Received 6 November 2023, accepted 24 November 2023, date of publication 27 November 2023, date of current version 8 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3336987

RESEARCH ARTICLE

Performance Prediction of Students in Higher Education Using Multi-Model Ensemble Approach

NAVEED ANWER BUTT[1], ZAFAR MAHMOOD[1], KHAWAR SHAKEEL[1], SULTAN ALFARHOOD[2], MEJDL SAFRAN[2], AND IMRAN ASHRAF[3]

[1]Department of Computer Science, University of Gujrat, Gujrat 50700, Pakistan
[2]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[3]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea

Corresponding authors: Imran Ashraf (ashrafimran@live.com) and Sultan Alfarhood (sultanf@ksu.edu.sa)

This research is funded by the Researchers Supporting Project Number (RSPD2023R890), King Saud University, Riyadh, Saudi Arabia.

ABSTRACT Many stakeholders including students, teachers, and educational institutions, benefit from accurately predicting student performance and facilitating data-driven policies. In this field, providing users with accurate and understandable predictions is challenging, but equally important. The goals of this study are multifaceted: to identify students at-risk; to identify differences in assessment across different environments; methods for assessing students; and to determine the relationship between teacher employment status and student achievement. This study performs an empirical comparison of the performance and efficiency of ensemble classification methods based on bagging, boosting, stacking, and voting for successful predictions. An ensemble model is developed and validated using double, triple, and quadruple combinations of classification algorithms using Naive Bayes, J48 decision trees, Adaboost, logistics, and multilayer perceptron. This study uses primary quantitative data from the learning management system of a university in Pakistan to analyze the performance of these models. The boosted tree detection method outperforms bagged trees when the standard deviation is higher and the data size is large, while stacking is best for smaller datasets. Based on behavioral analysis results of students, academic advice can be given for selected case studies. These will help educational administrators and policymakers working in education to introduce new policies and curricula accordingly.

INDEX TERMS Ensemble learning, learning analytics, predicting student performance, student education risk analysis.

I. INTRODUCTION

In the modern era, universities operate in highly sensitive, complex, and challenging environments due to the technological revolution. It is a challenge for universities to be aware of their specificities, to truly evaluate their performance, and to plan their policies for future actions and achievements. The education industry is becoming increasingly competitive as the number of institutions continues to increase. To stay alive, these institutions focus more on improving many elements, one of which is quality learning. To provide quality education, institutions need to understand their strengths, both visible and hidden, as well as, their weaknesses. To be competitive, institutions must recognize their largely hidden capabilities and develop plans to exploit them.

As the volume of digital data in education increases, automated analytics that support teaching and learning practices are in high demand [1], [2], [3]. For example, dynamic studies of student transcript data can effectively highlight learning patterns and behaviors and provide predictive models to drive teaching and facilitation efforts, among other interventions [4], [5]. Learning analytics (LA) equips teachers with the tools they need to track each student's development, identify their strengths and areas for improvement, and provide personalized feedback based on each learner's progress [4], [6]. As a result, LA is increasingly

The associate editor coordinating the review of this manuscript and approving it for publication was Yichuan Jiang.

used for online or blended learning to provide active feedback [5], [7], [8], [9] and foundations achieving goals such as identifying learner needs and predicting learning outcomes [10]. According to [11], LA enables educators to identify issues students are struggling with and enables them to provide targeted instruction or process-oriented feedback. As such, LA has a great opportunity to directly impact student self-regulation and academic success. However, due to its complexity, visual reporting in LA has little impact on its audience [6], [12].

The primary uses of LA include learning and predicting student performance, which can be used to suggest changes to current educational practices [13]. The concept of analyzing educational data is not new, but recent developments in computing power, educational technology, and the ability to record fine-grained data have paved the way for the development of new techniques to assess as much data as possible related to the education environment [14]. Furthermore, LA may guide curriculum designers and educational specialists in creating pedagogical foundations while creating syllabuses and changing course materials following students' requirements and the learning environment [15]. LA expands the role and performance of the conventional instructor by allowing him to build alternative models based on the characteristics of the learners in their authentic learning environment [16], [17]. According to the researchers, there is a lack of evidence that LA improves learning outcomes such as "knowledge acquisition, skill development, and cognitive enhancement, as well as learning support and coaching [18], [19], [20].

The purpose of this study is to determine whether there are patterns in relevant data that could be used to predict students' performance based on students' college preparatory and academic characteristics and to examine how teachers evaluate students on multiple campuses. In particular, this study examines educational data from information systems to identify effective and efficient key performance indicators to support complex decision-making by building predictive models to provide various services to all stakeholders in higher education. This study solves the problem by dividing it into two stages. In the first stage, identifying relevant factors as variables of the model helps make better decisions that affect student learning. In this regard, the following research questions (RQs) are formulated

- **RQ1**: Which machine learning models can show better predictions of whether a student will complete his degree or will be dropped from the university?
- **RQ2**: What are the reasons for the poor performance of the majority of the students in some courses?
- **RQ3**: To what extent does the evaluation of the students studying on the main campus differ from the students studying in the sub-campuses?
- **RQ4**: How is the evaluation of the students affected by the job status of the teachers? (Visiting /Contractual/Permanent)?

To answer the formulated research questions, the second stage involves building a predictive modeling structure for selected variables using different classification and ensemble-based techniques. The following objectives are achieved in this study

- To predict student performance and identify under-performing students whose enrollment status is at risk and needs additional support.
- To examine the causes of the student's poor performance in specific courses or domains.
- To analyze the evaluation differences of students in each examination category across multiple campuses and the job status of the teachers.

To summarize, this study aims to investigate how professors assess students across several campuses and to find trends in pertinent data that may be utilized to forecast students' performance based on their academic and college readiness. This study specifically looks at educational data from information systems to find useful and efficient KPIs to help with difficult decision-making. Predictive models are then built to offer a range of services to all parties involved in higher education. The problem is broken down into two steps to solve it. Making better judgments that impact student learning in the initial phase is facilitated by the identification of pertinent aspects as model variables.

In this study, the accuracy and effectiveness of individual classifiers of different types and ensemble classifiers are also empirically tested and compared using data from different related to different research questions. In general, this work adequately addresses the essential questions: which variable combinations are the most reliable predictors of student academic performance? and advances the understanding of how learner data may be used to predict student performance using ensemble approaches. What are the chances of bagging, boosting, stacking, and voting ensemble approaches being used in a trustworthy manner?

The rest of the paper is organized into 3 sections. The literature review is presented in Section II. It is followed by a description of machine learning models in Section III. The methodology adopted in this study is described in Section IV. Results and discussions are given in Section V while Section VI provides the conclusion.

II. LITERATURE REVIEW

Machine learning and deep learning models have been employed in several domains including medical image analysis, object detection, data mining, risk analysis, etc. For example, [21] used machine learning models like SVM and K-NN risks related to global software development projects. Similarly, [22] performs risk prediction related to time, cost, and resources. These factors are investigated in the context of global software development using deep learning models like neural networks, Bayesian approaches, etc. The use of machine learning and deep learning models for academic performance prediction is discussed here.

A. LEARNING ANALYTICS

According to [23], the LA and knowledge is "the measurement, collection, analysis, and reporting of data about learners and their attitudes with the aim of understanding and optimizing learning and the context in which it occurs". For example, LA can automatically analyze large volumes of student log data in an online environment to identify student learning behaviors and trends, allowing teachers to provide adaptive learning materials or experiences accordingly [13], [17], [24]. LA can also be used to assess learning progress and develop predictive models by evaluating and correlating large and complex datasets [4], [25], [26], [27], [28]. Data visualization, learning suggestions, prompts, predictions, relationship mining, and self-assessment are just some of the unique capabilities of LA [2], [29]. These attributes provide exciting new perspectives on how students learn. LA provides educators with a variety of skills, including identifying problem students, simulating healthy learning practices, and ultimately encouraging student achievement [11], [30].

Another potential benefit of LA is personalized feedback. It allows trainers to provide tailored guidance and advice [30]. More importantly, it enables the development and implementation of adaptive and personalized learning based on the different needs of students and the progression of dynamic processes [2], [31]. Students can also use LA to monitor their progress toward personal learning goals [31]. In addition, LA can help learners identify their strengths and weaknesses, thereby gradually guiding them to become autonomous learners. Although the currently available studies are insufficient to assess how LA affects learning [18], it facilitates both students and teachers. In addition, scholars have recognized some of the difficulties in implementing LA, including linking LA outcomes to learning science, optimizing different learning environments, and LA's ethics and privacy concerns.

B. MACHINE LEARNING AND DEEP LEARNING MODELS

The development of learning analytic was first proposed and then driven by research that seemed to reveal an encouraging intuition about the conventional behavior of students in the teaching process. In addition, it provides a basis for improvement for a wide range of investors, from interns and educators to utility providers, designers, and managers [32]. LA to improve performance prediction is the kind of actionable intelligence that instructors and students need to drive learning, and it inevitably requires interpretation and contextualization of data [25]. The use of computer-based methods to predict student performance is very common. To lay the groundwork for this research, here are some working examples using these strategies.

More recently, however, the focus on predicting student performance has been using their cognitive abilities, activity records in learning management systems (LMS), and student demographic attributes. Various models are used to identify at-risk students at an early stage. Since it is only a binary classification, the model must determine whether the student under consideration is at risk. For early prediction, a variety of machine learning classification model approaches are applied based on cross-sectional data attributes [33], [34]. Such as Bayesian classifiers, decision tree (DT) algorithm, artificial neural network (ANN), logistic regression (LR), k-nearest neighbor (kNN), support vector machine (SVM), extreme gradient boosting (XGB), adaptive boosting, random forest (RF) [35], [36], [37], [38], [39]. Semi-supervised learning is also seen in the early identification of at-risk individuals [40]. Interpretable classification rule mining algorithms [41], genetic programming and evolutionary algorithms [26], multi-view learning [42], [43], multi-objective optimization [44], ensemble models [45], and deep learning models [46], [47] are examples of developed algorithmic techniques. However, the answer to this challenge is not focused on a single model due to the huge disparity in educational data. It is usually determined by various parameters such as data size, data type, and pedagogy.

