

Article

Predicting Academic Performance Using an Efficient Model Based on Fusion of Classifiers

Ansar Siddique ^{1,*}, Asiya Jan ², Fiaz Majeed ², Adel Ibrahim Qahmash ³, Noorulhasan Naveed Quadri ⁴  and Mohammad Osman Abdul Wahab ⁵

¹ Department of Software Engineering, University of Gujrat, Gujrat 50700, Pakistan

² Department of Information Technology, University of Gujrat, Gujrat 50700, Pakistan; 19025956-003@uog.edu.pk (A.J.); fiaz.majeed@uog.edu.pk (F.M.)

³ College of Education, King Khalid University, Abha 61413, Saudi Arabia; aqahmash@kku.edu.sa

⁴ College of Computer Science, King Khalid University, Abha 61413, Saudi Arabia; qnaveed@kku.edu.sa

⁵ Department of English, Faculty of Languages and Translation, King Khalid University, Abha 61413, Saudi Arabia; moothman@kku.edu.sa

* Correspondence: dr.ansarsiddique@uog.edu.pk

Abstract: In the past few years, educational data mining (EDM) has attracted the attention of researchers to enhance the quality of education. Predicting student academic performance is crucial to improving the value of education. Some research studies have been conducted which mainly focused on prediction of students' performance at higher education. However, research related to performance prediction at the secondary level is scarce, whereas the secondary level tends to be a benchmark to describe students' learning progress at further educational levels. Students' failure or poor grades at lower secondary negatively impact them at the higher secondary level. Therefore, early prediction of performance is vital to keep students on a progressive track. This research intended to determine the critical factors that affect the performance of students at the secondary level and to build an efficient classification model through the fusion of single and ensemble-based classifiers for the prediction of academic performance. Firstly, three single classifiers including a Multilayer Perceptron (MLP), J48, and PART were observed along with three well-established ensemble algorithms encompassing Bagging (BAG), MultiBoost (MB), and Voting (VT) independently. To further enhance the performance of the abovementioned classifiers, nine other models were developed by the fusion of single and ensemble-based classifiers. The evaluation results showed that MultiBoost with MLP outperformed the others by achieving 98.7% accuracy, 98.6% precision, recall, and F-score. The study implies that the proposed model could be useful in identifying the academic performance of secondary level students at an early stage to improve the learning outcomes.

Keywords: educational data mining; supervised learning; secondary education; academic performance



Citation: Siddique, A.; Jan, A.; Majeed, F.; Qahmash, A.I.; Quadri, N.N.; Wahab, M.O.A. Predicting Academic Performance Using an Efficient Model Based on Fusion of Classifiers. *Appl. Sci.* **2021**, *11*, 11845. <https://doi.org/10.3390/app112411845>

Academic Editors: Kamran Shaukat, Suhuai Luo and Federico Divina

Received: 21 October 2021

Accepted: 9 December 2021

Published: 13 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Educational data mining (EDM) is a growing area of research that is being used to explore educational data for different academic purposes. The main application of EDM is the prediction of students' academic performance [1,2]. In data mining, the analysis and interpretation of student academic performance are regarded as suitable analysis, evaluation, and assessment tools [3]. In the present era of a knowledge economy, the students are the key element for the socio-economic growth of any country, so keeping their performance on track is essential. Data mining (DM) methods are applied to learn hidden knowledge and patterns which assist administrators and academicians in decision making regarding the delivery of instructions. DM techniques have applications in numerous areas including retail business, the health sector, marketing, banking, bioinformatics, counter-terrorism, and many others are also using it to enhance productivity and efficiency [4].

Education plays a vital role in the development of any nation. The education in Pakistan has improved in the past few years; it is further striving to enhance academic performance to produce a well-educated and competitive workforce to meet the requirements of the market [5]. The academic performance of high school students is a critical concern for parents, teachers, the education department, and the government. The performance of students at the secondary level is highly influenced by their demographic, schooling, social background, family background, and psychological factors [6]. Individuals vary from one another in terms of these factors, and hence their academic performance differs accordingly [5]. It is therefore important to predict students' academic performance, taking into account the aforementioned parameters.

The performance of secondary level students in science subjects is expected to be high in order to provide quality entrants at a higher level of education, which is key to prosperity and the knowledge economy [7]. Therefore, the lower secondary level is very crucial for students aiming for science as a major subject, because it is a baseline for setting their academic goals. Ninth-grade students need to consider grades as imperative at the lower secondary level as they act as a stepping-stone in subsequent educational levels. Academic performance can be enhanced by identifying weak areas concerning academics and personal traits. To achieve this, the design of a prediction system is indispensable to estimation of students' academic progress at an early stage before they in board examination. Such prediction may be beneficial in multiple ways. For example, it can help in reinforcing students to improve their performance in their weak areas. Secondly, it may assist in the selection of subjects where students can perform better. Thirdly, it can be useful in deciding on career goals. Finally, it can assist in awarding accurate grades to students based on their previous performance in COVID-19-like situations where administering examinations is impossible. Therefore, a prediction system should be in place to predict the performance of ninth-grade students at an early stage of the academic year to keep them on track and enable them to perform better in board examinations. The basic objective is to minimize the dropout and failure ratios of students at the lower secondary level. Additionally, the selection of subjects in grade nine is considered important in students' academic growth and professional advancement. To support education in a traditional setting, numerous systems such as massive open online courses (MOOCs), intelligent tutoring systems (ITS) and web-based educational systems have been designed but there is a dearth of systems to predict students' academic progress [8]. As per our knowledge, in Pakistan, there is no prediction system available to measure students' progress at the lower secondary level.

It has been described that data mining approaches analyze an organization's historical data and figure out the required pattern or information that is otherwise impossible [9]. In DM, both classification and regression are used to build predictive models. The classification techniques are used on pre-classified data to characterize the unclassified data [10]. The classification can be performed using different learning schemes such as decision trees (ID3, REP Tree and C4.5, etc.), logistic regression (LogitBoost), backtracking (MLP), and probability (Naïve Bayes, Bayes Network). These algorithms are known as single classifiers which have some limitations for the accuracy of the model. It is therefore important to improve the performance of single classifiers. To achieve such a goal, the ensemble method has been introduced which combines different classifiers into a single unit. However, in the ensemble structure, many areas must be explored to increase the accuracy of prediction models. A research study has been conducted to discover other algorithms of learning schemes in different ensemble methods which can help in predicting students' academic performance with greater accuracy [11]. This research intends to propose a classification model developed by fusion of single-based classifiers and ensemble-based classifiers to predict the academic performance of secondary-level students based on their academic and personal traits.

The prediction of students' grades from their academic data along with other features is a useful application in EDM; it is therefore becoming a suitable source of information that can be utilized in multiple ways to improve the quality of education. The major

contributions of this research study towards EDM include (a) the identification of the personal traits of secondary school students. (b) Development of a dataset by gathering data from four different secondary schools located in three different cities. The academic data were collected through a student information system and data related to personal characteristics and socioeconomic conditions was gathered using online and physical surveys. (c) Building of a model which analyzes students' academic and personal data and predicts their academic progress with higher accuracy and precision.

The rest of the paper is organized as follows: Section 2 presents a literature review that highlights the findings of prior research studies in order to identify the set of the most common factors affecting students' academic performance. Secondly, it determines the most frequently used data mining techniques in previous research. Section 4 presents the research methodology, and the section is about data collection and data preprocessing. In Section 5, the conclusion and future prospects of this research study have been presented.

2. Literature Review

EDM is burgeoning due to the massive growth of educational resources, the internet, and the usage of online tools to impart education [12]. Consistent research efforts are being made to improve the quality of educational tools. This section provides an overview of the factors that may affect students' academic progress and technologies that are helpful in making predictions regarding students' academic progress. A systematic review has been conducted to determine factors that affect student performance through information mining procedures [13]. The study handled numerous subjects, one of which was to recognize the significant characteristics which can be utilized in anticipating student performance. The results indicated that the internal assessments and aggregate evaluation points are the most incessant qualities utilized for predicting academic performance. Additionally, other significant properties were identified including personal and inner appraisal, previous record, extra-curricular activities, and social attributes. The decision tree and neural networks were found as the most regularly utilized information-digging strategies [14].

Another research study [15] considered the cumulative grade point average as a prominent factor for measuring student performances in each semester. Additionally, the study also interprets normal class tests and assignments, previous academic failure, and study duration as the appreciable factors for predicting student performances. In [16] author stated that academic attainments are somehow linked with students' extracurricular activities and presented the presence of the student in a class as the strongest predictor for academic performance.

