Review Article

PREDICTING STUDENTS' PERFORMANCE USING MACHINE LEARNING ALGORITHMS: A REVIEW

Abstract

Educational Data Mining is a discipline focused on developing ways for studying the unique and increasingly large-scale data generated by educational settings and applying those methods to better understand students and the environments in which they learn. Predicting student performance is one of the most critical concerns in educational data mining, which is gaining popularity. Student performance prediction attempts to forecast a student's grade before enrolling in a course or completing an exam. The goal of this paper is to present a systematic literature review on predicting student performance using machine learning techniques and how the prediction algorithm can be used to identify the most important attribute(s) in a student's data.

Introduction

Data mining or Knowledge discovery in data (KDD) is the science of studying patterns and other essential information from large data sets. It is also the science of discovering useful patterns in large datasets[1]. There exist many data mining applications and techniques that can be utilized and they include clustering, prediction, relationship mining, outlier detections, social media analysis, text mining, process mining, data distillation for judgment[2], classification, association rules [3], correlation analysis, decision trees, regression trees, Markov chains, differential sequence mining, sequential patterns, Bayesian networks, neural networks, linear regression [4], prediction[5], [6], and machine learning models [7], [8].

Currently, one of the popular fields of interest is Educational Data Mining (EDM) [9]. Using data mining techniques to analyze and extract hidden knowledge from educational data context is what s termed Educational Data Mining (EDM). An example of this is that discovered knowledge can assist education instructors to enhance instructional techniques, improve the learning process, understand learners, and it could also be adopted by learners to enhance their learning activities. It can also help the administrator make correct decisions to produce high-quality outcomes. Different sources can be used to collect educational data. Some of these sources include traditional surveys, web-based education, and educational repositories. EDM is able to utilize different DM techniques. Each of these techniques can be employed for specific educational problems [10]. The educational data needed can be collected from different sources such as e-learning or web-based learning, traditional education, learning management systems, adaptive educational systems, test questionnaires, and text contents [11]. In EDM, the most common method used is prediction techniques that develop a platform that can guess/predict a single likely outcome of the data from a combination of other outcomes; structure discovery algorithms which attempt to find structure in the data without any ground truth or a prior idea of what should be found and relationship mining methods that discover relationships between variables in a dataset with a large number of variables. Educational data mining applications can be used widely. A vast use of educational mining is in enhancing the studying process, improving course organization, supporting students in course selection, finding problems leading to dropping out, and as a support for decision-making at student admissions [12].

Literature Review

One of the main topics of educational data mining is Predicting students' academic performance. With technology advancing, so has there been an increase in technological investments in the field of education. E-Learning platforms such as Web-based online learning and multimedia have evolved along with technological developments. Both have impacted the cost of learning, increasing it while eliminating the challenges of time and space limitations. Digital data has increased tremendously due to the increase of online course trainings and the increase of online transactions and interactive transactions in schools.

Student's success prediction aids educational institutions to improve teaching and learning methodologies by identifying instructional methods that suit students from varied background information [13]. The student's performance prediction is an essential area as it can help teachers identify students that need additional academic assistance [14]. Instructors attain an understanding into how well or poorly students may perform and necessary proactive measures can be adopted to improve teaching and learning when student's performance are predicted. Accurately predicting students' future performance based on their ongoing academic records is crucial for effectively carrying out necessary pedagogical interventions to ensure students' on-time and satisfactory course completion [15]. Machine learning techniques can be uutilized to predict the output of the students and identifying the at risk students as early as possible so appropriate actions can be taken to enhance their performance. Choosing the attributes or the descriptive features used as input to the machine learning algorithm is one of the most important steps when using these techniques. The attributes can be classified into grades and GPA, demographics, psychological profile, cultural, academic progress, and educational background.

Related Works

Waheed et al [16] deployed a deep artificial neural network on the OULA dataset to forecast at-risk learners and provide intervention measures in early stages in such cases. The outcome showed the suggested model achieved a categorization efficiency of 84%-93%. which is far beyond the 79.82% - 85.60% baseline logistic regression, and 79.95% - 89.14% support vector machine model. Output shows demographic characteristics and student's clickstream activity, after the application of the module and the impact on student performance. Zorić [12] performed a study on forecasting academic performance of students based on enrolment data using neural network. He randomly divided the dataset into three sets: training (68.75%), validation (15.63%) and testing(15.63 %). After training the network, Mean Correct Classification Rate was 93,421053 %. Different neural network topologies were tested (different number of hidden layers, different algorithms and parameters), and the results obtained from all was almost the same.