Many research works utilize machine learning models to predict students' academic performance using a variety of features. For example, [48] predicts student performance based on their previous performance in particular courses. The authors utilize several different machine learning models including NB, ID3, C4.5, and SVM. Models are evaluated using accuracy, precision, error rate, etc. Similarly, several machine learning models are analyzed for their efficacy in educational data mining to predict student grades in [49]. The authors utilized an optimum number of attributes to gauge the performance of models. Regression and classification approaches are used for predicting marks, and grades, respectively. The results approve the use of machine learning models for predicting students' performance. Results show that genetic algorithm-based DT and regression show better results.

C. ENSEMBLE ALGORITHMS

Using an ensemble algorithm improves the prediction accuracy and stability of a single learning algorithm [50]. A meta-algorithm that combines similar or different types of independently trained models to provide the final prediction serves as the basis for an ensemble learning model [51]. In a distance learning setting, [52] attempted to bridge the gap between empirical predictions of student performance and existing machine learning approaches. Academics have recently used several ensemble models to predict student performance. They propose an online voting-based classifier ensemble that integrates incremental versions of the NB, 1-NN, and Winnow algorithms. The authors concluded that the proposed method is the best option for building software tools. A method to choose an appropriate ensemble learner from a collection of different machine learning algorithms was proposed by [53] and is based on the Gini index and *p*-value. Experimental findings show that ensemble models attain good accuracy with a low false-positive rate (FPR).

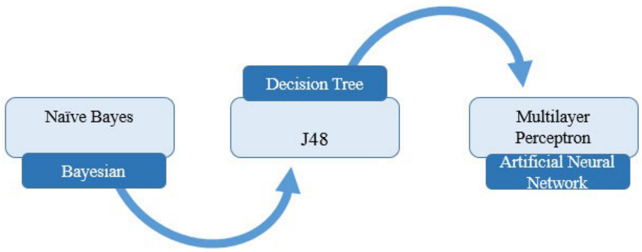


FIGURE 1. **Single base classifier.**

Furthermore, identifying students at risk of attrition has been difficult due to the poor efficiency and accuracy of prediction models. This poor performance has led to the suggestion of using an ensemble algorithm as a consequence of the poor performance being attributed to both the exclusive use of the base classifier and insufficient usage of variables [54], [55].

The study [56] focuses on using an ensemble model to predict student academic performance. The authors use performance factor analysis using RF, AdaBoost, and XGBoost mode to enhance prediction accuracy. Experiments involve using three different datasets indicating an enhanced performance compared to the original PFA algorithm.

While predominantly, existing works focus on machine learning ensemble learning, the study [57] proposes a student academic performance predicting system using an ensemble of machine learning and deep learning models. The proposed approach utilizes long short-term memory (LSTM), Rf, and GB. For experiments, the OULAD dataset and a self-formulated dataset are used. Performance comparison with other deep learning and existing models indicates the superior performance of the proposed approach with 96% accuracy.

III. MACHINE LEARNING MODELS

To build the prediction models, various machine learning models from various families can be applied. In the theoretical procedure used to create the model, these families differ. To answer the proposed research questions, in this study, classifiers are applied from two different categories, i.e., statistical classifiers (classifiers with a single-based classifier) and ensemble-based classifiers. The hyperparameters are individually adjusted for three base classifiers with various traits and decision boundaries on the extracted dataset to optimize their performance. Figure 1 shows single base classifiers used in this study.

A. BAYESIAN-BASED CLASSIFIERS

This is a classifier-building family based on probability theory. The classes of the probability of given features are mirrored in the classifier development process by employing rule-based network models in this genre. NB and Bayes Net employ the prior probabilities of various data given the hypothesis in the creation process. This family bases its classifier construction on the probability theory. This family constructs a classifier based on rule-based or network models that represent the probability of classes given particular features. The purpose of choosing these particular classifiers is to assess various classification philosophies, such as Naive Bayes, which was chosen to assess probability-based classification for the independence of the characteristics. Bayes theorem [58] is the foundation of NB work

$$P\left(\frac{c}{x}\right) = \frac{P\left(\frac{x}{c}\right) \times P(c)}{P(x)} \tag{1}$$

where $P\left(\frac{c}{x}\right)$ is the posterior probability of class (c, target) given predictor (x, attributes), $P(c)$ is the prior probability of class, $P(\frac{x}{c})$ is the likelihood which is the probability of predictor given class. $P(x)$ is the prior probability of the predictor. C is the target (class variable, whereas x is the features of the dataset, represented as

$$x = (x_1, x_2, x_3, x_4, \cdots, x_n) \tag{2}$$

where $x_1, x_2, x_3, x_4, \cdots, x_n$ represent the different features that will be mapped to the respective target class, as given

$$P\left(\frac{c}{x_1, \cdots x_n}\right) = \frac{P\left(\frac{x_1}{c}\right)P\left(\frac{x_2}{c}\right), \ldots\ldots P\left(x_1\right)P\left(c\right)}{P\left(x_1\right), P\left(x_2\right)\ldots\ldots P\left(x_n\right)} \tag{3}$$

Here each feature is just substituted for the target class. For all entries in the dataset, the denominator does not change, it remains static. Therefore, the denominator can be removed and proportionality can be injected as

$$P\left(\frac{c}{x_1, \cdots x_n}\right) \propto P\left(c\right)\prod_{i=1}^{n} P\left(\frac{x_i}{c}\right) \tag{4}$$

B. TREE-BASED CLASSIFIERS

This family creates classifiers in the shape of trees, with the exception that the leaf nodes serve as labels and the arcs (or edges) serve as the values of each attribute. The characteristic is chosen using a range of strategies that are based on entropy data that indicates the significance of an attribute with very little disorder. Many ways set themselves apart from one another when it comes to selection methods including $J48$ decision trees [59], LMT, and RF. A collection of samples that have already been classified serves as the training data for $J48$ is given as

$$S = s_1, s_2, s_3, \cdots, s_n \tag{5}$$

Each sample s_i consists of a p-dimensional vector and is given as

$$V = x_1, i, x_2, i, x_3, i, \cdots, x_p, i \tag{6}$$

where the associated sample's class, together with its attribute values or characteristics, are represented by the x_j.

The attribute with the most information is the one to divide to achieve the maximum classification accuracy. If a dataset consists of a single class label, then the dataset is pure or homogeneous, and if the dataset is multiclass, then the dataset is impure or heterogeneous. To measure the impurity of the dataset, three well-known indices, entropy measure, classification error, and Gini index are used [60] as

$$Entropy = \sum_{i=1}^{n} p_i \times log(p_i) \tag{7}$$

Since there is just one class in a pure table and probability is 1 ($log(1) = 0$), there is no entropy in a pure table.

Entropy achieves its peak value when all classes in the table exhibit the same probability. Entropy is the quantity of data required to completely characterize a sample. Therefore, if the sample is homogeneous-that is, if all the elements are similar-then the entropy is 0, and if the sample is evenly divided, then the entropy is at its highest. The information gain metric is applied to select a suitable attribute for each node. Gain (S, A) of an attribute (A) with a set of examples (S) is the information gain, as given here

$$Gain(S, A) = Entropy(S) - \sum_{\upsilon \in Values(A)} \frac{|S_\upsilon|}{|S|} Entropy(S_\upsilon) \tag{8}$$

where values (A) is the A's set of all possible values and S_v has values of the subset of S. In Equation 8, the first term defines the entropy of the S and the second term in the equation is the expected value of the entropy. This expected entropy is the sum of the entropy of each subset of S, weighted by the fraction of examples $\frac{|S_\upsilon|}{|S|}$ and is given as

$$Split\ information(S, A) = \sum_{i=1}^{n} \frac{|S_\upsilon|}{S} log_2 \frac{|S_\upsilon|}{S} \tag{9}$$

$$Gain\ ratio(S, A) = \frac{Gain(S, A)}{Splitinformation(S, A)} \tag{10}$$

The suitable attribute for a specific node in the tree is determined by the information gain metric and is given in Equation 10. For each non-terminal node, the specified values will be repeated for dividing and choosing the new attribute [61].

C. FUNCTION-BASED CLASSIFIERS

This family of classifiers aims to provide a model of the function. Both input characteristics and output labels are part of the function. There are various ways to translate inputs into outputs, including neural networks that update themselves via feed-forward and back-propagation methods, which improve prediction by lowering loss function error. Due to their tendency to provide stochastic solutions, MLP is useful in research because it frequently enables one to obtain estimated solutions for incredibly difficult issues like fitness approximation. The MLP is an improved feed-forward neural network that has three different types of layers: input, output, and hidden [62].

The back-propagation learning algorithm is used to train the neurons in the MLP. MLPs are intended to approximate any continuous function and to address issues that cannot be solved linearly. Pattern classification, recognition, prediction, and approximation are the most common applications of MLP. The computations performed by each neuron in the output and hidden layers [63] are given as

$$ox = Gb2 + W2h \tag{11}$$

$$hx = \Phi x = sb1 + W1x \tag{12}$$

With bias vector $b(1)$, $b(2)$, weight matrices $W(1)$, $W(2)$ and activation function of G and s, the set of parameters to learn is the set as given as

$$\theta = W(1), b(1), W(2), b(2) \tag{13}$$

A typical choice for s includes $\tan h(a)$ function or logistic sigmoid function.

TABLE 1. Weightage for data variable.