The study [17] found that family attributes and academic attributes were the deciding factors for prediction. The cumulative grade point average of students and their external and internal assessments marks were also the most frequently used attributes by the researchers. Another research [18] also showed the performance prediction of fourth-year undergrad students using pre-university marks and considering marks obtained in courses of the second year. They only considered grades for performance prediction and ignored the family and socio-economic attributes of the students.

A comprehensive analysis of supervised machine learning techniques was conducted and applied to predict students' performance in the examination. They considered different factors including demographics and social interest to predict students' expected score in the final term as well as students at risk [19].

A survey study was conducted using a sample of 1500 United State students to identify the impact of the COVID-19 pandemic on higher education. It has been revealed through analysis and a classification model that the shocks related to health and economics brought about by COVID-19 varied by socioeconomic factors [20]. Another study gathered students' pre-university marks, first and second-year marks through a large-sized sample and applied a predictive model on it to envisage students' CGPA at the final semester [21].

Table 1 shows the set of attributes used at the secondary level to predict students' academic performance.

Table 1. Student attributes used for predicting performances.

Student Attributes	Possible Values Used in All Research Papers While Implementing the DM Algorithm
Academic Attributes	Internal and external assessment, lab marks, sessional marks, attendance, Cumulative Grade Point Average (CGPA), semester marks, grade, seminar performance, assignment, attendance, schools marks, previous academic marks, etc.
Personal Attributes	Age, gender, height, weight, Emotional Intelligence (EI), student interest, level of motivation, communication, sports person, hobbies and ethnicity, etc.
Family Attributes	Qualification, occupation, income, status, support, siblings, responsibilities, etc.
Social Attributes	Number of friends, social networking, girls'/boys' friends, movies, travel outings, friends' parties, etc.
School Attributes	Teaching medium, accommodation, infrastructure, water and toilet facilities, transportation system, class size, school reputation, school status, class size, school type, teaching methodology, etc.

Another research study [22] proposed a predictive investigation framework to gauge the satisfaction level of university students regarding an online summer program. The students were given a questionnaire, consisting of questions related to socioeconomic, demographics, and some other indicators. Regression and ANOVA analysis were applied in that prediction model to interpret learners' interaction. Data mining approaches are being widely used by academic analysts to assess the effectiveness of education by processing and interpreting the huge volumes of data [23]. For the timely completion of students' degree requirements, their future performance was predicted based on their academic record. A novel machine learning method was used to predict the students' performance in a degree program [24]. A brief overview of multiple machine learning techniques along with comparison of time complexity of models was presented. The work shows the current limitations and challenges of machine learning techniques [25].

It has been shown that composite methods are vital for improving single classifiers and accuracy of predictive models. Bagging, Boosting and stacking, etc. are different types of ensemble methods that use a blend of models to improve composite models. Among them, bagging is utilized for classification and prediction purposes. The study handled the imbalance dataset with a SMOTTEEN technique with ensemble classifiers to produce high results. Since every model comprises some limitations, so the ultimate purpose of ensemble methods was to join the strength of single different models with the aim to achieve higher accuracy [26].

Another research presented two prediction models for estimating the student performance in final examinations. A K-Nearest Neighbor algorithm and support vector machine algorithm were used as a prediction technique to estimate the students' performance in final examinations on the basis of their demographic, class and social attributes [27].

In the field of EDM, researchers made efforts to study different kinds of attributes and properties that influence students' learning and performance results [28].

There are so many data mining approaches, but some mining approaches are being considered more effective. Such as machine learning-based ensemble methods also known as composite methods which are being considered as a vital approach to strengthening single classifiers. This approach leverages the power of multiple models to attain improved prediction accuracy than any of the individual models could achieve independently [29].

3. Research Methodology

This section describes the research methodology in detail. The implementation of the data mining approach is completely described in this section. The Waikato Environment for Knowledge Analysis (WEKA) [30] was used to perform data mining tasks. The research methodology consists of different phases and experiments conducted during this research. The pictorial representation of research methodology is given in Figure 1.

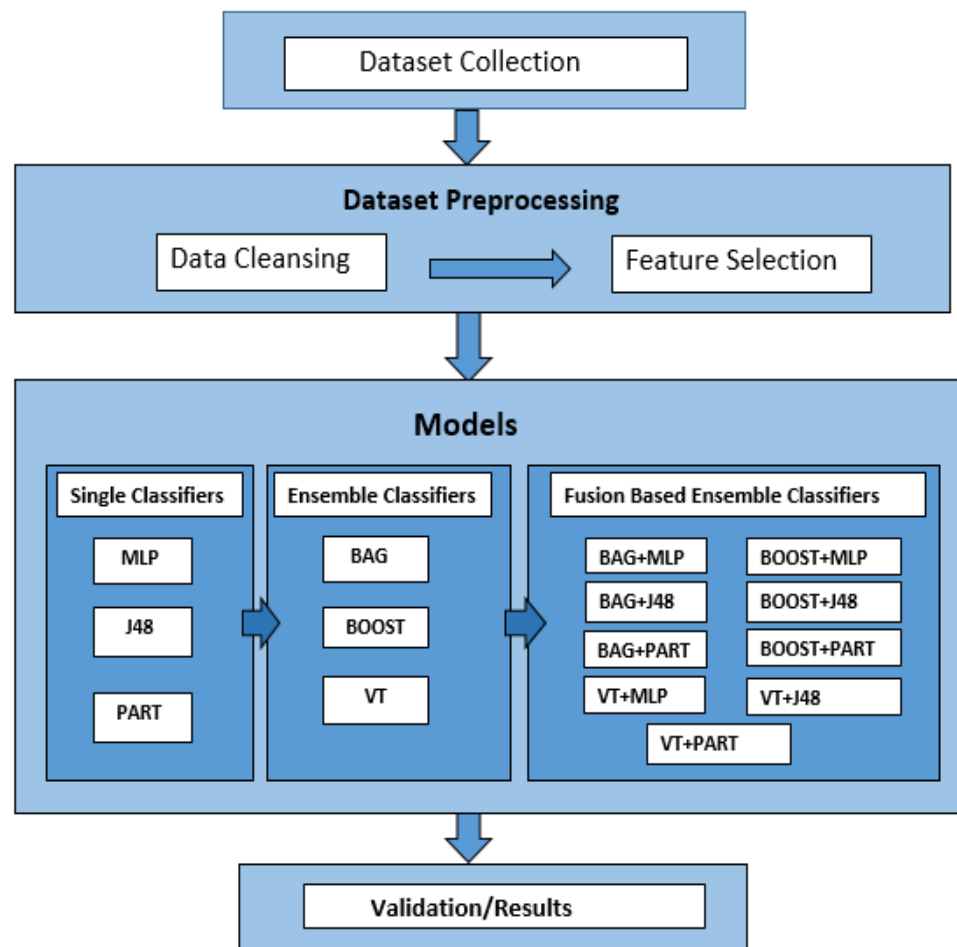


Figure 1. Proposed methodology.

The data collection was based on attributes suggested by researchers as the most rational attributes to predict academic performance at secondary level of education. The dataset was collected using an online and physical survey from four different schools based on students' academic, demographic, social, and family attributes. A dataset comprising 1227 records was collected that is currently available online at Kaggle. It includes 21 attributes in total of four different types including demographics, family, social, and academic. A total of 16 attributes remained after a feature selection process. The outcome is categorized into seven classes of grades including A+, A, B+, B, C, D and F. This section demonstrates the detailed description of selected attributes. Table 2 provides a brief description of every attribute that was used in this study.

The first step in data pre-processing is data cleansing, which is used to remove all the irrelevant attributes. The dataset consists of 1227 records in total. After detecting the missing values from different features, some records were removed from the dataset. The removal of records does not lead a model towards biasness, if dataset is used for training as a whole for each algorithm instead of subset of the dataset [31,32]. To reduce the computational complexity while implementing the mining techniques, missing values were also removed. The second step in data preprocessing is feature selection which is used to reduce dimensionality in feature space and obtain better classification results [26] because training on high-dimensional data leads to overfitting of the model. The subset of original features have been picked up through feature selection method which leads to the removal of redundant and obsolete characteristics without losing any important information [26]. This study applied filter-based methods using information gain-based selection to evaluate crucial features which may help in developing good performance models. Filter-based

feature selection is a ranking method, used to rank the attributes according to their rank values by overlooking the remaining ones and then through application to the learning ones. In a rank-based method, values were given to each attribute according to their ranks in building a good model. Information gain is a filter-based feature ranking technique that is based on the information theory where information is provided about the target class attribute given the value of the dependent class attribute [33]. Out of 21 features, 16 were selected based on the rank that is close in its relationship with the final predicted outcome and gives better results. The foremost motive of the proposed model is to predict the students' performances under the classes such as A+, A, B+, B, C, D, and F, where nominal and numeric data are converted into ordinal values with the help of the discretization technique. The class distribution is shown in Table 2.