Umer [17] proposed a process mining approach to help in making early predictions to improve students' learning experience in massive open online courses (MOOCs). The research evaluated four machine learning classification techniques (Naïve Bayes (NB), logistic regression (LR), K-nearest neighbour and random forest (RF)) to observe weekly progression and predict their overall performance outcome. The Naive Bayes method gave an output better than all the other methods with an efficiency of 89%. Rincon-Flores et al [18] utilized two algorithms (Random Forest and K-Nearest Neighbor) to monitor a predictive model for students pursuing engineering's academic performance. Mean Absolute Percentage Error (MAPE) and Mean Absolute Deviation (MAD) were applied to set error parameters. Random Forest predicted an average of 16.4% error.

Musso et al [19] used background information of students, together with their cognitive and non-cognitive measures to construct predictive models for monitoring performance using artificial neural networks (ANN). Three neural network models were constructed with two (2) of the models identifying the top 33% and the lowest 33% groups respectively. were able to reach 100% correct identification of all students in each of the two groups. The third model (identifying low, mid and high -performance levels) reached precisions from 87% to 100% for the three groups. Amrieh et al [20] utilised Artificial Neural Network, Naïve Bayesian and Decision tree to predict student's performance and also applied ensemble methods (Bagging, Boosting and Random Forest (RF)) to enhance the output. Results obtained made apparent the fact that there is a strong relationship between learner's behaviours and their academic achievement. There was a of 25.8% improvement record with the application of the predictive model using behavioural features as compared to a 22.1% when such features were removed.

Phua and Batcha [21] developed a model to predict students' grades of modules from past results. The study used Linear Regression, K-Nearest Neighbor and Decision Table and ensemble algorithms (Stacking, Bagging and Random Forest). Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) were used to benchmark of each algorithm. The Decision Table performed the best with the lowest MAE of 8.2194 for the small dataset and the Stacking performed the best with the lowest overall MAE of 7.5989 for large dataset. Hamoud and Humadi [22] presented a prediction model based on an artificial neural network (ANN) by adopting feature selection (FS). Based on the FS algorithms, out of the first 30 academic factors,14 influenced the final outcome (Grade). The model indicated an efficiency of 87% after dividing the data set into 70% for the model training and 30% for data testing.

Affendey et al [23] used WEKA open source data mining tool to determine the key courses that potentially affect the overall academic performance within the Bachelor of Computer Science program. In evaluating the prediction accuracy, a cross-validation with 10 folds was used. The outcome indicated that Naive Bayes algorithm provided the highest percentage score of 95.29%.

Collaborative Filtering (CF), Matrix Factorization (MF), and Restricted Boltzmann Machines (RBM) techniques were used by Iqbal et al to assess the academic performance. Empirical validation showed the effectiveness of RBM technique. Pereira et al [24] deployed a deep learning method for early performance prediction of introductory programming students and compared to an evolutionary algorithm (Pereira et al. (2019). The deep learning model achieved an efficiency range of 81.9% to 82.7% whereas the evolutionary algorithm attained an efficiency range of 77.8% to 78.6%. These results were obtained using features extracted from the same dataset, which was collated within the very first two weeks of introductory programming courses to give way for early intervention. A performance comparison adopting Linear support vector machines (LSVM) with ten other algorithms (coarse decision tree, medium decision tree, fine decision tree, logistic regression, Gaussian Naïve Bayes, Kernel Naive Bayes, quadratic SVM, cubic SVM, fine Gaussian SVM, and medium Gaussian SVM) in predicting student's performance was conducted by Naicker et al. the outcome was thus: LSVM produced the highest accuracy of 90.1% with the Gaussian Naïve Bayes recording the lowest accuracy 68.4%.