Exam Category	Marks Range	Weights
Assignment	10	10 %
Quiz	5	5 %
Presentation	10	10 %
Mid Term	25	25 %
Final Subjective	25	25 %
Final Objective	25	25 %
Total	100	100 %

D. CLASSIFICATION WITH MULTIPLE CLASSIFIERS (ENSEMBLE CLASSIFICATION)

Contrary to single classifiers, the purpose of selecting ensemble classifiers is to evaluate the various flavors of ensemble approaches based on the aforementioned classifiers to evaluate them with the different types of data sets. The key objective of the ensemble techniques is to reduce bias and variance. Some of the important ensemble approaches are stacking, bagging, boosting, and voting. In boosting-based methods, different weak learners are combined into strong learners to minimize the training loss by reducing bias and variance. A random sample of data is chosen, fitted with a model, and then trained sequentially, each model attempts to compensate for the shortcomings of its predecessor. Each iteration combines the weak rules from each separate classifier to generate a single, strong prediction rule. Bagging-based algorithms are expressed as

$$f_{bag} = f_1(X) + f_2(X) + \cdots\cdots\cdots\cdots + b(X) \tag{14}$$

The bagged prediction is on the left, and the individual learners are on the right side of Equation 15. More precisely,

$$f_0(x) = xargmin_{\hat{\gamma}} \times argmin\hat{\gamma} \times L(y, \hat{\gamma}) \tag{15}$$

where $f(x)$ is the model, y is the actual value, gamma is the predicted value and L is the loss function. First, model $f_0(x)$ is built as shown in Equation 15, on (x_i, y_i). The second model will be calculated as given

$$f_1(x) = f_0(x) + h_1(x) \tag{16}$$

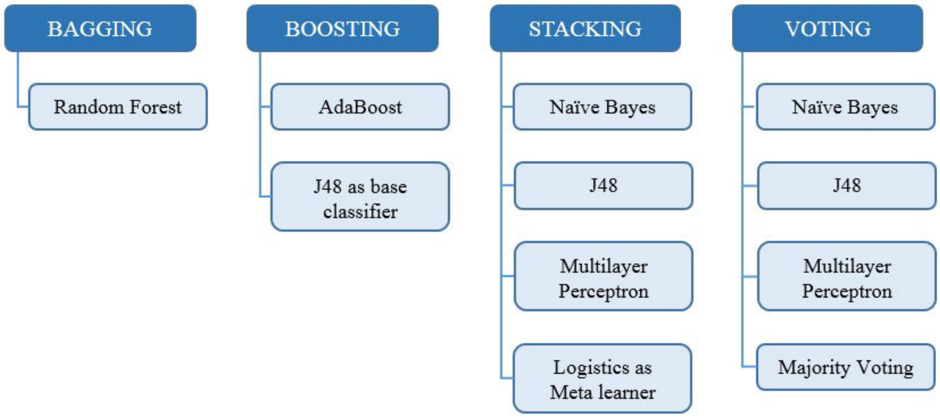


FIGURE 2. Ensemble classifiers used in this study.

TABLE 2. Data set for research question 1.

No.	Assignment	Quiz	Presentation	Mid Term	Subjective	Objective	Previous Academics	Enrollment
1	85.68	84.55	87.36	83.27	77.35	78.36	80.51	PROMOTED
2	81.82	73.64	80.73	69.64	59.59	65.9	77.21	PROMOTED
3	60	58.75	68.75	62	55.14	61.78	79.44	PROMOTED
4	80	65	85	78	37.78	49.09	78.26	DROPPED
5	66.11	81.67	64.17	52.89	41.75	53.2	68.05	DROPPED
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

where the residual ($ri = (y - \gamma)$) is calculated and the second model $h_1(x)$ is built on (x_i, y_i). Add $h_1(x)$ to $f_0(x)$ and get a new improved model $f_1(x)$. The above process will be repeated again and again until a generalized model, as given

$$f_m(x) = f_{(m-1)}(x) + h_m(x) \tag{17}$$

This study employs a variety of learning approaches drawn from several families. In terms of family perspective, the classifiers are summarized here. Some renowned ensemble classifiers used in this study are highlighted in Figure 2. By integrating supervised learning and ensemble learning, the objective is to build a classifier with strong classification capabilities.

On the collected dataset, six base classifiers are selected with varying characteristics and decision bounds and modify the hyperparameters independently. The ensemble technique used in this study includes stacking, bagging, boosting, and voting to combine the strengths of different heterogeneous classifiers to empirically demonstrate the higher prediction.

IV. MATERIALS AND METHODS

A. DATA AGGREGATION AND PREPARATION

This study was performed at a Pakistani provincial university. This study uses quantitative primary data from the University Information System to collect all essential variables for analysis, understand the relationship between variables in each dataset, and present the characteristics of the dataset.

The data was drawn from student batches enrolled in a bachelor of science (BS) (four years) and Master's (two-year) programs in the 2018-2019, 2019-2020, and 2020-2021 academic years. The selection of departments for data extraction is based on a higher order of the percentage of dropouts for RQ1. If the student's $CGPA \leq 1.5$, the student has been removed from the university roll. The data of the same instructor ad program is selected to answer RQ2. Moreover, to answer RQ3 and RQ4, data from sub-campuses of the same programs are used.

A total of 3130 student records were collected. The first two academic years were used to create the model and the last year to assess the performance. More generally, the idea is to use historical data from the past two years to predict the new academic year. We rely on a limited set of academic indicators of past performance. First, consider the course assignments, tests, presentations, projects, midterms, and final subjective and objective marks. Table 1 shows the grade distribution and weight percentages. Second, some information about their past academics, from matriculation to their last academic activity is available. Few factors are taken into account, not only voluntarily, but because of context. This study aims to identify students with potential difficulties as early as possible, that is before the academic year begins. Therefore, this limits the factors that can be collected. It also avoids some private information, such as demographics, finances, and distance.

To measure why the majority of students under-performed in a given course, we collected data on the same variables, namely the campus where the student is located and the job status of the faculty teaching the relevant course, as shown in

TABLE 3. Data set for research question 2.

No.	Assignment	Quiz	Presentation	Mid Term	Subjective	Objective	Enrollment
1	70	90	90	60	68	80	HIGHGRADE
2	30	60	100	56	30	60	LOW GRADE
3	30	60	100	60	55.2	72	HIGHGRADE
4	30	80	100	52	60	72	HIGHGRADE
5	30	60	100	60	45.2	60	LOW GRADE
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
3130	70	100	70	66	27.6	60	LOW GRADE

TABLE 4. Data set for research question 3.

No.	Assignment	Quiz	Presentation	Mid Term	Subjective	Objective	Enrollment
1	70	100	80	84	57.5	48	SUB-CAMPUS
2	80	70	90	76	64	92	SUB-CAMPUS
3	90	80	90	64	60	48	SUB-CAMPUS
4	75	30	80	48	76	84	MAIN-CAMPUS
5	75	20	70	60	50	76	MAIN-CAMPUS
6	90	80	80	78	65	100	MAIN-CAMPUS
.	.	.	.	.	.	.	.
18364	80	80	75	44	78	68	MAIN-CAMPUS

TABLE 5. Data set for research question 4.

No.	Assignment	Quiz	Presentation	Mid Term	Employee Job Status
1	80	60	80	72	PERMANENT
2	85	85	85	84	PERMANENT
3	70	80	80	70	PERMANENT
4	80	80	80	88	CONTRACT / VISITING
5	70	60	70	80	CONTRACT / VISITING
6	80	80	75	84	CONTRACT / VISITING
.	.	.	.	.	.
2000	75	80	80	82	CONTRACT / VISITING

Table 2. The courses taught are divided into general courses, elective courses, compulsory courses, and core courses. Each course is graded A-F. Here, grades C, D, and F are considered to be ''Low Grades'' when marks tend to be below 60, and ''High Grades'' otherwise. Table 3 illustrates marks distribution and percentage weights.

The courses are examined and ordered by the percentage of lower grades (C, D, and F). i. e. how many students are registered in a course and what is the number of students securing low grades, that is how the courses with the percentage of low grades are obtained. The data of courses having higher percentages of low grades are selected for analysis, and for the selected courses, the data of all students is collected. Table 4 shows the data collected to answer RQ3 and it contains information regarding marks distribution for assignments, quizzes, presentations, mid-term, etc.

For selected departments, data are recorded on all students who were dropped and promoted. In addition, the data of the departments with higher dropout percentage rates were selected. Students with missing grades in any of the exam categories of any course have been cleaned up as this can sometimes lead to biased decisions. Within each grade level category, each student's marks are aggregated by category and converted to percentages. This is done for all variables. In the case of the variable ''Previous Academics'', all marks for each student from matriculation to the last academic activity are recorded and converted to percentages. At the time of data collection, there were six sub-campuses of the university under study.

Table 5 shows the data samples from the data recorded for departmental selection of students studying in the same department as ''MAIN CAMPUSES'' and ''SUB CAMPUSES''. To normalize the data, each student's marks are converted to percentages.

B. RATIONALE FOR SELECTING ENSEMBLE MODELS

Ensemble learning is a machine learning strategy that involves integrating the predictions of multiple individual models to enhance predictive accuracy. By combining diverse models, ensemble learning aims to minimize the limitations of individual models, leading to improved generalization and performance in various machine learning tasks [64]. Using an ensemble algorithm improves the prediction accuracy and stability of a single learning algorithm [50]. A meta-algorithm that combines similar or different types of independently trained models to provide the final prediction serves as the basis for an ensemble learning model [51].

The key advantage of ensemble learning is its ability to reduce overfitting [64], enhance predictive accuracy [65], and improve model robustness [50], handling complex relationships [66], flexibility [67], making it a valuable technique in machine learning. Ensemble methods, such as bagging and boosting, are commonly employed to address these challenges [64]. Furthermore, ensemble learning can effectively handle complex relationships in data and is less sensitive to noise, making it suitable for real-world, noisy datasets [66]. Here are some key ensemble learning methods and references:

- **Bagging (Bootstrap Aggregating)**: Bagging involves training multiple instances of the same base model on different subsets of the training data. The predictions from these models are then aggregated, often using techniques like majority voting for classification or averaging for regression [64].
- **Boosting**: Boosting is a technique where base models are trained sequentially, with each subsequent model focusing on correcting the errors made by the previous models. Popular boosting algorithms include AdaBoost, Gradient Boosting, and XGBoost [68].
- **Stacking**: Stacking involves training multiple diverse base models and then training a meta-model (often a simple model like linear regression) on the predictions made by the base models. The meta-model learns how to best combine the base model predictions to make the final prediction [69].
- **Voting**: Voting ensembles combine multiple independently trained models by taking a majority vote (for classification) or an average (for regression) to make the final prediction. Voting can be hard (using majority voting) or soft (considering class probabilities) [70].

C. PERFORMANCE EVALUATIONS

To evaluate the performance of all algorithms, 10-fold cross-validation and 10 different runs for each partition are used. The preferred and widely used metrics for measuring the model's classification performance are given here. These metrics make use of the confusion matrix's true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values.

Accuracy determines how frequently the model's predictions come true and is calculated using

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FN + FP)} \tag{18}$$

The precision determines how often the model correctly predicts a positive outcome and is calculated using

$$Precision = \frac{(TP)}{(TP + FP)} \tag{19}$$

Recall measures how often it correctly predicts when it is positive and is calculated using

$$Recall = \frac{(TP + TN)}{(TP + FN)} \tag{20}$$

TABLE 6. Dataset details for student performance.