Table 2. Dataset description and possible values.

S. No	Attributes	Description
1	GE	Gender (Male, Female)
2	HA	Home Address Urban, Rural)
3	PCA	Parent Cohabitation Status (Living together, Apart)
4	QFR	Quality of family Relationship (Very Good, Good, Not Good)
5	MJ	Mother Job (Yes, No)
6	FJ	Father Job (Yes, No)
7	ME	Mother Education (None, Elementary, Secondary, Higher)
8	FE	Father Education (None, Elementary, Secondary, Higher)
9	FS	Family Size (Less than 3, Greater or equal to 3)
10	GF	Going out with friends (Yes, No)
11	PF	Past Failures (Yes, No)
12	NS	Attended Nursery School (Yes, No)
13	HE	Want to take Higher Education (Yes, No)
14	R	Relationship (Yes, No)
15	IA	Internet access at home (Yes, No)
16	ECA	Extra-Curricular Activities (Yes, No)
17	DST	Daily Study Time (<2 h, 2 to 5 h, 5 to 10 h, >10 h)
18	HST	Home to school Travel Time (<15 min, 15 to 30 min, 30 min to 1 h, >1 h)
19	EG	8th Class Grades (A+, A, B+, B, C, D)
20	NG1	9th Class First Term Grades (A+, A, B+, B, C, D, F)
21	NG2	9th Class Final Term Grades (A+, A, B+, B, C, D, F)

The research methodology is mainly premised on ensemble methods including bagging, boosting, and stacking, which is a different kind of ensemble method which uses a blend of models [29]. Among these methods, bagging, boosting, and stacking can be utilized for classification and prediction. Each model has some strengths and limitations, so the ultimate objective of ensemble methods is to complement the models, in order to achieve higher prediction accuracy. The bagging method is used to sort the tuples randomly into different bags while developing a model. The process is known as bootstrap aggregation. All models in bagging are built in parallel and, for the overall decision, an average is taken from all models, which lowers the variance in the model. It has been shown that bagging achieves the highest efficiency relative to the other methods [34,35].

Bootstrap improves on bagging. It is developed sequentially by assigning weights to the tuples classified incorrectly by previous classifiers and thus receive more attention

from the next classifier. The weighted average is finally taken to build the final decision. The boosting algorithms are highly capable, take weak and low-performing models, and convert them into robust models. The boosting classifier methods include AdaBoost, GradientBoost and MultiBoost, etc. [36].

The effectiveness of the boosting method has been examined and it was found that the method is an efficient and effective strategy for classification and prediction [37,38].

Random forest is also an ensemble method which is an improved version of bagging, and it is used for classification and regression. In the training phase, it creates multiple decision trees and produces the mode of classes as well as generating a mean prediction of an individual in regression problems. It also performs random sampling on features with the help of feature engineering. It builds prediction models based on the aggregation of decision trees [39]. Another ensemble method is voting, which combines the output predictions from multiple models. This technique is used to enhance the performance of models in comparison to any single classifier model. The technique is mainly used for regression and classification problems. In the classification method, the prediction from multiple models of each label is aggregated and the majority vote label is predicted, which may be considered as a meta-model. The method of combining the decision of different algorithms requires stacking algorithms. The most common way to develop the training dataset for the meta-model is through k-fold cross-validation of the base models, where the out-of-fold predictions are used as a premise for the training dataset for the meta-model. After training, a meta-model from different models is assembled and is trained on the resultant of component models. Thus the heterogeneous ensemble model is created using this approach as the component model comprises diverse algorithms [37–39].

The importance of different ensemble techniques such as bagging, boosting, and voting classifiers can be viewed through various recent studies where hybridization of ensemble classifiers with base classifiers is a current research trend. Livieris et al. proposed a prediction model which was based on Bagging and Boosting and created two strategies to successfully combine the predictions of weight-constrained neural networks (WCNNs) [40]. Similarly, another study has explored bagging, boosting, and voting classifiers for the automated classification of news articles, in particular concerning identification of fake content from real content. They highlighted that the novel aspect of their research is the use of various ensemble methods including bagging, boosting, and voting classifiers to investigate their performance over multiple datasets [41]. Yang et al. proposed a two-layer ensemble approach to enhance the performance of the software defect prediction process. In the inner layer, they have combined decision tree and bagging to form a random forest model. In the outer layer, they used random under-sampling to train various random forest models and applied staking to ensemble them once again [42].

3.1. Phase 1: Classification Using Base Classifier

Several classification learning schemes are being used for classification and prediction, such as decision trees (REP Tree, C4.5, CART, and J48, etc.), probability (Naïve Bayes and Bayes Networks, etc.), backtracking (ANN like MLP, etc.), logistic regression (Logistic Boost, etc.), and so many other schemes are highly embraced. Any of them used independently are referred to as a single base classifier.

Previous research studies revealed that, among other classification learning schemes, the Multilayer perceptron, J48, and PART are the most efficient and frequently used classifiers for performance prediction. The performance of such classifiers was studied regarding training time, efficiency, and accuracy of prediction. It has been found that J48 took less training time for each data instance than the MLP classifier. The MLP, PART, and J48 showed higher accuracy for both large and small size data sets. Moreover, it has also been examined that the PART algorithm provided better accuracy in every case with and without noise [38].

3.1.1. Multilayer Perceptron Classifier

A multilayer perceptron (MLP) is a feed-in class to the artificial neural network. It utilizes back-propagation for training, referred to as a supervised learning technique. MLP consists of an input layer, a hidden layer, and an output layer. Each node is known as a neuron excluding the input nodes and uses a non-linear ReLu activation function [43]. This non-linear activation along with its multiple layers distinguishes it from linear perceptron. The ReLu activation function is simple and efficient, as it has been empirically observed that training a network with this function tended to converge quickly and reliably in comparison to other activation functions. Furthermore, it also helps in detecting data that cannot be separated linearly [43].

3.1.2. J48 Classifier

The J48 algorithm was developed by Ross Quinlan to classify different datasets and applications to enhance the results of classification. J48 is used to generate decision trees that are based on C4.5 algorithms. Every aspect of data is divided into small subsets based on a decision tree. It employs greedy search and top-down search by all branches to build a decision tree for modeling of the classification process [44]. This decision tree can estimate the missing attributes and deal with certain distinctive and varying features. Furthermore, it can be used to examine the data continuously [45,46].

3.1.3. PART Classifier

The developed version of the ripper algorithm and C4.5 is a partial decision tree algorithm that does not require global optimization to produce appropriate rules for classification [38]. It helps in building a partial decision tree on different sets of instances and produces rules for decision trees [47]. PART is an algorithm that uses a divide-and-conquer mechanism to build a partial C4.5 decision tree in each iteration, i.e., it generates a PART decision list, and makes the best leaf into a rule [44].

3.2. Phase 2: Building Model by Ensemble Methods

Ensemble methods are an influential and efficient development in data mining and machine learning. The philosophy of ensemble classification is based on the decision of a group of experts instead of a single expert. This research uses three main techniques including Bagging (BAG), Boosting (BST), and voting (VT) as these techniques are highly recommended by previous research in order to attain high accuracy and low errors in prediction [46]. Basically, ensemble methods are classified as homogenous and heterogeneous ensemble methods. Homogenous ensemble methods apply a single algorithm on various training datasets to construct multiple classifiers such as bagging and boosting. Conversely, different algorithms are used to manipulate training datasets to make various models in heterogeneous ensemble models including voting and stacking [48].

3.2.1. Bagging

Bagging, or bootstrap aggregation, is an efficient ensemble technique used for classification. The bagging technique takes a sample of data randomly, puts them into different sample-sized bags, and then trains them on a classifier. It is mostly used as an ensemble method to reduce variance in data, randomize the design process, and finally create an ensemble from them [49]. It has also been observed that bagging provides high efficiency [34]. In this research study, the bagging technique samples the data erratically into different sample-sized bags and trains them on random forest, a basic classifier for bagging, and aggregates their specific forecasts into a final prediction [50].

3.2.2. Boosting

Boosting is another robust algorithm of ensemble learning that creates a strong learner from weak learners. It generates many weak learners with the help of a decision tree and combines them to form a strong learner. It helps in reducing errors during prediction and

makes a model less biased. The effectiveness of the boosting method has been examined and it was concluded that it is an efficient and effective strategy for classification and prediction [38]. The boosting technique used in this research is MultiBoostAB (MB), as it is an extension of the AdaBoost technique to form decision committees but with a lower error rate in comparison to AdaBoost [51,52].