Saheed et al [25] presented a method to forecast students' performance utilizing Classification and Regression tree (CART), Iterative dichotomiser 3 (ID3), C4.5 algorithms. Results obtained from the experiment indicated that ID3 recorded an accuracy of 95.9%, C4.5 gave an accuracy of 98.3% and CART results showed an accuracy of 98.3%. The time taken to build the model of each was: ID3, 0.05 seconds, C4.5, 0.03 seconds and CART, 0.58 seconds. C4.5 outperforms other classifiers and requires reasonable amount of time to build the model as indicated by the results of this experiment. Ünal [26] proposed a wrapper method for feature subset selection before decision tree, random forest, and naive Bayes to forecast the final grades of students on two educational datasets related to mathematics lesson and Portuguese language lesson. The best results for five-level grading version without wrapper feature selection was decision tree (J48) algorithm with an accuracy rate of 73.42% whiles with feature selection was obtained with random forest 78.99%. The best results were obtained with random forest in binary label dataset with accuracy rate increased from 91.39 to 93.67% after feature selection.

Bithari et al [9] sought to forecast the academic performance of engineering students using their previous educational records, family background, demographic factors and other related factors. A predictive model is firstly built using the traditional classifiers Decision Tree, SVM, and Linear Regression, and an ensemble method, voting. The Ensemble Voting obtained an accuracy of 82, Decision Tree 74, SVM 78 and Linear Regression 72. The result obtained shows significant improvement in the performance when the ensemble method was implemented. Vairachilai and Vamshidharreddy [27] implemented algorithms(decision tree, Support Vector Machine (SVM), and Naive Bayes) of machine learning classification to forecast the academic performance of students. 71% accuracy was recorded by Decision tree provided, with Support Vector Machine provided and Naïve Bayes38% accuracy, and Naïve Bayes givds 77% accuracy. The Naïve Bayes algorithm gave the better result.

Imran et al [28] proposed an ensemble classification model that classifies a students' performance as Pass or Fail. Three supervised learning algorithms (J48, Non-Nested Generalization (NNge) and MLP) were employed in this study. The result shows that the proposed ensemble model Decision tree (J48) classifier achieved the high accuracy which is 95.78 % and lowest accuracy of 92.81% was achieved by NNge. Altabrawee et al [29] built a classifier using four machine learning techniques to predict computer science students' performance in one of their subjects. The techniques of machine learning used was made up of: Naïve Bayes, Artificial Neural Network, Logistics Regression and Decision Tree. The ANN model attained the best accuracy of all the classifiers that is equal to 77.04% as compared to Naïve Bayes model which had lowest accuracy that is of 66.52%

Soni et al [30] prepared a model which analysed the performance of pupils from their last output using Algorithms of classification such as: Naïve Bayes, Decision Tree, and Support Vector Machine for students' performance prediction. For the extraction process, twenty (20) out of the 48 features using the classifiers (NB, SVM and Decision Tree) were selected to analyse the influence of each feature for predicting the performance of students. Support vector machine with 83.33% accuracy was far better than was far better than decision tree and Naïve Bayes. Livieris et al [31] studied and analyzed the of two wrapper methods effectiveness for semi supervised learning algorithms for forecasting the performance of students in their final examinations. Two-wrapper methods for semi supervised (self-training and YATSI) learning and five classification algorithms; multilayer perceptron (MLP) naive Bayes (NB), sequential minimum optimization (SMO), C4.5and RIPPER algorithm were tested in their work. In conclusion, it was observed that semi supervised algorithms can improve the classification accuracy utilizing a few labelled and many unlabelled data for developing reliable prediction models.

Sekeroglu et al [32] focused on the forecasting and categorization of diverse types of educational data using machine learning algorithms in order to assess raw data efficiency without any data selection or preparation algorithm. Backpropagation (BP), Support Vector Regression (SVM) and Gradient Boosting Classifier (GBC) were used in testing the data. BP had the highest accuracy rate of 80.91% and 87.78% for 40% and 30% of testing ratio of instances respectively. Hashim et al [33] analysed the effectiveness of supervised machine learning algorithms in students' success prediction and academic performance in higher education using demographic, academic background and behavioural features. The performances of several supervised machine learning algorithms, such as Naïve Bayes, Decision Tree, Support Vector Machine, Logistic Regression, K-Nearest Neighbour, Sequential Minimal Optimisation and Neural Network. Research Results showed an accuracy of (68.7% for passed and 88.8% for failed) by logistic regression classifier as the most accurate in students' final grades of prediction with an AUC of 0.9541

Zohair [34] compared predicting performance of students' in any course with or without excluding the dissertation grade. Multiple machine learning classifiers were applied to train the datasets, including: Support Vector Machines (SVM), MLP-ANN, K Nearest Neighbour (KNN), Naïve Bayes (NB), and LDA. The highest accuracy was recorded by the LDA classifier with values 79% when dissertation grade was included and without dissertation grade, SVM recorded 69.7%. Okereke et al [35] asserted the necessity to adopt feature selection mechanism due to the high number of predicting variables determining student's performance. Decision tree was used in training and testing and achieved an accuracy of 92.27% and 70% on two different dataset. It was therefore concluded that the greatest factor in achieving higher accuracy is the type of datasets not actually the type of classification algorithm.