Research Question 1			
Instances	**Attributes**	**Dropout**	**Promoted**
1201	8	236	995
Research Question 2			
Instances	**Attributes**	**High Grades**	**Low Grades**
3130	8	1931	1199
Research Question 3			
Instances	**Attributes**	**Main Campus**	**Sub Campuses**
18364	8	9178	1199
Research Question 4			
Instances	**Attributes**	**Regular**	**Visiting/Contractual**
2000	8	1001	999

F measure is the harmonic mean of precision and recall and is calculated using

$$F\ measure = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \tag{21}$$

Besides these metrics, this study also utilizes statistical indicators including mean absolute error (MAE) and root mean squared error (RMSE). MAE is calculated as

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{22}$$

where y_i is the prediction, x_i is the true value, and n is the total number of data points.

RMSE is calculated using the following equation

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{n}} \tag{23}$$

where i is the prediction, x_i is the actual observation time series, $\hat{x}_i$ is the estimated time series, and n is the number of non-missing data points.

V. RESULTS AND DISCUSSIONS

This is a fact that each predictive model's performance and accuracy are dependent on several factors, one of which is the dataset. Unreliable data can lead to incorrect results [71]. To comprehensively investigate and strengthen this research, various traditional prediction algorithms were executed, and four experiments were conducted against four research questions. Initially, a descriptive analysis of the datasets used in this study is performed.

A. DATA ANALYSIS

Data analysis is a computational process that establishes the relationships between machine learning methods and relevant information for decision-making in acquired large datasets. It looks specifically at a dataset from the perspective of the class labels. Details about characteristics and instances can be found in the model structure information.

The dataset used for RQ 1 is an imbalanced dataset which is very common in educational data. There are many more promoted students than dropped students, which makes it difficult for a classifier to forecast dropout rates. Table 6 shows details for the datasets used for each research question.

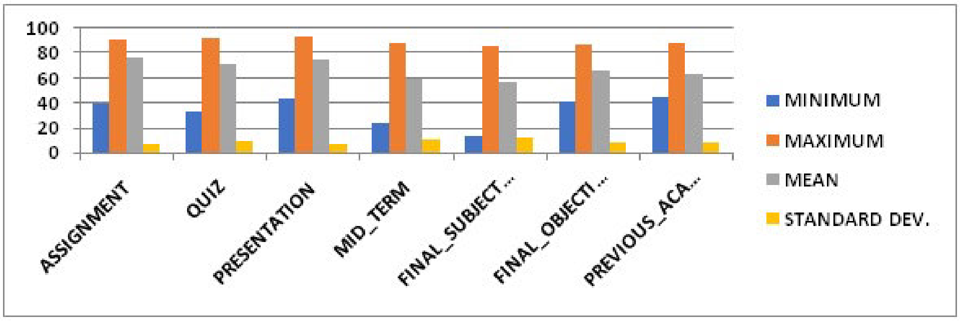


FIGURE 3. Basic statistics of the dataset (a) used for RQ1.

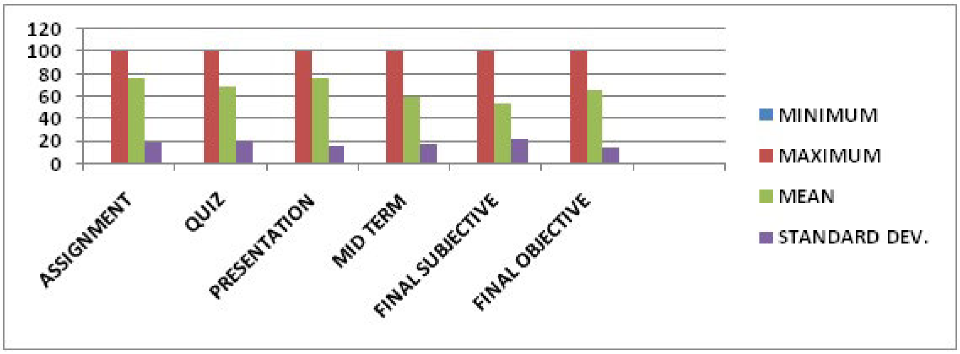


FIGURE 4. Basic statistics of the dataset (b) for RQ2.

TABLE 7. Basic statistics of dataset(a) for RQ1.

Exam Category	Min. Marks %	Max. Marks %	Mean	Std. Dev.
Assignment	40	90.83	76.425	6.166
Quiz	33.33	92	71.399	8.739
Presentation	43.68	93.33	75.306	6.541
Mid Term	23.35	87.82	59.53	10.39
Final Subjective	13.25	85.49	56.702	12.016
Final Objective	41.36	87.12	65.767	7.646
Previous Academics	44.88	88.05	63.558	7.318

Similarly, the dataset used for RQ 3 is also imbalanced with sub-campus instances of only 1199 compared to 9178 instances for the main campus.

Table 7 shows the basic statistics of the dataset used for RQ1. It shows the distribution of dataset instances in terms of minimum marks, maximum marks, mean, and standard deviation (Std. Dev.) for each category of examination.

A visual presentation of the exam category distribution is shown in Figure 3. It represents the basic statistics of the data set used for RQ 1.

Table 8 shows the mean marks achieved in each category for promoted and dropped students. These statistics are for the dataset used for RQ1.

Table 9 and Figure 4 represent the basic statistics of the dataset which is gathered and used to investigate and answer RQ 2.

TABLE 8. Mean of marks of the dataset (a).

Exam Category	Promoted	Dropped
Assignment	77.4229	72.3414
Quiz	73.1795	64.1045
Presentation	76.5942	70.047
MidTerm	62.4105	47.7569
Final Subjective	60.528	41.055
Final Objective	67.6329	58.1423
PreviousAcademics	64.5385	59.5509

TABLE 9. Basic statistics of dataset (b) used for RQ2.

Exam Category	Minimum	Maximum	Mean	Std. Dev.
Assignment	0	100	76.102	19.021
Quiz	0	100	68.787	21.04
Presentation	0	100	76.127	17.024
Mid_Term	0	100	60.554	18.569
Final_Subjective	0	100	54.073	22.237
Final_Objective	0	100	65.251	14.398

The mean of each attribute (class label-wise) for dataset (b) which is used for RQ2 is given in Table 10. It shows the mean marks achieved in each category for high-grade and low-grade students.

Table 11 and Figure 5 represent the basic statistics of the data set used for RQ 3.

Table 12 shows the mean marks achieved in each category for sub-campuses and the Main campus.

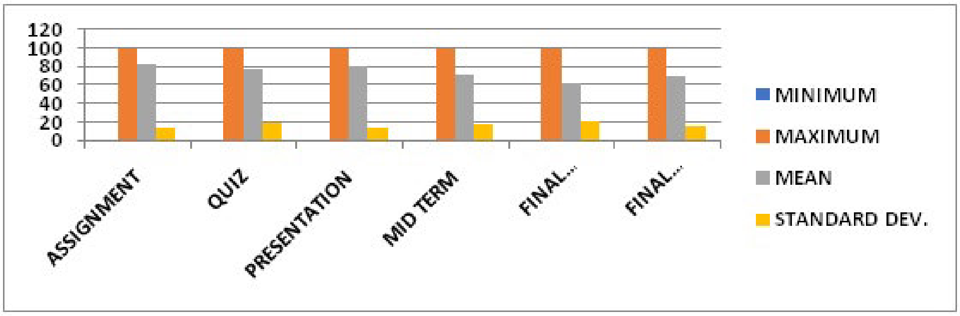


FIGURE 5. Basic statistics of dataset (c) for RQ 3.

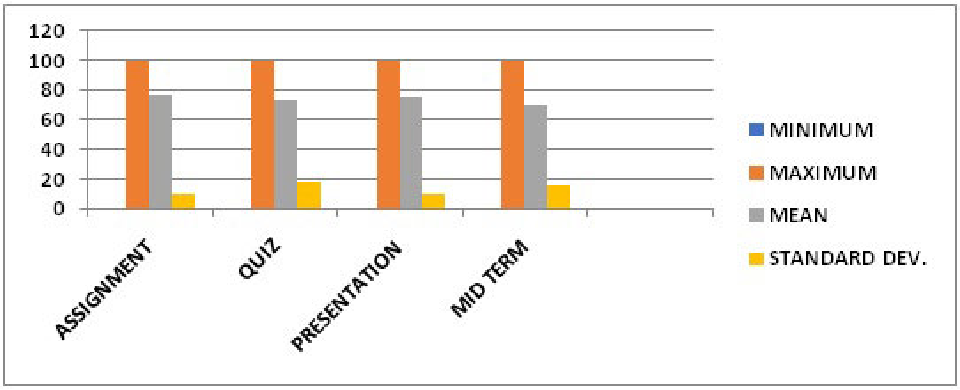


FIGURE 6. Basic statistics of dataset (d) used for RQ 4.

TABLE 10. Mean of marks of dataset (b).

Exam Category	High Grade	Low Grade
Assignment	81.4414	67.5114
Quiz	74.486	59.6167
Presentation	82.0291	66.7432
Mid_Term	69.8718	45.574
Final_Subjective	65.783	35.3155
Final_Objective	70.9326	56.0114

TABLE 11. Basic statistics of dataset (c) used for RQ 3.

Category	Minimum	Maximum	Mean	Std. Dev.
Assignment	0	100	81.871	13.847
Quiz	0	100	77.309	19.091
Presentation	0	100	81.149	13.118
Mid Term	0	100	70.971	17.07
Final Subjective	0	100	61.691	20.949
Final Objective	0	100	69.986	15.873

TABLE 12. Mean of marks of dataset (c).

Category	Sub Campuses	Main Campus
Assignment	85.5299	78.4417
Quiz	80.6631	73.966
Presentation	82.948	78.9539
Mid Term	74.1849	67.8352
Final Subjective	60.9376	62.4413
Final Objective	66.6987	73.3158

Table 13 and Figure 6 represent the basic statistics of the data set used for research question 4.

Table 14 shows the mean marks achieved in each category for permanent and contract teachers.

TABLE 13. Basic statistics of dataset (d)) for RQ 4.

Category	Minimum	Maximum	Mean	Std. Dev.
Assignment	0	100	76.556	919
Quiz	0	100	72.603	17.769
Presentation	0	100	75.175	9.435
Mid Term	0	100	69.552	16.197

TABLE 14. Mean of marks of dataset (d).