3.2.3. Voting

Voting is also known for its heterogeneous nature where classifiers comprise different algorithms which are used to predict the final outcome. It creates two or more models and each of them make their predictions [36]. It is based on an aggregating network that is used to determine the weights' mean. It is a model that combines the result of multiple classifiers on basis of weights [53]. There are three types of voting including majority voting, unanimous voting, and plurality voting. The majority voting reflects more than 50% votes for the final decision; in unanimous voting, all classifiers develop an agreement for final decisions, whereas polarity voting considers the majority of votes to decide the final outcome. In this research, majority voting was used to combine classifiers because it provides better results in terms of accuracy, as indicated by prior research [53,54]. Moreover, three different algorithms including Naive Bayes, IBk, and ZeroR were used in this study for voting.

3.3. Phase 3: Building Model by Hybrid Ensemble Methods

This phase includes the building, training and testing of hybrid ensemble-based models by hybridizing them with base classifiers. The fusion of base classifiers with ensemble models enhances the generalization and prediction capability of ensemble models. The hybrid models of machine learning, build accurate and efficient machine learning models and feed their output to each other [55].

These models include BAG + J48, BAG + MLP, BAG + PART followed by BST + J48, BST + MLP, BST + PART. The last hybrid ensemble-based models were VT + J48, VT + MLP, VT + PART.

3.4. Phase 4: Performance Comparison Analysis

To determine the performance of an algorithm, evaluation metrics were used. The performance validation of models was generated through 10-fold cross-validation. The k-fold cross-validation procedure divides a limited dataset into k non-overlapping folds. Each of the k folds are allowed to be used as a held back test set whilst all other folds are collectively utilized as a training dataset. A total of k models are fit and evaluated on the k holdout test sets and report the mean performance. The evaluation in this study was conducted using a 10-fold cross-validation technique. The technique divided the whole dataset into 10 subsets of equal size; out of 10 subsets, 9 were used for training and the 1 remaining was used for testing. The process was iterated ten times; the final result was estimated as the average error rate on test examples [55].

The evaluation metrics include accuracy, precision, recall, and F-score, which were used to examine the performance of each predictive model. Such predictive models were figured out based on True-Positive (TP), False-Positive (FP), True-Negative (TN), and False-Negative (FN).

4. Experiments and Evaluation

WEKA was used to evaluate the proposed classification model and to make comparisons. In this study, different experiments were conducted sequentially to assess students' performance. The comparison was made through various single base classifiers, ensemble-based classifiers, and fusion ensemble classifiers. The time complexity of each algorithm is also represented in terms of Big O notation which plays an important role in finding the efficiency of algorithms [25]. Additionally, a comparative analysis has been performed to discover performance improvements in different models. The experiments detected the

efficient model in predicting student academic performance at the secondary level. To acquire precise results during evaluation, 10-fold cross-validation was used.

4.1. Experiments with Base Classifiers and Ensemble Base Classifiers

The three base classifiers including MLP, J48, and PART were applied after the data preprocessing stage. The evaluation results showed that among these three base classifiers, MLP outperformed the other classifiers, achieving greater accuracy (i.e., 88.52) as shown in Figure 2. The figure has two parts; in the first part, the bar chart presents the performance of classifiers in terms of accuracy, precision, recall, and F-score. The second part presents the performance of classifiers in tabular form through the same measures. The MLP classifier also performed better in terms of other measures such as precision, recall, and F-score. The time complexity of the MLP classifier is $O(emnk)$, which is a composition of interconnected neurons, whereas J48 and PART have a time complexity of $O(mn^2)$, which works on the If and Then rule until the predicted class has not been obtained.

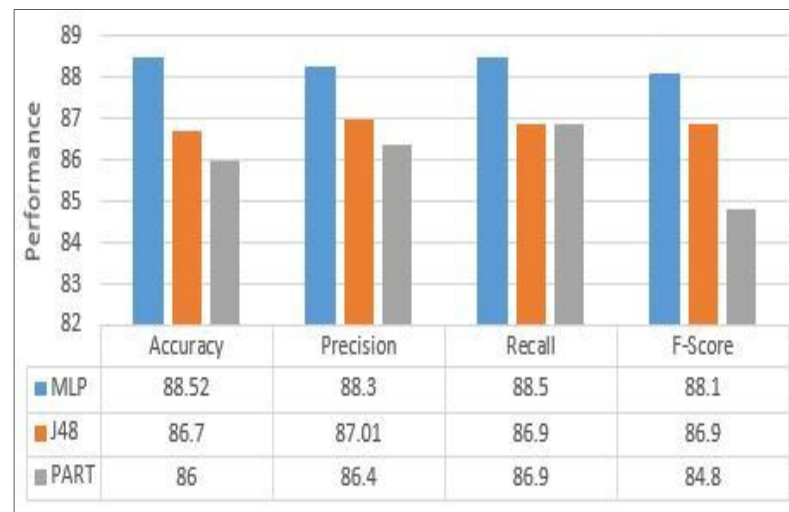


Figure 2. Single-based classifiers.

Furthermore, three different ensemble classifiers including bagging, multiboost, and voting were built. Among these three ensemble classifiers, multiboost outperformed the other classifiers, achieving higher accuracy (i.e., 95.7) as shown in Figure 3. The figure comprises two parts; in the first part, the bar chart shows the performance of classifiers in terms of accuracy, precision, recall, and F-score. The second part indicates the performance of classifiers in tabular form through the same measures. The classifier also performed better in terms of other measures such as precision, recall, and F-score. The time complexity of bagging is $O(k\log n)$, where k is the number of bags.

4.2. Experiments with Fusion Ensemble-Based Models

The aim of this phase was to develop hybridization of ensemble classifiers with single-based classifiers. This experiment evaluated nine fused ensemble models including fusion of BAGGING (BAG) with MLP, PART, and J48, MultiBoost fusion with MLP, PART, and J48 as well as Voting (VT) fusion with MLP, PART, and J48. The results of these models are shown in Figures 4–6.

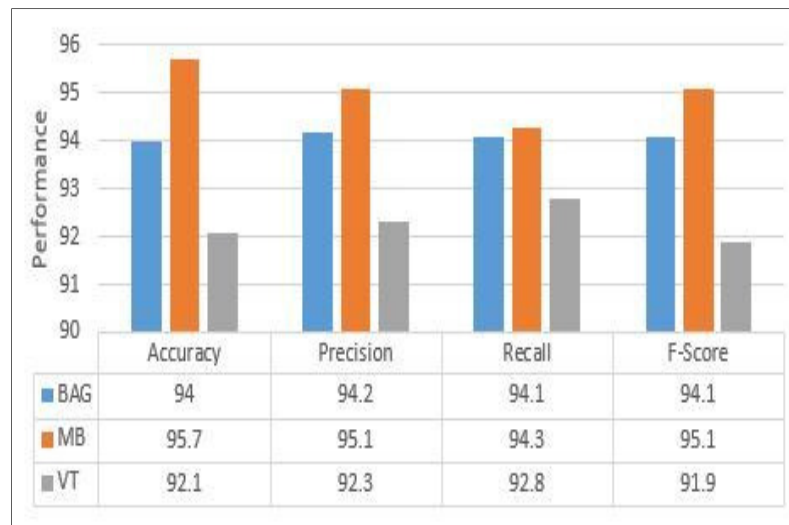


Figure 3. Ensemble-based classifiers.

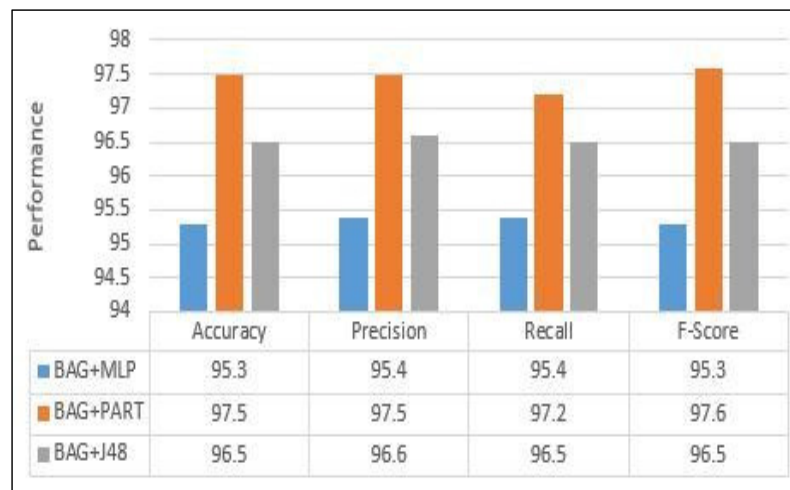


Figure 4. Bagging with single-based classifiers.