Sokkhey and Okazaki [36] introduced a principal component analysis (PCA) in conjunction with four machine learning (ML) as a hybrid approach. The four machine (ML) algorithms used were: random forest (RF), C5.0 of decision tree (DT), and naïve Bayes (NB) of Bayes network and support vector machine (SVM) with 10-fold cross-validation. to predict student performance. The Hybrid Random forest had the highest accuracy of 99.72% as compared to 89.2% without the hyprid approach. The proposed hybrid models produced very prediction results which shown itself as the optimal prediction and classification algorithms. Sungar et al [37] utilized CART and SVM to predict placement possibilities of a student. Using a MLP mapping function, CART provided the best results with an accuracy of 93.24%.

Agrawal and Mavani [38] used Multivariate Linear Regression to predict future marks of this student and obtained an accuracy Rate of 70.48%. Student's Secondary Education Grade, Living Location, Medium of Teaching were found to be of most important in the prediction. A comparison was made and it was concluded that in general, the neural networks tend to outperform Bayesian classification. Walia et al [39] built a predictive classification models using: Decision Tree, Naive-Bayes, ,JRip, Random-Forest, and ZeroR on student academic performance dataset. The JRip showed the best results with 81.77 % accuracy and using eight most significant characteristics, its accuracy increased to 85.06 %.

Sikder et al [40] predicted student's yearly Cumulative Grade Point Average (CGPA) performance in the form of using neural network and compare that with real CGPA. The result showed an accuracy of more than 90% and a RMSE is 0.1765. Chen and Do [41] investigated the neural network's prediction ability of neural networks by two heuristic algorithms, the cuckoo search(CS) and gravitational search algorithms(GSA). Based on the obtained results regarding prediction accuracy, the ANN–CS outperformed the ANN–GSA, and the results show that its prediction outcome is more accurate and reliable. Yanka [42] hypothesised that the failure or success on e-learn courses could be forecasted using data from activity logs. From the applied algorithms (Neural Net, Random Forest, XG Boost, Logistic Regression) the best accuracy performance was recorded by XG Boost and neural net with 0.9062. Number of actions taken, tasks from average and files uploaded from average are some of the significant variables identified by the models.

Cruz-Jesus et al [43] presented a novel approach to predict the academic achievement of virtually every public high school student in Portugal, i.e., 110,627 students in the academic year of 2014/2015. The RF, ANN, SVM, and LR were tested and the RF had the highest accuracy of 87%. Osmanbegovic and Suljic [44] compared diverse techniques and methods of mining data mining for students success prediction applying the collected data from the conducted surveys during the summer semester at the Faculty of Economics - University of Tuzla, academic year 2010-2011, among first year students and the data taken during the enrolment. The results indicated that the Naïve Bayes classifier outperforms decision tree and neural network methods with an accuracy of 76.65%. Polyzou and Karypis [45] explored if students' poor performance could be accurately predicted by projecting the challenge as binary classification and also, engineer a couple of human interpretable features that quantify these factors. The Gradient Boosting and Random Forest classifiers gave the best outcome with an Area under the ROC curve (AUC) of 0.854. Xu and Moon [46] developed a bi – layered model consisting of a number of base indicators and a cascade of outcome predictors for forecasting students' evolving performance. A method was developed to discover relevant courses for constructing base predictors. This method is the latent factor model-based course clustering method. An ensemble-based progressive prediction architecture was developed to incorporate students' evolving performance into the prediction.

Urkude and Gupta [47] adopted Decision Tree, support vector machines (SVM) and Naïve Bayes (NB) categorization algorithms to forecast the performance of students in an examination to ascertain whether a student will graduate or otherwise. The statistical analytical method (F1 score) was utilized to assess the adopted model's performance. Out of all the algorithms applied, Support vector machine with a F1 score of 0.7838 gave a better prediction. Iyanda [48] compared two neural network models (Generalized Regression Neural Network and Multilayer Perceptron) to ascertain which of them was the best model for forecasting the academic performance of students based on single performance factor. Generalized Regression Neural Network outperformed the Multilayer Perceptron model with an accuracy of 95% according to the study. This was then recommended to educationists to predict students' academic performance using single performance factor.