Category	Permanent	Contract / Visiting
Assignment	74.7789	76.1632
Quiz	72.4281	73.3063
Presentation	75.0647	74.9324
Mid Term	65.3972	73.5764

B. ANALYSIS OF RESULTS FROM PREDICTION MODELS

Several machine learning and ensemble techniques are applied to each data set following the research questions to illustrate the performance of the proposed system. The output of different classification methods is analyzed and carefully examined, having a study of core measures. The selected supervised classifiers are NB, J48, and MLP. In the ensemble category, RF is used as the base classifier for bagging, ADA and J48 for boosting, NB, J48, MLP, and LG for stacking, and NB, J48, MLP, and LG for voting. In the training phase, base learners are constructed using their default settings. Four ensemble models are compared with different base classifiers. The comparison is based on the four

TABLE 15. Class label-wise comparison of supervised machine learning algorithms and ensemble classifiers with the dataset (a).

Metrics		Algorithms						
		Individual Classifiers			Ensemble Classifiers			
		Naïve Bayes	J48	MLP	Bagging	Boosting	Stacking	Voting
Precision	Promoted	92	93	95.2	94	95	95.3	96
	Dropped	70.6	81.8	85.1	86	83.1	90	83.1
	Weighted Avg.	81.3	87.4	90.15	90	89.05	**92.65**	89.55
Recall	Promoted	90.9	95.8	96.6	96.8	96.1	97.5	95.9
	Dropped	83.4	78	85	80.5	79.2	89.1	82
	Weighted Avg.	87.15	86.9	90.8	88.65	87.65	**93.3**	88.95
F Measure	Promoted	93.9	94.2	95.9	96	95.5	96.4	95.9
	Dropped	78.9	79.8	82.5	83.2	81.1	86.4	82.3
	Weighted Avg.	86.4	87	89.2	89.6	88.3	**91.4**	89.1
TP Rate	Promoted	90.9	95.8	96.6	96.8	96.1	97.5	95.9
	Dropped	89.4	78	80.1	80.5	79.2	83.5	80.5
	Weighted Avg.	90.15	86.9	88.35	88.65	87.65	**90.5**	88.2
FP Rate	Promoted	10.6	22	19.9	19.5	20.8	19.5	16.5
	Dropped	9.1	4.2	3.4	3.2	3.9	2.5	4.1
	Weighted Avg.	91	13.1	11.65	11.35	12.35	11	**10.3**
Accuracy		90.6	92.3	93.3	93.6	92.8	95.2	93.4
MAE		0.109	0.1006	0.0767	0.1001	0.0712	0.0658	0.0712
RMSE		0.2644	0.2594	0.2307	0.218	0.2606	0.2148	0.2565

distinct datasets with each base classifier tuned for different classification tasks. Further results are represented as per the research questions for this study.

1) RQ 1: WHICH MACHINE LEARNING MODELS CAN SHOW BETTER PREDICTIONS OF WHETHER A STUDENT WILL COMPLETE HIS DEGREE OR WILL BE DROPPED FROM THE UNIVERSITY?

The precision, recall, F measure, TP rate, and FP rate test statistics of the learning algorithms are used in the study. The classification performance of each test and ensemble approach is evaluated for these metrics. All the classifiers are examined using the standard approach known as 10-fold cross-validation. Tables 15 show the results for dataset (a) used to answer RQ 1. The highest performance achievement is marked with bold style.

The aggregated results presented in Table 15 show that for predicting high and low-grade students, the stacking ensemble technique outperforms all other classifiers and across all metrics. The stacking ensembles method shows higher accuracy, high precision, high recall, high F measure, and lower classification error and RMSE than any other model. Stacking with an accuracy of 95%, precision, recall, and F measures lie in the range of 91% to 95%. The next best approach, closely followed, is voting. On the other hand, MLP proved best among individual machine learning algorithms.

2) RQ 2: WHAT ARE THE REASONS FOR THE POOR PERFORMANCE OF THE MAJORITY OF THE STUDENTS IN SOME COURSES?

The aggregated results presented in Table 16 show that for predicting promoted or dropped students, the stacking ensemble technique outperforms all other classifiers across all metrics. The stacking ensembles method shows higher accuracy, high precision, high recall, high F-measure, and lower classification error and RMSE than any other model. Stacking shows an accuracy of 95% which is the best among all models. Similarly, precision, recall, and F measure fall between 91% to 93%. The best performance from stacking is followed by voting. Among individual classifiers, MLP proved best with an average accuracy of 99%.

3) RQ 3: TO WHAT EXTENT DO THE STUDENTS STUDYING ON THE MAIN CAMPUS DIFFER FROM THE STUDENTS STUDYING IN THE SUB-CAMPUSES?

Table 17 shows the results for the performance of machine learning models for predicting the performance of students studying on different campuses. The stacking, bagging, and boosting ensemble methods show a marginal difference in predicting the evaluation of students between sub-campuses and main campuses. The accuracy, precision, recall, and F measure of all the ensemble methods lie in the range of 68% to 73%. The next best approaches, closely followed are NB and J48.

Figure 7 represents a comparison of weighted average values of accuracy, precision, recall, F measure, TP rate, and FP rate of all techniques used in all the datasets.

C. COMPUTATIONAL PERFORMANCE TESTING

The performance testing experiments involved the execution of the two learning paradigms, namely supervised machine learning algorithms and ensemble methods. To perform this experiment the four datasets comprising a different number of instances are used. This experiment proceeded with three machine learning classifiers and four ensemble methods with different variants as base classifiers. These experiments comprised simulation runs on a system with processor Intel(R) Core(TM) i3-2370M CPU @ 2.40 GHz, with 8.00 GB RAM and a 64-bit Windows operating system.

Figure 8 shows the comparison of the computational complexity of the models. The stacking with the highest

TABLE 16. Class label-wise comparison of supervised machine learning algorithms and ensemble classifiers for student's performance.

Metrics		Algorithms						
		Individual Classifiers			Ensemble Classifiers			
		Naïve Bayes	J48	MLP	Bagging	Boosting	Stacking	Voting
Precision	High Grade	95.7	94.5	98.2	97.8	96.9	99	98.3
	Low Grade	95	91.8	97.8	93.8	95.6	98.7	97
	Weighted Avg.	95.35	93.15	98.0	95.8	96.25	**98.1**	97.65
Recall	High Grade	95.7	94.5	99.2	97.9	96.9	99	98.3
	Low Grade	95.2	91.8	98.7	93.8	95.6	98.7	97.9
	Weighted Avg.	95.5	93.2	99.0	95.9	96.3	**98.9**	98.1
F Measure	High Grade	93.9	95.2	95.9	96	95.5	96.4	95.9
	Low Grade	78.9	79.8	82.5	83.2	81.1	84.4	83.3
	Weighted Avg.	86.4	87.5	89.2	89.6	88.3	**90.4**	89.6
TP Rate	High Grade	95.7	94.5	99.2	97.9	96.9	99	98.3
	Low Grade	95.2	91.8	97	93.8	95.6	98.7	97.9
	Weighted Avg.	95.45	93.15	98.1	95.1	96.25	**98.1**	98.1
FP Rate	High Grade	4.8	8.3	1.3	6.3	4.4	1.3	2.1
	Low Grade	4.3	5.5	0.8	2.1	3.1	1	1.7
	Weighted Avg.	4.55	6.9	1.05	4.2	3.75	**1.15**	1.9
Accuracy		95.5	93.5	99.0	96.3	96.4	98.9	98.2
MAE		0.0935	0.0763	0.0119	0.0746	0.0352	0.0172	0.0185
RMSE		0.192	0.2485	0.0975	0.1692	0.1767	0.0897	0.1361

accuracy rate also takes a longer period to develop a model from all of the algorithms utilized in this study. Voting comes in second place when employing ensemble methods. MLP is the third-placed algorithm overall. NB is the algorithm with the minimum run time in all experiments. The most time-effective models are NB, J48, and boosting.

D. DISCUSSION

The experiments conducted in this research allow us to compare the models' predictions using academic and demographic. This research examines various machine learning and ensemble architectures to discover the model capable of behavioral statistical analysis of learning analytics. This study yielded a host of promising outcomes. The following major observations are emphasized in particular research questions.

1) HOW CAN THE CHANCE OF STUDENT DROPOUT HAVING A RISK OF NOT DEGREE COMPLETION BE PREDICTED?

It can be observed that all students in the assignments, quizzes, and presentation exam categories have average grades in the range of over 75%, which indicates that students in these exam categories achieve higher grades, as shown in Table 16. Taking into account all three attributes, it can be concluded that students achieved good grades in their course grades. The standard deviation is also between 6 and 8, indicating that the data points are distributed within a normal range of values. In the midterm and final subjective exam categories, students are examined on subjective methods, with average scores ranging from 56-60% below the session marks. The standard deviation is between 10 and 12, indicating that the data points cover a wider range of values than the unit ranks. In the target category, the range of average scores is close to 66%, indicating that students are doing well when given multiple-choice to answer questions.

On the other hand, students' previous academic development is highlighted in the "Previous Academics" category, with an average range of 63%, which reflects the level at which all students are enrolled in certain courses. The standard deviation of these two attributes is about 7, indicating that the data points are distributed within a normal range of values. Very interesting observations and reports show that, as shown in Table 17, the range of student achievement on assignments and presentations reaches 70% of all students, whether they are "dropped" or "promoted" students. The same applies to "quizzes". It is in the range of the '60s and '70s. Therefore, it is assumed that students will receive a very healthy range of grades in the session grade category, regardless of how well they perform on other types of exams such as mid-term or final term exams.

It is also important to note that 'Dropped' students have scored in the range of less than 50% in the 'Mid Term' and 'Final Subjective' exam categories, indicating that students' subjective methods must be strong enough to perform better to promote in the next semester. In addition, the marks range in the "Promoted" category does not even reach 65%. On the other hand, it can be seen that students in the "Dropout" category also perform well, within the 58% range in the "Final Objective" exam category. This fact shows that even underperforming students do well on exams when options are offered. Finally, in the "Previous Academics" category, previous academic performance is not decisive because both category label categories have a range of 60%. After summarizing these facts, it is stated that students must enhance their subjective methods, such as in "Mid Term" and "Final Subjective" to receive good grades and be promoted to the upcoming semesters. In addition, the evaluation of different levels of students should be balanced.

2) TO WHAT EXTENT DOES THE EVALUATION OF THE STUDENTS STUDYING ON THE MAIN CAMPUS DIFFER FROM THE STUDENTS STUDYING ON THE AFFILIATED CAMPUSES?

It can be observed that the average grade range for all students in the assessment categories is above 68% for all

TABLE 17. Class label-wise comparison of supervised machine learning algorithms and ensemble classifiers for students' performance studying in different campuses.