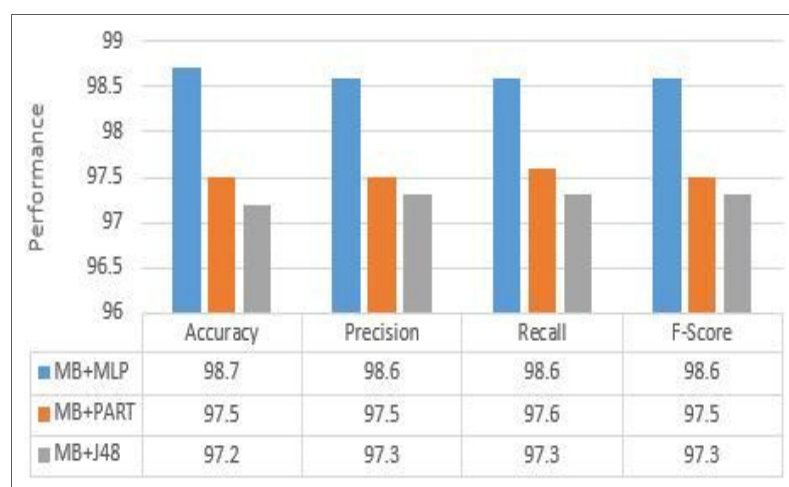


Figure 5. MultiBoostAB with single-based classifiers.

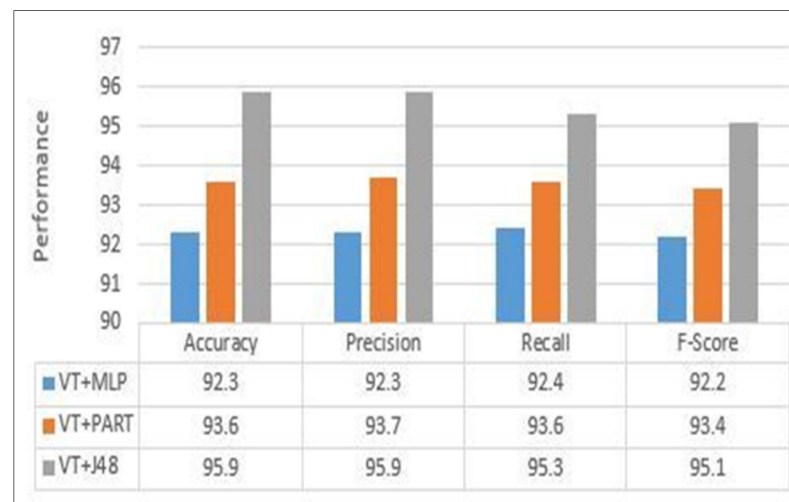


Figure 6. Voting with single-based classifiers.

The evaluation results related to BAG fusion with PART showed the highest accuracy (i.e., 97.50%). The model also performed very well with respect to precision, F-score and recall, as shown in Figure 4. The experiment results related to MB fusion with MLP achieved the highest accuracy (i.e., 98.7), as shown in Figure 5. This model has achieved good performance with regard to precision, recall, and F-score. The results related to the fusion of VT with different single classifiers showed that VT + J48 achieved greater accuracy (i.e., 95.9%), as shown in Figure 6. This model also showed better performance in terms of precision, recall, and F-score. Each of Figures 4–6 consist of two parts; in the first part, the bar chart presents the performance of classifiers in terms of accuracy, precision, recall, and F-score. The second part shows the performance of classifiers in tabular form through the same measures.

4.3. Comparative Analysis of Applied Techniques

A comparative analysis was performed to analyze the performance of different classifiers evaluated during this study. To analyze the performance, the comparison of the evaluation results of single classifiers, ensemble-based classifiers, and fusion-based ensemble models is presented in this section. First, this section presents the comparison between single-based models and ensemble-based models. Secondly, a comparison is performed between fusion ensemble-based models.

The experimental results shown in Figure 7 depict that all of the ensemble-based models outperformed concerning all measures including accuracy, recall, precision, and F-score in comparison with single-based classifiers. Figure 7 shows the performance of models through bar charts and numeric terms.

The purpose of this experiment was to identify the high-performing fusion model by comparing fusion-based models. The results showed that fusion-based models improve the precision and accuracy of the student prediction model.

Figure 8 presents the performance of classifiers using a bar chart and table of numeric terms. The results shown in the above figure indicate that MB and MLP performed very well in terms of all measures. The fusion model MB + MLP achieved 98.7 accuracy and 98.6% precision, recall, and F-score which is higher than those for all other fusion models. The rest of the fusion models also showed relatively good performance. Overall, the results showed that fusion-based models improve the accuracy and precision of student predictions in comparison to single- and ensemble-based models. For effective communication, it is to be considered that the false positive rate, also known as the rate of sensitivity, should be nearly zero. The sensitivity rate for the MB + MLP model can be seen very close to zero, which strengthens the performance of the model.

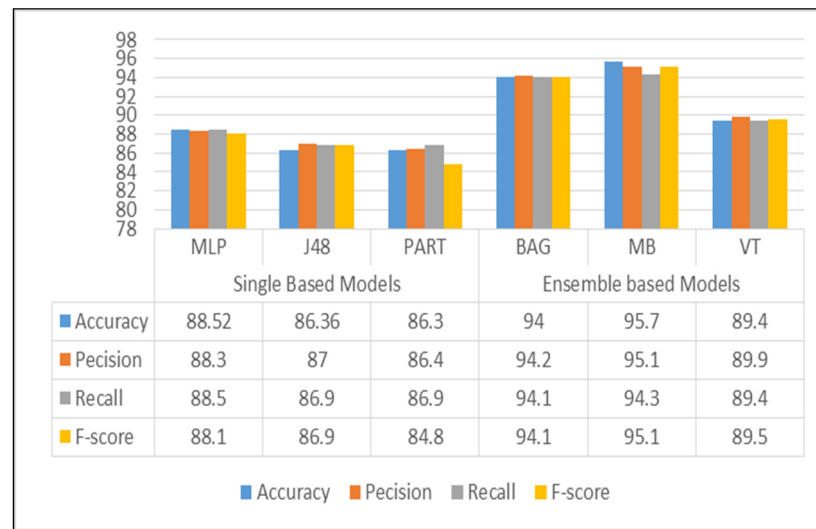


Figure 7. Comparison of single and ensemble-based classifiers.

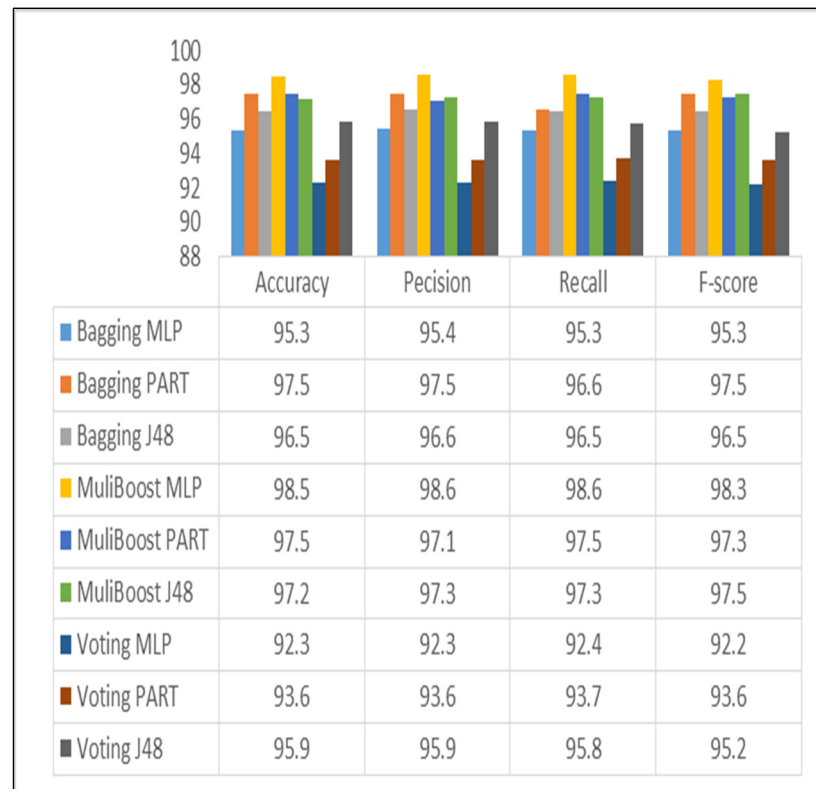


Figure 8. Comparison of fusion-based ensemble classifiers.