Vijayalakshmi and Venkatachalapathy [49] used Deep Neural Network to develop a student performance prediction system. The model was trained and tested with Kaggle dataset using different algorithms such as Decision Tree (C5.0), Naïve Bayes, Random Forest, Support Vector Machine, K-Nearest Neighbor and Deep neural network in R Programming. The Deep Neural Network outperformed the other classifiers with an accuracy of 84%. Rasheed et al [50] evaluated data collected from Electroencephalogram (EEG) using Machine Learning algorithm. The data was collected at age 4 and projected in predicting academic achievement at age 8 among rural children in Pakistan. K- Nearest Neighbor (KNN) classifier using 5 Stratified Folds was used on different model combinations of EEG, sociodemographic and home environment variables based on the sensitivity and specificity. Sensitivity was similar when EEG variables were combined with sociodemographic, and home environment (Math = 58.7%, Language = 66.3%) variables but specificity improved (Math = 43.4% to 50.6% and Language = 32% to 60%). Shanthini et al [51] projected the academic performance of University students based on their previous academic records instead of applying course dependent formulae to predict the final grades of the students. Meta decision tree classifier techniques based on four representative learning algorithms, namely Adaboost, Bagging, Dagging and Grading were used to construct different decision trees. Adaboost was observed to be the best classier of the meta decision model for predicting the student's result based on the marks obtained in the semester.

Findings

From the above review, it can be seen that neural networks was the most used algorithm followed by decision tree as presented in Table 1 and Figure 1. The Random Forest was the most ensemble algorithm used as compared to the rest. All neural network methods were classified as Deep learning methods