Metrics		Algorithms						
		Individual Classifiers			Ensemble Classifiers			
		Naïve Bayes	J48	MLP	Bagging	Boosting	Stacking	Voting
Precision	Affiliated	63.7	68.3	68.2	68.7	68.3	70.8	67.5
	Uni-Campus	69.1	69.2	71.2	69.7	68.8	70	68.9
	Weighted Avg.	66.4	68.75	69.7	69.2	68.55	70.4*	68.2
Recall	Affiliated	74.3	70	64.3	70.6	69.3	71.8	70.1
	Uni-Campus	57.5	67.5	70.1	67.8	67.8	65.9	66.3
	Weighted Avg.	65.9	68.75	67.2	69.2*	68.55	68.1	68.2
F Measure	Affiliated	68.6	69.2	66.2	69.7	68.8	69.8	68.8
	Uni-Campus	62.8	68.4	68.1	68.8	68.3	67.9	67.5
	Weighted Avg.	65.7	68.8	67.15	69.25*	68.55	68.11	68.15
TP Rate	Affiliated	74.3	70	64.3	70.6	69.3	71.8	70.1
	Uni-Campus	57.5	65.5	70.1	67.8	67.8	65.9	66.3
	Weighted Avg.	65.9	67.75	67.2	69.2*	68.55	68.11	68.2
FP Rate	Affiliated	42.5	32.5	29.9	32.2	32.2	34.1	33.7
	Uni-Campus	25.7	30	35.7	29.4	30.7	28.2	29.9
	Weighted Avg.	34.1	31.25	32.8	30.8*	31.45	31.15	31.8
Accuracy		65.9	68.8	67.2	69.2*	68.6	69.0	68.2
MAE		0.413	0.3937	0.4083	0.3809	0.3588	0.4054	0.3183
RMSE		0.4652	0.4637	0.4563	0.4483	0.4664	0.4496	0.5642

FIGURE 7. Comparison of accuracy, precision, recall, F-measure, TP rate ad FP rate of all the datasets.

students in the assessment categories 'Assessment', 'Quiz', and 'Presentation', indicating a wider range of scores in these exam categories. Students, especially in assignments and presentations, scored over 75%, as shown in Table 9. Looking at all three of these qualities, it can be said that students achieve good results in session marks. The fact that the standard deviation is also between 17 and 21 indicates that the data points are spread over a wider range of values. Typical grades for assessing students' subjective methods in the mid-term and final subjective test categories range

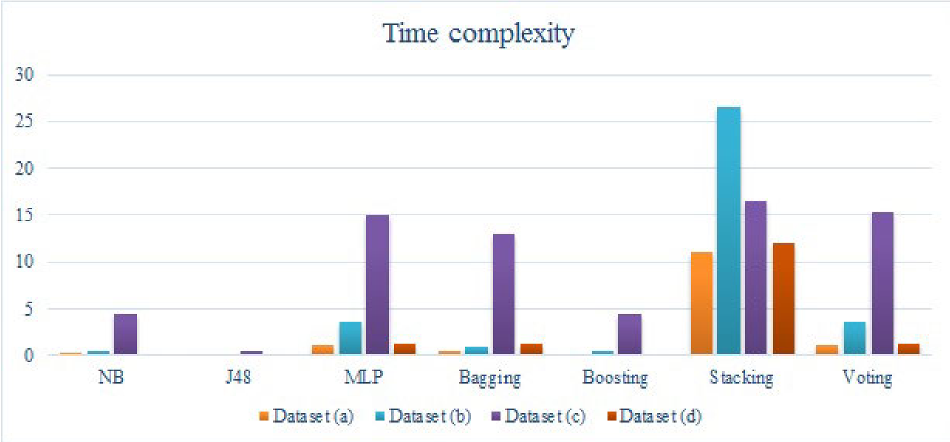


FIGURE 8. Results for the computational complexity of models.

from 54-60% lower than session marks. Data points are more widely distributed than session marks, ranging from 18 to 22 as indicated by the standard deviation. When questions in the target category have "objective", the average score ranges from 65%, indicating that students are performing well and have a good understanding of the objective-based questions.

As shown in Table 10, students' grades in "Assignment" and "Presentation" range from almost 67% of all students, regardless of whether they are "Low" or "High" grades, and for "Quiz", the range is from 60% to 70%. Therefore, regardless of how well students perform on other types of tests, for example, students appear to have a relatively healthy range of grades in the session assessment categories in the mid-term or final term exam. It is also worth noting that the range of grades for Low-Grade students is around 45% in the mid-term and around 35% in the Final Subjective, suggesting that the student's subjective approach must be good enough to achieve a good grade to achieve a "high grade" courses.

When the range of grades in the "High Grade" category is examined, it is discovered that it does not exceed 70%. On the other hand, in the "Final Objective" exam category, students from the "Low Grade" group are also performing well, with scores at the 56 percent level. This demonstrates that when options are presented, even weak performers perform well in exams. After summarizing these facts, it can be concluded that students must enhance their subjective methods, such as in "Mid Term" and "Final Subjective," to receive "High Grades" and be promoted to the next semester. In addition, the evaluation of different levels of students should be balanced.

3) HOW IS THE EVALUATION OF THE STUDENTS AFFECTED BY THE JOB STATUS OF THE TEACHERS? (VISITING/CONTRACTUAL/PERMANENT)?

The value ranges for different qualities are predicted in Table 11. Average student grades in the assignments, quizzes, and presentation exam categories ranged over 77% indicating that student achievement in these exam categories varied, even more, especially assignments and presentations, which accounted for over 80%. Taking into account all three attributes, it can be said that students achieved good grades in the session category. The standard deviation is also between 13 and 19, indicating that the data points are distributed within a normal range of values.

Compared to semester grades, average grades in the test categories "Midterm" and "Final Subjective", which assess students' subjective methods, range from 60% to 70% lower. In addition, the standard deviations for these study categories were lower between 17 and 20, indicating that the data points were distributed within a more normal range of values. When questions in the target category have options, the "Objective" category ranges up to 70%, indicating that students are performing well and have a good understanding of the objective questions.

The range of student grades for assignments, quizzes, and presentations on sub-campuses is greater than 80%, but below 80% on the university main campus, confirming that session grades on sub-campuses are higher than "Main Campus", which can lead to good grades. In terms of mid-term exams, it is worth mentioning that the score range of the sub-campuses is around 74%, which is relatively higher than that of the main campus, and the mid-term exam scores are slightly lower than 68%. On the other hand, students from the main campus performed better than students from sub-campuses in the exam categories of "Final Subjective" and "Final Objective".

In conclusion, it can be said that students from sub-campuses outperformed the main campus in the 'session' and 'midterm' categories and that those grades were exclusively from 'sub-campuses'. The affiliated campus system does not include an evaluation of any part of the main campus of the University. On the other hand, the university has some checks and balances in the evaluation of "Finals", so students from "main-campus" performed relatively better.

TABLE 18. Winning algorithms in each dataset for a single classifier.

Classifier	Dataset (a)	Dataset (b)	Dataset (c)	Dataset (d)
NB				
J48			⊙	⊙
MLP	⊙	⊙		

4) WHAT ARE THE REASONS FOR THE POOR PERFORMANCE OF THE MAJORITY OF THE STUDENTS IN SOME SPECIFIC COURSES?

As shown in Table 13, there was no significant difference in the scores given to students by either "faculty" or "contract" teachers in "Assignments", "Quiz" and "Presentation". Summarizing these facts in the range of 73%, it can be seen that in the "Course Grades" category teachers are free to set grading levels for all students. The "contract teachers" give more grades to students compared to "permanent" instructors, but the difference is not significant, suggesting that there is no concern in this regard.

E. WINNING ALGORITHMS IN EACH DATASET

One of the most pressing issues confronting machine learning researchers nowadays is "whether a combined classifier model outperforms the best of the base level classifiers". This vital question is attempted to be addressed in this study.

It has been found that MLP works best as a single classifier when there are more attributes and a smaller dataset. Additionally, it has been noted that MLP is the best algorithm in this study when standard deviations are high. The decision tree approach is useful with larger datasets to improve testing performance. However, it must be kept in mind that a large data collection, especially one with numerous attributes, will generate a massive decision tree, resulting in a longer computation time. Table 18 shows the winning single classifiers for each dataset used in this study.

Table 19 shows the comparison of ensemble classifiers used in this study, indicating which classifier performed better for which dataset. Stacking performs optimally in smaller datasets for ensemble classification, but bagging performs better in larger datasets. The boosting strategy, which is likewise based on the decision tree in this study, works well for fewer attributes. Another observation is that the performance of the ensemble classifiers suffers from the increased rate of standard deviation. In such cases, however, MLP as a single classifier performs far better.

The fundamental conclusion of this study is that ensemble classification approaches outperform base-level classifiers. The idea of classifier merging is presented as a novel path for improving the performance of single classifiers. These classifiers might be developed using a variety of methods of classification and could provide varying percentages of correctly categorized things. Algorithms for combining categorization results are designed to create more accurate, precise, and right results for the system.

TABLE 19. Winning algorithms in each dataset for ensemble classifier.

Classifier	Dataset (a)	Dataset (b)	Dataset (c)	Dataset (d)
Bagging			⊙	⊙
Boosting				⊙
Stacking	⊙	⊙		
Voting				

F. FUTURE RESEARCH DIRECTIONS

Using machine learning techniques to predict students' performance in higher education is a promising field with several potential directions for the future. The following are some topics of interest and possible lines of study for more investigation and development:

- **Personalized Learning Pathways**: Personalised education is becoming more popular. By determining each student's strengths, shortcomings, and preferred methods of learning, machine learning models can be used to customize learning paths for them. For each student, predictive models can be very helpful in suggesting courses, resources, and instructional techniques.
- **Early Intervention and Support**: developing models that can detect underachieving students early in the semester or academic year. This makes it possible to provide struggling students with prompt interventions like counseling, tutoring, or extra resources.
- **Multi-modal Data Analysis**: integrating information from multiple sources, such as wearable technology, social media, and academic records, to create more complete models for performance prediction. This could offer a comprehensive picture of the life and conduct of a student.
- **Explainable AI for Education**: developing interpretable models to explain their predictions so that teachers and students can better understand the reasoning behind a given prediction. Trust and decision-making may both benefit from this.
- **Global and Cross-Cultural Applications**: extending research to take cross-cultural and global aspects of higher education into account. Diverse educational systems and cultural differences may need to be taken into consideration by predictive models.

