk Sikaping, Indonesia. She receives master degree from Postgraduate Program at Islamic State University Imam Bonjol Padang. She works on Islamic and finance and Islamic business. She can be contacted at email: nuraiman@stai-ydi.ac.id.