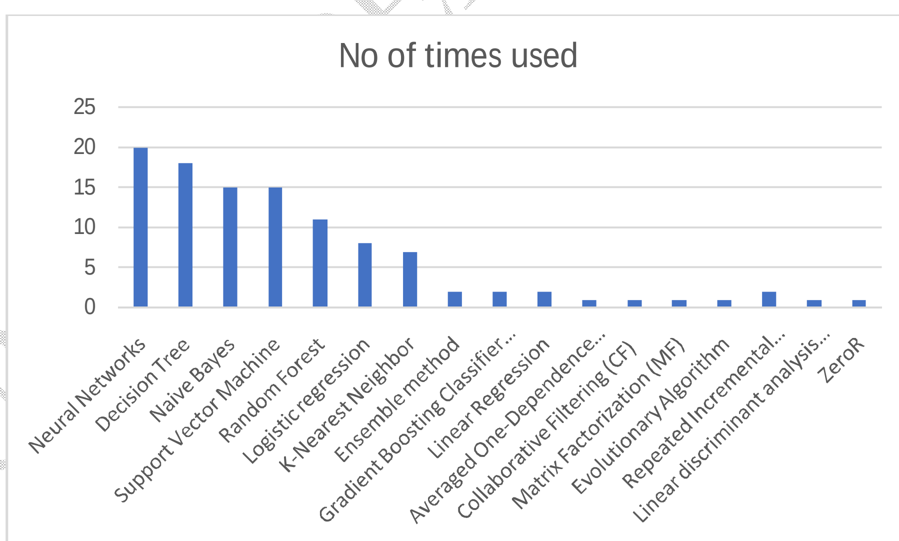

Figure 1: Frequency of Algorithms Used

Table 1: Summary of Papers Review

No	Algorithm	No of times used	Classification	References	**Best Performance**
1	Neural networks	20	Deep learning	[12], [16], [33], [34], [40]–[44], [48], [49], [52], [19], [20], [23], [24], [28], [29], [31], [32]	[12], [16], [24], [29], [32], [40], [42], [48], [49], [52]
2	Decision tree	18	Supervised learning	[8], [9], [33], [35], [36], [39], [44], [47], [49], [20], [25]–[31]	[25], [28], [35]
3	Naive Bayes	15	Supervised learning	[8], [17], [36], [39], [44], [47], [49], [20], [23], [26], [27], [29], [30], [33], [34]	[23], [27], [44]
4	support vector machine	15	Supervised learning	[8], [9], [37], [43], [47], [49], [16], [27], [30]–[34], [36]	[8], [30], [47]
5	Random Forest	11	Supervised learning	[17], [18], [49], [20], [26], [36], [39], [42], [43], [45], [46]	[26], [36], [43], [45]
6	Logistic regression	8	Supervised learning	[8], [16], [17], [29], [33], [42], [43], [46]	[33]
7	K-Nearest Neighbor	7	Supervised learning	[17], [18], [33], [34], [46], [49], [50]	
8	Ensemble method	2	Supervised learning	[9], [20], [42], [46], [51]	[9], [46], [51]
9	Gradient Boosting Classifier (GBC)	2	Supervised learning	[32], [45]	
10	Linear Regression	2	Supervised learning	[9], [38]	[38]
11	Averaged One-Dependence Estimators (AODE)	1	Supervised learning	[23]	
13	Collaborative Filtering (CF)	1	Unsupervised learning	[52]	
14	Matrix Factorization (MF)	1	Unsupervised learning	[52]	
16	Evolutionary Algorithm	1	N/A	[24]	
17	Repeated Incremental Pruning to Produce Error Reduction (RIPPER) algorithm	2	Supervised learning	[31], [39]	[39]
18	Linear discriminant analysis (LDA)	1	Supervised learning	[34]	[34]
20	ZeroR	1	Supervised learning	[39]	

The algorithms were also classified as either supervised learning or unsupervised. As presented in Figure 2, 87% of the algorithms used were supervised learning as compared to 13% for unsupervised learning. This shows in predicting students' performance, labelled data is used most of the time.

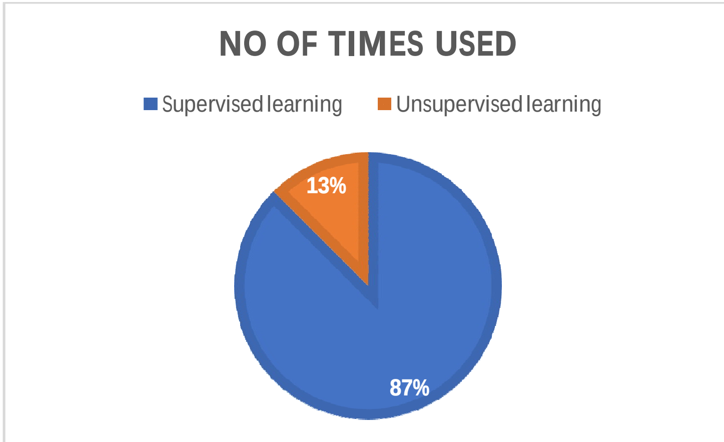

Figure 2: Frequency of Category of Algorithms Used

Feature selection is one of the effective and efficient way to classify relevant features and various studies has proven that features can be reduced without affecting performance. There is a high threat of dimensionality curse on the learning process of features when there is a great dimensionality of data leading to a huge computational time. The number of features in the dataset is directly proportional to both the time for training and the accuracy of the model. Thus, a large number of features will directly affect the training and performance of the model [53].

Employing feature selection techniques have the following advantages:

- The size of the dataset is reduced hence improving computational speed and also reducing storage size
- Since relevant features are selected, the number of features used is reduced which saves time and memory in subsequent tasks
- The general performance of learning tasks is increased hence improving the accuracy
- Features that adequately distinguish class labels are seen thereby understanding the dataset well

From the review, 59% of the studies employed various feature selection(ranking) methods to improve the performance of their model as shown in Figure 3.