VI. CONCLUSION

Predominantly, existing studies on student academic performance prediction utilize machine and deep learning models, but ensemble models are not very well investigated. This work has briefly presented the research on the application of base and ensemble classifiers to teacher assessment patterns and student academic performance prediction by examining the types of datasets unique to academia. In this study, data from a learning management system of a university was used, which included data from multiple campuses of a university. The aim is to provide a feasible solution to synthesize more accurate predictive models. Eight representative standard machine learning algorithms from different families were used for behavioral statistical analysis.

This work has also been significantly expanded to include the use of base classifiers and ensemble classifiers. The accuracy and potency of ensemble classifiers and individual classifiers of various kinds are also actually examined and compared in relation to several research inquiries. Overall, this effort advances knowledge of the potential applications of learner data utilized to forecast student achievement through ensemble techniques and adequately responds to the main queries: which combination factors are the most trustworthy indicators of student academic achievement? What prospects exist for the trustworthy use of voting ensemble, bagging, boosting, and stacking strategies? On the other hand, to find the ideal ensemble system, bagging, boosting, staking, and voting methods are trained. Using these methods, at-risk students are assessed and reasons for under-performance; student performance on a variety of assessment methods; and assessment patterns for teachers of different professional statuses across multiple campuses. Overall, the stacking ensemble method is the best overall and can be used to improve the performance prediction model, thereby increasing the accuracy, reducing the error rate, and improving the prediction efficiency. The methods presented in this study will help teachers, educators, and administrators develop new higher education regulations and instructional programs. At-risk students can get the support and feedback they need to avoid falling behind or failing with the help of these policies and instructional interventions, which will be decided upon. This fact has the potential to increase higher education's quality, efficiency, and effectiveness. Because it gives a high-precision prediction, the approach of this study will assist administrators, educators, and policymakers in developing new guidelines and pedagogical approaches in higher education.

ACKNOWLEDGMENT

The authors extend their appreciation to King Saud University for funding this research through Researchers Supporting Project Number (RSPD2023R890), King Saud University, Riyadh, Saudi Arabia.

REFERENCES

[1] J. Zilvinskis, J. Willis III, and V. M. Borden, "An overview of learning analytics," *New Directions Higher Educ.*, vol. 2017, no. 179, pp. 9–17, 2017.

[2] C. Schumacher and D. Ifenthaler, "Features students really expect from learning analytics," *Comput. Hum. Behav.*, vol. 78, pp. 397–407, Jan. 2018.

[3] K. M. L. Jones, "Learning analytics and higher education: A proposed model for establishing informed consent mechanisms to promote student privacy and autonomy," *Int. J. Educ. Technol. Higher Educ.*, vol. 16, no. 1, Dec. 2019.

[4] J.-L. Hung and K. Zhang, "Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching," *MERLOT J. Online Learn. Teach.*, vol. 4, no. 4, pp. 426–437, Dec. 2008.

[5] G. Sedrakyan, J. Malmberg, K. Verbert, S. Järvelä, and P. A. Kirschner, "Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation," *Comput. Hum. Behav.*, vol. 107, Jun. 2020, Art. no. 105512.

[6] F. G. K. Yilmaz and R. Yilmaz, "Student opinions about personalized recommendation and feedback based on learning analytics," *Technol., Knowl. Learn.*, vol. 25, no. 4, pp. 753–768, Dec. 2020.

[7] O. H. T. Lu, J. C. H. Huang, A. Y. Q. Huang, and S. J. H. Yang, "Applying learning analytics for improving students engagement and learning outcomes in an MOOCs enabled collaborative programming course," *Interact. Learn. Environ.*, vol. 25, no. 2, pp. 220–234, Feb. 2017.

[8] A. Pardo, J. Jovanovic, S. Dawson, D. Gaševic, and N. Mirriahi, "Using learning analytics to scale the provision of personalised feedback," *Brit. J. Educ. Technol.*, vol. 50, no. 1, pp. 128–138, Jan. 2019.

[9] N. Ahmad Uzir, D. Gaševic, W. Matcha, J. Jovanovic, and A. Pardo, "Analytics of time management strategies in a flipped classroom," *J. Comput. Assist. Learn.*, vol. 36, no. 1, pp. 70–88, Feb. 2020.

[10] S. Dawson, D. Gaševic, G. Siemens, and S. Joksimovic, "Current state and future trends: A citation network analysis of the learning analytics field," in *Proc. 4th Int. Conf. Learn. Anal. Knowl.*, Mar. 2014, pp. 231–240.

[11] D. Gaševic, S. Dawson, T. Rogers, and D. Gasevic, "Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success," *Internet Higher Educ.*, vol. 28, pp. 68–84, Jan. 2016.

[12] I. Jivet, M. Scheffel, M. Specht, and H. Drachsler, "License to evaluate: Preparing learning analytics dashboards for educational practice," in *Proc. 8th Int. Conf. Learn. Anal. Knowl.*, Mar. 2018, pp. 31–40.

[13] G. Siemens, "Learning analytics: The emergence of a discipline," *Amer. Behav. Scientist*, vol. 57, no. 10, pp. 1380–1400, 2013.

[14] J. Broadbent, "Comparing online and blended learner's self-regulated learning strategies and academic performance," *Internet Higher Educ.*, vol. 33, pp. 24–32, Apr. 2017.

[15] L. Dörrenbächer and F. Perels, "Self-regulated learning profiles in college students: Their relationship to achievement, personality, and the effectiveness of an intervention to foster self-regulated learning," *Learn. Individual Differences*, vol. 51, pp. 229–241, Oct. 2016.

[16] A. Tlili, M. Denden, F. Essalmi, M. Jemni, M. Chang, and N.-S. Chen, "Automatic modeling learner's personality using learning analytics approach in an intelligent moodle learning platform," *Interact. Learn. Environ.*, vol. 31, no. 5, pp. 1–15, 2019.

[17] I. G. Ndukwe and B. K. Daniel, "Teaching analytics, value and tools for teacher data literacy: A systematic and tripartite approach," *Int. J. Educ. Technol. Higher Educ.*, vol. 17, no. 1, pp. 1–31, Dec. 2020.

[18] O. Viberg, M. Hatakka, O. Bälter, and A. Mavroudi, "The current landscape of learning analytics in higher education," *Comput. Hum. Behav.*, vol. 89, pp. 98–110, Dec. 2018.

[19] D. Ifenthaler and J. Y.-K. Yau, "Utilising learning analytics to support study success in higher education: A systematic review," *Educ. Technol. Res. Develop.*, vol. 68, no. 4, pp. 1961–1990, Aug. 2020.

[20] L.-K. Lee, S. K. S. Cheung, and L.-F. Kwok, "Learning analytics: Current trends and innovative practices," *J. Comput. Educ.*, vol. 7, no. 1, pp. 1–6, Mar. 2020.

[21] A. Iftikhar, S. Musa, M. M. Alam, R. Ahmed, M. M. Su'ud, L. M. Khan, and S. M. Ali, "Risk classification in global software development using a machine learning approach: A result comparison of support vector machine and K-nearest neighbor algorithms," *J. Inf. Technol. Res.*, vol. 15, no. 1, pp. 1–21, Apr. 2022.

[22] A. Iftikhar, M. Alam, R. Ahmed, S. Musa, and M. M. Su'ud, "Risk prediction by using artificial neural network in global software development," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–25, Dec. 2021.

[23] E. Duval and K. Verbert, "Learning analytics," *Eleed*, vol. 8, no. 1, pp. 1–12, 2012.

[24] A. Rubel and K. M. L. Jones, "Student privacy in learning analytics: An information ethics perspective," *Inf. Soc.*, vol. 32, no. 2, pp. 143–159, Mar. 2016.

[25] Á. F. Agudo-Peregrina, S. Iglesias-Pradas, M. Á. Conde-González, and Á. Hernández-García, "Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning," *Comput. Hum. Behav.*, vol. 31, pp. 542–550, Feb. 2014.

[26] W. Xing, R. Guo, E. Petakovic, and S. Goggins, "Participation-based student final performance prediction model through interpretable genetic programming: Integrating learning analytics, educational data mining and theory," *Comput. Hum. Behav.*, vol. 47, pp. 168–181, 2015.

[27] G. Siemens, "Learning analytics and open, flexible, and distance learning," *Distance Educ.*, vol. 40, no. 3, pp. 414–418, Jul. 2019.

[28] S. Aslam and I. Ashraf, "Data mining algorithms and their applications in education data mining," *Int. J.*, vol. 2, no. 7, pp. 1–10, 2014.

[29] A. Rafique, M. S. Khan, M. H. Jamal, M. Tasadduq, F. Rustam, E. Lee, P. B. Washington, and I. Ashraf, "Integrating learning analytics and collaborative learning for improving student's academic performance," *IEEE Access*, vol. 9, pp. 167812–167826, 2021.

[30] F. G. Karaoglan Yilmaz, "Modeling different variables in flipped classrooms supported with learning analytics feedback = Ögrenme analitigi geribildirimleri ile desteklenmis ters-yüz ögrenme ortaminin çesitli degiskenler açisindan modellenmesi," *Online Submission*, vol. 1, no. 2, pp. 78–93, 2020.

[31] D. Ifenthaler and C. Widanapathirana, "Development and validation of a learning analytics framework: Two case studies using support vector machines," *Technol., Knowl. Learn.*, vol. 19, nos. 1–2, pp. 221–240, Jul. 2014.

[32] G. Siemens and P. Long, "Penetrating the fog: Analytics in learning and education," *EDUCAUSE Rev.*, vol. 46, no. 5, p. 30, 2011.

[33] C. Romero, S. Ventura, M. Pechenizkiy, and R. S. Baker, *Handbook of Educational Data Mining*. Boca Raton, FL, USA: CRC Press, 2010.

[34] C. Silva and J. Fonseca, "Educational data mining: A literature review," in *Europe and MENA Cooperation Advances in Information and Communication Technologies*, 2017, pp. 87–94.

[35] F. Marbouti, H. Diefes-Dux, and J. Strobel, "Building course-specific regression-based models to identify at-risk students," in *Proc. ASEE Annu. Conf. Expo.*, 2015, pp. 304–326.

[36] V. T. N. Chau and N. H. Phung, "A random forest-based self-training algorithm for study status prediction at the program level: Minsemi-RF," in *Proc. Int. Workshop Multi-Disciplinary Trends Artif. Intell.* Cham, Switzerland: Springer, 2016, pp. 219–230.