The F-measure of fusion-based MultiBoostAB and MLP has produced an average of 0.987 score that is a very significant result. It gives an average weight of the true positive as 0.988 and false positive rate as 0.004, as shown in Table 3.

As an average, the overall accuracy rate of the proposed model is evaluated as 98.7%. It shows that fusion-based ensemble models provide better precise outcomes, while classifying instances is very helpful in the evaluation of students' performance.

Table 3. Result obtained for MB + MLP.

Correctly Classified Instances		1185	98.7 %		
Incorrectly Classified Instances		15	1.24 %		
TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.997	0.008	0.976	0.997	0.986	A+
0.996	0.001	0.996	0.996	0.996	B
0.983	0.001	0.994	0.983	0.988	B+
0.994	0.007	0.983	0.994	0.989	A
0.969	0.000	1.000	0.969	0.984	C
0.909	0.000	1.000	0.909	0.952	F
0.943	0.000	1.000	0.988	0.971	D
Weighted Average					
0.988	0.004	0.988	0.986	0.987	

Figure 9 below shows the fusion-based ensemble model of ninth-class grade division based on the performance of MB + MLP and the class column represents the final performance prediction grades of the ninth class. In machine learning, sensitivity or recall is termed as the true positive rate and is used to measure the percentages of actual positives, which are identified correctly. The sensitivity rate should be nearer to one in order to obtain the true positive values.

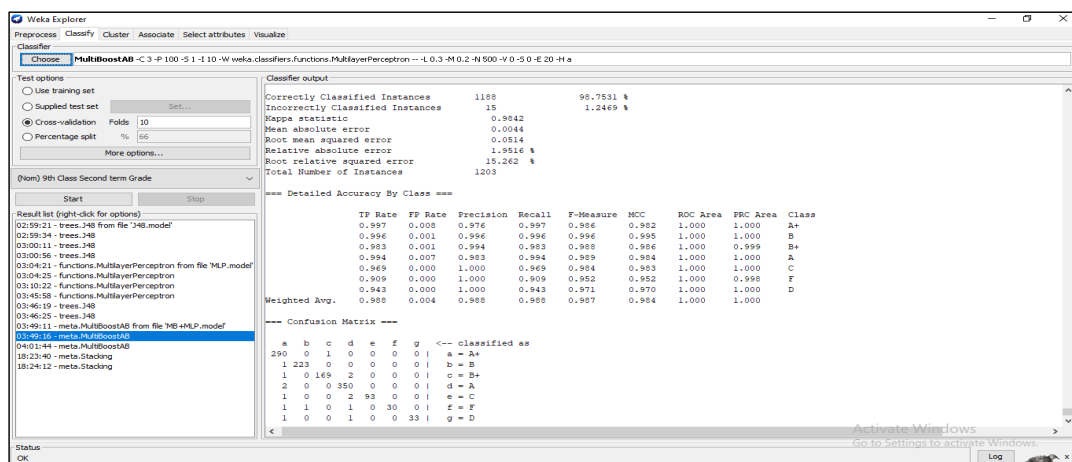


Figure 9. Fusion-based ensemble models (MB + MLP) result.

As shown in the above figure, the A+ grade has a true positive value of 290, whereas the false negative and false positive values are 7 and 1. The A grade has a true positive value of 350 with 8 false negative and positive values. Similarly, the values including 169, 223, 93, 33 and 30 are the true positive values of grades B+, B, C, D, and F.

The false-positive ratio (FPR) is the proportion of incorrectly classified negative instances. It has been discovered that a good model should have a false positive rate nearer to 0.0, which indicates less incorrectly classified negative instances.

A confusion matrix is a technique in which the performance of a classifier is summarized. From the confusion matrix, it was derived that, amongst the students from the dataset, 24.1% students secure an A+ grade, 18.5% students achieve a B grade, 14.0%, 29.1%, 7.7%, 2.4% and 2.7% students secure B+, A, C, F, D grades, respectively. This distribution obtains more insight from the model analysis.

Statistical hypothesis testing is used to make a claim related to the distribution of data or whether a set of results vary from one another. The null hypothesis is essential in interpreting the results and ensuring the strength of the claim of model performance by some statistical analysis. An Analysis of Variance test (ANOVA) is used in this study where statistical hypothesis testing is performed by estimating the p -value which is used to interpret the results of a test to either reject or fail to reject the null hypothesis. The p -value is selected as 0.05 for this study. The activity is conducted by comparing the p -value to the pre-chosen threshold value called alpha.

If the p -value < alpha, the null hypothesis would be rejected or we would fail to reject the null hypothesis in the p -value > alpha case.

The issue is addressed as whether the single and ensemble-based classifiers' performances are similar to fusion-based ensemble classifiers or not. The performances are evaluated by the four performance evaluation metrics (accuracy, precision, recall and F-measure) with the use of a collected dataset.

After testing, it was revealed that the four performance evaluation metrics are relatively higher for fusion-based ensemble classifiers models. By using the one-way ANOVA test, the p -value turned out to be 0.045 for single-based classifiers and 0.002 for ensemble-based classifiers, which is less than the significance value of alpha. Hence, the null hypothesis was rejected and the claim about the efficiency of the fusion-based ensemble model was strengthened.

4.4. Comparison of Applied Approach with Existing Approaches

Numerous research studies in EDM have been conducted using multiple classifiers to predict the performance of students. Sakri et al. have recently proposed an ensemble model to identify at-risk students and advise them to regulate their learning. They hybridized four single classifiers with four ensemble algorithms including bagging, random subspace, multilayer perceptron, and random forest. The evaluation results showed that the ensemble model achieved 91.70% accuracy, 86.1%, and the F-score was 87.3% [26]. Another study identified at-risk students, predicting their learning performance through their learning behavior based on their logging data history. They used Logistic Regression along with Random Forest, Multilayer Perceptron, and Gaussian Naive Bayes. The results showed that Random Forest surpasses the baseline Logistic Regression and other models with 89% accuracy, 89% precision, 88% recall, and 88% F1 score [56]. Emmanuel et al. introduced a model to predict students' success based on their daily activities. They hybridized different single classifiers with bagging, boosting, and random forest. The experiment results showed that the model achieved 96.9% accuracy [57]. Another recent study predicted students' intermediate results based on their academic characteristics. The evaluation results indicated that the model attained 96.64% accuracy [58]. A model has been suggested to estimate the institutional performance based on key performance indicators using data mining techniques. The results showed that the artificial neural networks performed better in achieving accuracy (i.e., 82.9%) in comparison to other machine learning models employed in the study [59]. The student performance in a learning management system based on behavioral features was predicted by applying ensemble methods including bagging, boosting, and random forest to augment the performance of classifiers. An accuracy of 91.5% was achieved through the application of ensemble methods to the classifiers to enhance academic performance [60]. Ragab et al. introduced a data mining-based forecast model to determine students' accomplishments. The data mining techniques used to evaluate students' performance include a decision tree, logistic regression, a naïve Bayes tree, an artificial neural network, a support vector machine, and a k-nearest neighbor. To improve the productivity of these classifiers, they used ensemble methods such as bagging, boosting, random forest, and voting. The results showed that the decision tree algorithm accuracy increased with bagging from 90.4% to 91.4%. Similarly, recall results were increased from 0.904 to 0.914, and precision results were also improved from 0.905 to 0.914 [61]. Another study undertook the task of student performance prediction by extracting features from

an e-learning system. The proposed model comprises five traditional machine learning algorithms which were complemented by four well-established ensemble techniques including bagging, boosting, stacking, and voting. The F1 score measured by the NB model by the integration of boosting and GBT with AdaBoost were 0.71% and 0.75%, respectively [62]. Adejo et al.'s research focused on the prediction of students' performance using data mining techniques along with the support of ensemble methods. They also proposed novel hybrid classifiers to gain accurate predictions of student performance. The results showed that the hybrid model outperformed the other classifiers in terms of accuracy (i.e., 81.67) precision (i.e., 79.62), recall (i.e., 75.86), and F-score (i.e., 77.69) in comparison to base classifiers and ensemble techniques applied in the same research [54].

The fusion ensemble-based approach introduced in this research study to improve academic performance attained the highest accuracy (i.e., 98.7%) precision (i.e., 98.6%) recall (i.e., 98.6%) and F-score (i.e., 98.6) in comparison to state-of-the-art ensemble approaches proposed in EDM. Thus, the result acquired in this study demonstrates the reliability of the proposed predictive model. The performance of our approach represents improvements in terms of all measures relative to existing approaches which emphasize that the fusion of ensemble techniques can improve the fraction of prediction.