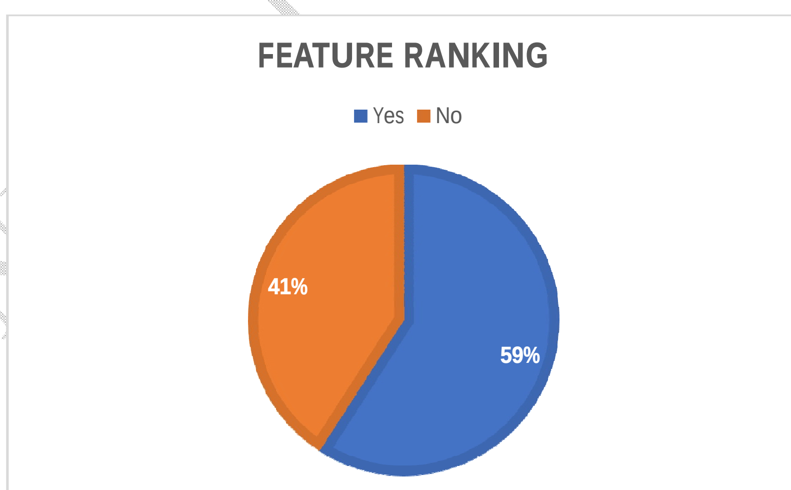

Figure 3: Feature Ranking

From Table 2, it can be seen that the attributes(features) used ranged from demographic factors, social factors, academic factors, and also online learning factors to the rise of e-learning as a result of the COVID pandemic. All these factors greatly influence the prediction of the academic performance of students.

Table 2: Features Used in Algorithms

Review	Algorithms	Attributes	Feature Ranking	Highest Accuracy
[16]	Deep Artificial Neural Network, Logistic Regression, And Support Vector Machine Models	Highest education, Index of Multiple Depravation band of the area, students age band, gender of the student, geographic region where student resided while taking that module, number of times a student has attempted a particular module, number of credits being currently studied by the student, if student has stated a disability, after course total clicks of a student for a particular module, before course total clicks of a student for a particular module, after course total clicks on the additional information such as videos, audios, sites etc., after course total clicks on the information on site and activity related to that information, after course total clicks on the files relevant to course, after course total clicks on the discussion forum, after course total clicks on the course homepage, after course total clicks on the online video discussions, after course total clicks on the contents of the assignment, after course total clicks on the Wikipedia content, after course total clicks on the information related to course, after course total clicks on the questionnaires related to course, after course total clicks on the course quiz, after course total clicks on the course contents from previous weeks, after course total clicks on the pdf resources such as books, after course total clicks on the other sites enabled in the course, after course total clicks on the links to audio/video contents, after course clicks on the shared information between courses and faculty, after course clicks on the external quiz activity, after course clicks on the online tutorial sessions, after course clicks on the interactive html page, after course clicks on the basic glossary related to contents of course, before course total clicks on the course homepage, before course clicks on the contents of the assignment, before course clicks on the subpage activity, before course clicks on the discussion forum, before course clicks on the links to audio/video contents, before course clicks on the information related to course, before course clicks on the Wikipedia content, before course clicks on the course quiz, before course clicks on the interactive html page, before course clicks on the ouelluminate activity, before course clicks on the basic glossary related to contents of course, before course clicks on the online video discussions, before course clicks on the external quiz activity, before course clicks on the questionnaires related to course, before course clicks on the additional information, before course clicks on the shared information between courses and faculty, before course clicks on the pdf resources such as books, before course clicks on the information on site and activity related to that information, number of clicks for a student, on the assessment day deadline submission, number of assessments in a module, number of assignments submitted late by a student, number of clicks for a student one day after the assessment deadline submission number of clicks for a student one day before the assessment deadline submission, total course activity clicks for each student for a particular	No	Deep Artificial Neural Network (84%-93%)

		module		
[12]	Neural Network	Status Of Study, Gender, Mother's Education, Father's Education, Average Grade During High School, State Matura Level - Croatian Language, State Matura Level - Mathematics, State Matura Level - Foreign Language, State Matura Grade - Croatian Language, State Matura Grade - Mathematics, State Matura Grade - Foreign Language, Average Monthly Income, Accommodation During Studies, Scholarship, Work During Studies	No	Neural Network (93.421053 %)
[17]	Naive Bayes, RF, LR, and KNN	Age, Education, Gender, Average score in the weekly quiz, Number of quizzes attempted, Quiz lag, Lecture lag, Total lecture attended, Video activity count, Efforts in seconds	Yes	
[18]	K-Nearest Neighbor and Random Forest	academic and biometric information from the participants, Biometric inputs (facial recognition and heartbeat), grades from different activities of the first period of evaluation	No	
[19]	Artificial Neural networks (ANN)	Gender, Secondary school from which the student graduated, Father's occupation, Mother's occupation, Highest level of education completed by mother, Highest level of education completed by father, Cognitive resources/Cognitive processing, Reaction Time Math, Time Management, Executive Control, Reaction Time Problem, Absolute Aospan (Sum of perfectly recalled sets), Anxiety Management, Alerting Attention,Reaction Time Operation,Orienting Attention, Study Techniques and use of help, Processing of information/ Generalization	Yes	
[20]	Artificial Neural Network, Naïve Bayesian and Decision tree, Ensemble methods (Bagging, Boosting and Random Forest (RF))	Nationality, Gender, Place of Birth, Parent responsible for the student, Educational Stages (school levels, Grade Levels, Section ID, Semester, Topic, Student Absence Days, Parent Answering Survey, Parent School Satisfaction, Discussion groups, Visited resources, Raised hand-on class, Viewing announcements	Yes	
[22]		Department, Age, Stage, Gender, Address, Status, Work, Live With Parent, Parent Alive, Father Work, Mother Work, Fail Courses, Absence Days, Credits, GPA, Completed Credits, Number of completed credits, Years of Study, List Important Points, Write Notes, Prep Study Schedule, Calm During Exam, low grades Not Make Me Fail		
[23]	Naive Bayes, AODE, and RBF Network classifiers		Yes	Naive Bayes (95.29%)
[52]	Collaborative Filtering (CF), Matrix Factorization (MF), and Restricted Boltzmann Machines (RBM)		No	Restricted Boltzmann Machines (RBM)
[24]	Deep Learning and Evolutionary	procrastination, amount of change, eventActivity, attempts, lloc, systemAccess,	Yes	Deep Learning