[37] A. Sandoval, C. Gonzalez, R. Alarcon, K. Pichara, and M. Montenegro, "Centralized student performance prediction in large courses based on low-cost variables in an institutional context," *Internet Higher Educ.*, vol. 37, pp. 76–89, Apr. 2018.

[38] S. Sarwat, N. Ullah, S. Sadiq, R. Saleem, M. Umer, A. A. Eshmawi, A. Mohamed, and I. Ashraf, "Predicting students' academic performance with conditional generative adversarial network and deep SVM," *Sensors*, vol. 22, no. 13, p. 4834, Jun. 2022.

[39] A. Rashid, S. Asif, N. Anwer Butt, and I. Ashraf, "Feature level opinion mining of educational student feedback data using sequential pattern mining and association rule mining," *Int. J. Comput. Appl.*, vol. 81, no. 10, pp. 31–38, Nov. 2013.

[40] G. Kostopoulos, S. Kotsiantis, and P. Pintelas, "Predicting student performance in distance higher education using semi-supervised techniques," in *Model and Data Engineering*. Cham, Switzerland: Springer, 2015, pp. 259–270.

[41] C. Márquez-Vera, A. Cano, C. Romero, A. Y. M. Noaman, H. Mousa Fardoun, and S. Ventura, "Early dropout prediction using data mining: A case study with high school students," *Expert Syst.*, vol. 33, no. 1, pp. 107–124, Feb. 2016.

[42] A. Cano and J. D. Leonard, "Interpretable multiview early warning system adapted to underrepresented student populations," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 198–211, Apr. 2019.

[43] G. Kostopoulos, S. Karlos, and S. Kotsiantis, "Multiview learning for early prognosis of academic performance: A case study," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 212–224, Apr. 2019.

[44] F. Jiménez, A. Paoletti, G. Sánchez, and G. Sciavicco, "Predicting the risk of academic dropout with temporal multi-objective optimization," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 225–236, Apr. 2019.

[45] F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, "Models for early prediction of at-risk students in a course using standards-based grading," *Comput. Educ.*, vol. 103, pp. 1–15, Dec. 2016.

[46] D. Monllaó Olivé, D. Q. Huynh, M. Reynolds, M. Dougiamas, and D. Wiese, "A quest for a one-size-fits-all neural network: Early prediction of students at risk in online courses," *IEEE Trans. Learn. Technol.*, vol. 12, no. 2, pp. 171–183, Apr. 2019.

[47] W. Xing and D. Du, "Dropout prediction in MOOCs: Using deep learning for personalized intervention," *J. Educ. Comput. Res.*, vol. 57, no. 3, pp. 547–570, Jun. 2019.

[48] H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, "Classification and prediction of student performance data using various machine learning algorithms," *Mater. Today, Proc.*, vol. 80, pp. 3782–3785, Jan. 2023.

[49] S. Hussain and M. Q. Khan, "Student-performulator: Predicting students' academic performance at secondary and intermediate level using machine learning," *Ann. Data Sci.*, vol. 10, no. 3, pp. 637–655, Jun. 2023.

[50] T. G. Dietterich, "Ensemble methods in machine learning," in *Proc. Int. Workshop Multiple Classifier Syst.* Cham, Switzerland: Springer, 2000, pp. 1–15.

[51] C. Kapucu and M. Cubukcu, "A supervised ensemble learning method for fault diagnosis in photovoltaic strings," *Energy*, vol. 227, Jul. 2021, Art. no. 120463.

[52] S. Kotsiantis, K. Patriarcheas, and M. Xenos, "A combinational incremental ensemble of classifiers as a technique for predicting students' performance in distance education," *Knowl.-Based Syst.*, vol. 23, no. 6, pp. 529–535, Aug. 2010.

[53] M. Injadat, A. Moubayed, A. B. Nassif, and A. Shami, "Systematic ensemble model selection approach for educational data mining," *Knowl.-Based Syst.*, vol. 200, Jul. 2020, Art. no. 105992.

[54] M. Pandey and S. Taruna, "A comparative study of ensemble methods for students' performance modeling," *Int. J. Comput. Appl.*, vol. 103, no. 8, pp. 26–32, Oct. 2014.

[55] R. Conijn, C. Snijders, A. Kleingeld, and U. Matzat, "Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 17–29, Jan. 2017.

[56] A. Asselman, M. Khaldi, and S. Aammou, "Enhancing the prediction of student performance based on the machine learning XGBoost algorithm," *Interact. Learn. Environ.*, vol. 31, no. 6, pp. 3360–3379, Aug. 2023.

[57] A. Kukkar, R. Mohana, A. Sharma, and A. Nayyar, "Prediction of student academic performance based on their emotional wellbeing and interaction on various e-learning platforms," *Educ. Inf. Technol.*, vol. 28, no. 8, pp. 9655–9684, Aug. 2023.

[58] I. Wickramasinghe and H. Kalutarage, "Naive bayes: Applications, variations and vulnerabilities: A review of literature with code snippets for implementation," *Soft Comput.*, vol. 25, no. 3, pp. 2277–2293, Feb. 2021.

[59] I. Karo, M. Fajari, N. Fadhilah, and W. Wardani, "Benchmarking naïve Bayes and ID3 algorithm for prediction student scholarship," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 1232, no. 1, 2022, Art. no. 012002.

[60] K. Gajowniczek and T. Zabkowski, "ImbTreeEntropy and ImbTreeAUC: Novel R packages for decision tree learning on the imbalanced datasets," *Electronics*, vol. 10, no. 6, p. 657, Mar. 2021.

[61] F. Qiu, G. Zhang, X. Sheng, L. Jiang, L. Zhu, Q. Xiang, B. Jiang, and P.-K. Chen, "Predicting students' performance in e-learning using learning process and behaviour data," *Sci. Rep.*, vol. 12, no. 1, pp. 1–15, Jan. 2022.

[62] E. Rozos, P. Dimitriadis, K. Mazi, and A. D. Koussis, "A multilayer perceptron model for stochastic synthesis," *Hydrology*, vol. 8, no. 2, p. 67, Apr. 2021.

[63] B. Ketzler, V. Naserentin, F. Latino, C. Zangelidis, L. Thuvander, and A. Logg, "Digital twins for cities: A state of the art review," *Built Environ.*, vol. 46, no. 4, pp. 547–573, Dec. 2020.

[64] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.

[65] D. Opitz and R. Maclin, "Popular ensemble methods: An empirical study," *J. Artif. Intell. Res.*, vol. 11, pp. 169–198, Aug. 1999.

[66] R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 161–168.

[67] L. Rokach, "Ensemble-based classifiers," *Artif. Intell. Rev.*, vol. 33, nos. 1–2, pp. 1–39, Feb. 2010.

[68] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proc. Int. Conf. Mach. Learn.*, vol. 96. 1996, pp. 148–156.

[69] D. H. Wolpert, "Stacked generalization," *Neural Netw.*, vol. 5, no. 2, pp. 241–259, Jan. 1992.

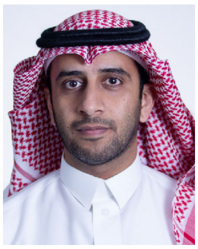
[70] L. I. Kuncheva, *Combining Pattern Classifiers: Methods Algorithms*. Hoboken, NJ, USA: Wiley, 2014.

[71] Ö. Gür Ali and U. Aritürk, "Dynamic churn prediction framework with more effective use of rare event data: The case of private banking," *Expert Syst. Appl.*, vol. 41, no. 17, pp. 7889–7903, Dec. 2014.

NAVEED ANWER BUTT received the Ph.D. degree in computer science from International Islamic University Islamabad. He has been an Assistant Professor with the Department of Computer Science, Faculty of Computing and Information Technology, University of Gujrat, since 2007. His research work and teaching career spans over 21 years in various organization and HEC recognized universities and institutes. His research interests include artificial intelligence, neuroscience, behavioral data modeling and prescriptive analytics, data visualization, exploríty data analytics, big data solution engineering, data mining, opinion mining, and process mining.

ZAFAR MAHMOOD is currently an Assistant Professor with the Department of Computer Science, Faculty of Computing and Information Technology, University of Gujrat. His research interests include data science and machine learning.

KHAWAR SHAKEEL is currently with the Department of Computer Science, Faculty of Computing and Information Technology, University of Gujrat. His research interests include data science and machine learning.



SULTAN ALFARHOOD received the Ph.D. degree in computer science from the University of Arkansas. He is currently an Assistant Professor with the Department of Computer Science, King Saud University (KSU). Since joining KSU, in 2007, he has made several contributions to the field of computer science through his research and publications. He has published several research articles on cutting-edge topics, such as machine learning, recommender systems, linked open data, and text mining. His work includes proposing innovative approaches and techniques to enhance the accuracy and effectiveness of recommender systems and sentiment analysis.

MEJDL SAFRAN received the M.Sc. and Ph.D. degrees in computer science from Southern Illinois University, Carbondale, in 2013 and 2018, respectively. He is currently an Assistant Professor in computer science with King Saud University. He has been a Faculty Member, since 2007. He has published articles in refereed journals and conference proceedings, such as *ACM Transactions on Information Systems*, *Applied Computing and Informatics*, *Biomedicine* (MDPI), *Sensors* (MDPI), IEEE International Conference on Cluster, IEEE International Conference on Computer and Information Science, International Conference on Database Systems for Advanced Applications, and International Conference on Computational Science and Computational Intelligence. His current research interests include computational intelligence, artificial intelligence, deep learning, pattern recognition, predictive analytics, developing efficient recommendation algorithms for large-scale systems, predictive models for online human activities, machine learning algorithms for performance management, and modeling and analyzing user behavior. Since 2018, he has been providing part-time consulting services in the field of artificial intelligence to private and public organizations and firms.

IMRAN ASHRAF received the M.S. degree (Hons.) in computer science from the Blekinge Institute of Technology, Karlskrona, Sweden, in 2010, and the Ph.D. degree in information and communication engineering from Yeungnam University, Gyeongsan, South Korea, in 2018. He was a Postdoctoral Fellow with Yeungnam University, where he is currently an Assistant Professor with the Information and Communication Engineering Department. His research interests include positioning using next-generation networks, communication in 5G and beyond, location-based services in wireless communication, smart sensors (LIDAR) for smart cars, and data analytics.