5. Limitations

This study is limited to predicting the performance of students studying in a physical learning environment and lacks envisagement of the performance of students in online learning. The data set should be extended to evaluate the performance of students in both physical and virtual educational settings. The proposed system is deficient in suggesting apposite learning streams to the students based on their learning performance to pursue further educational goals. Another limitation of this study is that the factors are identified only concerning a specific educational level rather than providing a framework of factors for all educational levels including elementary, secondary, higher secondary, and tertiary.

6. Conclusions and Future Work

Numerous DM techniques have been implemented in academia as a standard procedure for interpreting the bulk of students' data and then mining them into one meaningful datum and knowledge to support decision-making processes. An early performance prediction would be beneficial for at-risk students, facing difficulty in attaining good grades in the class. To support such students in their learning to improve their progress, it is important to periodically predict their performance so they can be supported. A robust fusion-based ensemble model was developed considering students' demographic, family, social and academic attributes to make predictions. The model is highly useful in assessing students at early stages. After building many models involving single, ensemble, and fusion-based ensemble classifiers, MultiBoostAB ensemble classifier with MLP base appears as the best model to predict students' performance at the lower secondary level.

A logical extension of this research would be the building of a meta-analysis system on a larger dataset for future study which can be considered as a decision support method based on the model that will achieve the highest efficiency and effectiveness.

Furthermore, the study can be enhanced by the use of hybrid feature selection methods to help predict student performances, so each feature becomes more optimal and significant in terms of student performance prediction. The advanced ensemble-based machine learning algorithm, in particular extreme gradient boosting, could also be used in this domain.

Author Contributions: Conceptualization, A.S.; Data curation, A.J. and A.S.; Formal analysis, A.J. and M.O.A.W.; Funding acquisition, N.N.Q.; Investigation, A.J.; Methodology, F.M.; Project administration, N.N.Q. and M.O.A.W.; Resources, A.I.Q., N.N.Q. and M.O.A.W.; Software, A.J.; Supervision, A.S. and F.M.; Validation, F.M. and A.I.Q.; Visualization, A.I.Q.; Writing—original draft, A.S.; Writing—review & editing, A.S. and F.M. All authors have read and agreed to the published version of the manuscript.

Funding: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding and support for this work under Research under grant number R.G.P.1/184/42.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The current study dataset is publicly available online (<https://www.kaggle.com/asiyajan001.student-performance-perdiction>, accessed on 8 December 2021) for research purposes.

Acknowledgments: Authors thankfully acknowledge all those who contributed this work scientifically, administratively, financially and academically during the whole research process.

Conflicts of Interest: The authors declare that they have no conflict of interest to report regarding the present study.

References

1. Kamran, S.; Nawaz, I.; Aslam, S.; Zaheer, S.; Shaikat, U. Student's performance in the context of data mining. In Proceedings of the 2016 19th International Multi-Topic Conference (INMIC), Islamabad, Pakistan, 5–6 December 2016; pp. 1–8.
2. Kamran, S.; Nawaz, I.; Aslam, S.; Zaheer, S.; Shaikat, U. *Student's Performance: A Data Mining Perspective*; LAP Lambert Academic Publishing: Koln, Germany, 2017.
3. Iqbal, M.S.; Luo, B. Prediction of educational institution using predictive analytic techniques. *Educ. Inf. Technol.* **2018**, *24*, 1469–1483. [[CrossRef](#)]
4. Kaur, A.; Umesh, N.; Singh, B. Machine Learning Approach to Predict Student Academic Performance. 2018. Available online: www.ijraset.com/734 (accessed on 16 July 2021).
5. Aslam, M.; Malik, R.; Rawal, S.; Rose, P.; Vignoles, A. Do government schools improve learning for poor students? Evidence from rural Pakistan. *Oxf. Rev. Educ.* **2019**, *45*, 802–824. [[CrossRef](#)]
6. Abid, A.; Kallel, I.; Blanco, I.; Benayed, M. Selecting relevant educational attributes for predicting students' academic performance. In *Intelligent Systems Design and Applications, Proceedings of the 17th International Conference on Intelligent Systems Design and Applications (ISDA 2017), Delhi, India, 14–16 December 2017*; Springer: Cham, Switzerland, 2018; pp. 650–660. [[CrossRef](#)]
7. Shahrazad, H. Knowledge economy: Characteristics and dimensions. *Manag. Dyn. Knowl. Econ.* **2017**, *5*, 203–225.
8. Baneres, D.; Rodriguez-Gonzalez, M.E.; Serra, M. An Early Feedback Prediction System for Learners At-Risk within a First-Year Higher Education Course. *IEEE Trans. Learn. Technol.* **2019**, *12*, 249–263. [[CrossRef](#)]
9. Imran, M.; Latif, S.; Mehmood, D.; Shah, M.S. Student Academic Performance Prediction using Supervised Learning Techniques. *Int. J. Emerg. Technol. Learn.* **2019**, *14*, 92–104. [[CrossRef](#)]
10. Phua, E.J.; Batcha, N.K. Comparative analysis of ensemble algorithms' prediction accuracies in education data mining. *J. Crit. Rev.* **2020**, *7*, 37–40.
11. Abu, A. Educational data mining & students' performance prediction. *Int. J. Adv. Comput. Sci. Appl.* **2016**, *7*, 212–220.
12. Romero, C.; Ventura, S. Educational Data Mining: A Review of the State of the Art. *IEEE Trans. Syst. Man Cybern. Part. C (Appl. Rev.)* **2010**, *40*, 601–618. [[CrossRef](#)]
13. Arun, D.K.; Namratha, V.; Ramyashree, B.V.; Jain, Y.P.; Choudhury, A.R. Student academic performance prediction using educational data mining. In Proceedings of the 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 27–29 January 2021; pp. 1–9.
14. Alturki, S.; Alturki, N. Using Educational Data Mining to Predict Students' Academic Performance for Applying Early Interventions. *J. Inf. Technol. Educ. Innov. Pract.* **2021**, *20*, 121–137. [[CrossRef](#)]
15. Trautwein, U.; Lüdtke, O.; Marsh, H.W.; Köller, O.; Baumert, J. Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *J. Educ. Psychol.* **2006**, *98*, 788–806. [[CrossRef](#)]
16. Li, F.; Zhang, Y.; Chen, M.; Gao, K. Which Factors Have the Greatest Impact on Student's Performance. *J. Phys. Conf. Ser.* **2019**, *1288*, 012077. [[CrossRef](#)]
17. Francis, B.K.; Babu, S.S. Predicting Academic Performance of Students Using a Hybrid Data Mining Approach. *J. Med. Syst.* **2019**, *43*, 162. [[CrossRef](#)]
18. Md Zubair Rahman, A.M.J. Model of Tuned J48 Classification and Analysis of Performance Prediction in Educational Data Mining. 2018. Available online: <http://www.ripublication.com> (accessed on 5 July 2021).
19. Tomasevic, N.; Gvozdenovic, N.; Vranes, S. An overview and comparison of supervised data mining techniques for student exam performance prediction. *Comput. Educ.* **2019**, *143*, 103676. [[CrossRef](#)]
20. Aucejo, E.M.; French, J.; Ugalde Araya, M.P.; Zafar, B. The impact of COVID-19 on student experiences and expectations: Evidence from a survey. *J. Public Econ.* **2020**, *191*, 104271. [[CrossRef](#)] [[PubMed](#)]
21. Zollanvari, A.; Kizilirmak, R.C.; Kho, Y.H.; Hernandez-Torrano, D. Predicting Students' GPA and Developing Intervention Strategies Based on Self-Regulatory Learning Behaviors. *IEEE Access* **2017**, *5*, 23792–23802. [[CrossRef](#)]

22. Abu Amrieh, E.; Hamtini, T.; Aljarah, I. Mining Educational Data to Predict Student's academic Performance using Ensemble Methods. *Int. J. Database Theory Appl.* **2016**, *9*, 119–136. [[CrossRef](#)]
23. Hutt, S.; Gardener, M.; Kamentz, D.; Duckworth, A.L.; D'Mello, S.K. Prospectively predicting 4-year college graduation from student applications. In Proceedings of the LAK '18: International Conference on Learning Analytics and Knowledge, Sydney, Australia, 7–9 March 2018; pp. 280–289. [[CrossRef](#)]
24. Xu, J.; Moon, K.H.; van der Schaar, M. A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs. *IEEE J. Sel. Top. Signal. Process.* **2017**, *11*, 742–753. [[CrossRef](#)]
25. Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Chen, S.; Liu, D.; Li, J. Performance Comparison and Current Challenges of Using Machine Learning Techniques in Cybersecurity. *Energies* **2020**, *13*, 2509. [[CrossRef](#)]
26. Hassan, H.; Ahmad, N.B.; Anuar, S. Improved students' performance prediction for multi-class imbalanced problems using hybrid and ensemble approach in educational data mining. *J. Phys. Conf. Ser.* **2020**, *1529*, 052041. [[CrossRef](#)]
27. Wood, L.; Kiperman, S.; Esch, R.C.; Leroux, A.J.; Truscott, S.D. Predicting dropout using student- and school-level factors: An ecological perspective. *Sch. Psychol. Q.* **2017**, *32*, 35–49. [[CrossRef](#)]
28. Nahar, K.; Shova, B.I.; Ria, T.; Rashid, H.B.; Islam, A.H.M.S. Mining educational data to predict students performance. *Educ. Inf. Technol.* **2021**, *26*, 6051–6067. [[CrossRef](#)]
29. Madni, H.A.; Anwar, Z.; Shah, M.A. Data mining techniques and applications—A decade review. In Proceedings of the International Conference on Automation and Computing (ICAC), Huddersfield, UK, 7–8 September 2017; pp. 1–7.
30. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA data mining software: An update. *ACM SIGKDD Explor. Newsl.* **2009**, *11*, 10–18. [[CrossRef](#)]
31. Cenitta, D.; Arjunan, R.V.; Prema, K.V. Missing data imputation using machine learning algorithm for supervised learning. In Proceedings of the 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 27–29 January 2021; pp. 1–5.
32. Alam, T.M.; Shaukat, K.; Hameed, I.A.; Luo, S.; Sarwar, M.U.; Shabbir, S.; Li, J.; Khushi, M. An Investigation of Credit Card Default Prediction in the Imbalanced Datasets. *IEEE Access* **2020**, *8*, 201173–201198. [[CrossRef](#)]
33. Kanchan, J.; Saha, S. Incorporation of multimodal multi objective optimization in designing a filter based feature selection technique. *Appl. Soft Comput.* **2021**, *98*, 106823.
34. Krishnan, N.; Karthikeyan, M. IEEE signal processing/computational intelligence/computer joint societies chapter. In Proceedings of the 2014 IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, India, 18–20 December 2014.
35. Salloum, S.A.; Alshurideh, M.; Elnagar, A.; Shaalan, K. Mining in educational data: Review and future directions. In Proceedings of the Joint European—US Workshop on Applications of Invariance in Computer Vision, Ponta Delgada, Portugal, 9–14 October 2020; pp. 61–70.
36. Sakri, S.; Alluhaidan, A.S. RHEM: A robust hybrid ensemble model for students' performance assessment on cloud computing course. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 388–396. [[CrossRef](#)]
37. Musiliu, B. Single Classifiers and Ensemble Approach for Predicting Student's Academic Performance. 2020. Available online: www.rsisinternational.org (accessed on 10 August 2021).
38. Ali, S.; Smith, K.A. On learning algorithm selection for classification. *Appl. Soft Comput. J.* **2006**, *6*, 119–138. [[CrossRef](#)]
39. Rokach, L. Taxonomy for characterizing ensemble methods in classification tasks: A review and annotated bibliography. *Comput. Stat. Data Anal.* **2009**, *53*, 4046–4072. [[CrossRef](#)]
40. Livieris, I.E.; Iliadis, L.; Pintelas, P. On ensemble techniques of weight-constrained neural networks. *Evol. Syst.* **2021**, *12*, 155–167. [[CrossRef](#)]
41. Ahmad, I.; Yousaf, M.; Yousaf, S.; Ahmad, M.O. Fake News Detection Using Machine Learning Ensemble Methods. *Complexity* **2020**, *2020*, 8885861. [[CrossRef](#)]
42. Yang, X.; Lo, D.; Xia, X.; Sun, J. TLEL: A two-layer ensemble learning approach for just-in-time defect prediction. *Inf. Softw. Technol.* **2017**, *87*, 206–220. [[CrossRef](#)]
43. Riestra-González, M.; del Puerto Paule-Ruiz, M.; Ortin, F. Massive LMS log data analysis for the early prediction of course-agnostic student performance. *Comput. Educ.* **2021**, *163*, 104108. [[CrossRef](#)]
44. Hoque, I.; Azad, A.K.; Tuhin, M.A.H.; Salehin, Z.U. University Students Result Analysis and Prediction System by Decision Tree Algorithm. *Adv. Sci. Technol. Eng. Syst. J.* **2020**, *5*, 115–122. [[CrossRef](#)]
45. Panigrahi, R.; Borah, S. Rank Allocation to J48 Group of Decision Tree Classifiers using Binary and Multiclass Intrusion Detection Datasets. *Procedia Comput. Sci.* **2018**, *132*, 323–332. [[CrossRef](#)]
46. Bauer, E.; Kohavi, R. An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Mach. Learn.* **1999**, *36*, 105–139. [[CrossRef](#)]
47. Li, X.; Zhang, Y.; Cheng, H.; Zhou, F.; Yin, B. An Unsupervised Ensemble Clustering Approach for the Analysis of Student Behavioral Patterns. *IEEE Access* **2021**, *9*, 7076–7091. [[CrossRef](#)]
48. Ashraf, M.; Zaman, M.; Ahmed, M. An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches. *Procedia Comput. Sci.* **2020**, *167*, 1471–1483. [[CrossRef](#)]
49. Shaukat, K.; Luo, S.; Varadharajan, V.; Hameed, I.A.; Xu, M. A Survey on Machine Learning Techniques for Cyber Security in the Last Decade. *IEEE Access* **2020**, *8*, 222310–222354. [[CrossRef](#)]

50. Sun, Y.; Li, Z.; Li, X.; Zhang, J. Classifier Selection and Ensemble Model for Multi-class Imbalance Learning in Education Grants Prediction. *Appl. Artif. Intell.* **2021**, *35*, 290–303. [[CrossRef](#)]
51. Schapire, R.E. A brief introduction to boosting. *Ijcai* **1999**, *99*, 1401–1406.
52. Shaukat, K.; Masood, N.; Mehreen, S.; Azmeen, U. Dengue Fever Prediction: A Data Mining Problem. *J. Data Min. Genom. Proteom.* **2015**, *6*, 3. [[CrossRef](#)]
53. Adejo, O.W.; Connolly, T. Predicting student academic performance using multi-model heterogeneous ensemble approach. *J. Appl. Res. High. Educ.* **2018**, *10*, 61–75. [[CrossRef](#)]
54. Dutta, S.; Bandyopadhyay, S.K. Forecasting of Campus Placement for Students Using Ensemble Voting Classifier. *Asian J. Res. Comput. Sci.* **2020**, 1–12. [[CrossRef](#)]
55. Alabi, E.O.; Adeniji, O.D.; Awoyelu, T.M.; Fasae, O.D. Hybridization of Machine Learning Techniques in Predicting Mental Disorder. *Int. J. Hum. Comput. Stud.* **2021**, *3*, 22–30.
56. Wasif, M.; Waheed, H.; Aljohani, N.R.; Hassan, S.-U. Understanding student learning behavior and predicting their performance. In *Cognitive Computing in Technology-Enhanced Learning*; IGI Global: Hershey, PA, USA, 2019; pp. 1–28. [[CrossRef](#)]
57. Emmanuel, A.A.; Aderoju, M.A.; Falade, A.A.F.; Atanda, A. An appraisal of online gambling on undergraduate students' academic performance in university of Ilorin, Nigeria. *Int. J. Innov. Technol. Integr. Educ.* **2019**, *3*, 45–54.
58. Yousafzai, B.K.; Hayat, M.; Afzal, S. Application of machine learning and data mining in predicting the performance of intermediate and secondary education level student. *Educ. Inf. Technol.* **2020**, *25*, 4677–4697. [[CrossRef](#)]
59. Alam, T.M.; Mushtaq, M.; Shaukat, K.; Hameed, I.A.; Sarwar, M.U.; Luo, S. A Novel Method for Performance Measurement of Public Educational Institutions Using Machine Learning Models. *Appl. Sci.* **2021**, *11*, 9296. [[CrossRef](#)]
60. Ajibade, S.-S.M.; Ahmad, N.B.B.; Shamsuddin, S.M. Educational Data Mining: Enhancement of Student Performance model using Ensemble Methods. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *551*, 012061. [[CrossRef](#)]
61. Ragab, M.; Aal, A.M.K.A.; Jifri, A.O.; Omran, N.F. Enhancement of Predicting Students Performance Model Using Ensemble Approaches and Educational Data Mining Techniques. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 6241676. [[CrossRef](#)]
62. Saleem, F.; Ullah, Z.; Fakieh, B.; Kateb, F. Intelligent Decision Support System for Predicting Student's E-Learning Performance Using Ensemble Machine Learning. *Mathematics* **2021**, *9*, 2078. [[CrossRef](#)]