	Algorithm	firstExamGrade events, correctness, correctnessCodeAct, copyPaste, syntaxError, ideUsage, keystrokeLatency, errorQuotient, watWinScore, deleteAvg, countVar		(82.7%)
[8]	Linear Support Vector Machine, Coarse Decision Tree, Medium Decision Tree, Fine Decision Tree, Logistic Regression, Gaussian Naïve Bayes, Kernel Naive Bayes, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, And Medium Gaussian SVM	gender, race/ethnicity, parental level of education, access to lunch, test preparation, mathematics score, reading score, and writing score.	Yes	Linear support vector machines (90.1%)
[25]	Iterative dichotomised 3 (ID3), C4.5 and Classification and Regression tree (CART)	Age, Marital Status Religion, Sex, Nationality, Genotype, Fathers occupation, Mothers Occupation, Course Applied for Admission, Course Admitted for, Level, Mode of entry, Year of entry	No	C4.5 (98.3%)
[26]	Decision Tree, Random Forest, and Naive Bayes	Sex, Age, School, Address, Parental status, Mothers education, Mothers job, Father's education, Father's job, Guardian, Family size, Family relationship, Reason, Travel time from home to school, Study time a week, Failures, Extra educational school support, Family educational support, Extracurricular activities, Paid class, Internet access at home, Nursery Attended, Wants to take higher education, With a romantic relationship, Free time after school, Going out with friends, Alcohol consumption of weekend, Alcohol consumption of workday, Health Status, Absences, Grade of the first period, Grade of the second period, Grade of the final period	Yes	Random Forest (93.67%)
[9]	Decision Tree, SVM, Linear Regression, Ensemble Voting,	Gender, Scholarship Type, Age, Entrance Rank, Interest, Plus two percent, Plus two locations, Gap, Main subject, Class ten percentage, School Location, School Type, Ethnicity, Batch, Program, Father Job, Class (target)	No	Ensemble Voting (82%)
[27]	decision tree, Support Vector Machine (SVM), and Naive Bayes	Gender, Nationality, PlaceofBirth, StageID, GradeID, SectionID, Topic, Semester, Relation, Raisedhands, VisitedResources, AnnouncementsView	No	Naïve Bayes (77%)
[28]	Decision tree (J48), NNge and MLP		No	Decision tree (J48) (95.78 %)
[29]	Artificial Neural Network, Naïve Bayes, Decision Tree, and Logistic Regression.	Attribute, Department, Gender, Studying Style, Using Internet For Study, Using Extra Learning Resources, Interest in studying computer, Has Computer Experience, Studying Hours, Family Members Education, Family Help In Studying, Educational Environment Satisfaction, Has A Job, Accommodation, Residence, Married, Sport Participation, Time Spent On Social Media (Hours, Computer Grade-Course1, English Grade-Course1, Final Computer Outcome	No	ANN (77.04%)

[30]	Decision Tree, Naïve Bayes and Support Vector Machine	Attendance, Marks above 80%, Marks above 40%, Student interested in study, Student understanding of things, Memorizes answers, Write their own words, Has taken extra courses, Does research, Parents Alive, Literate Parents, Parents can read or write, Parents are employed, Net income less than 30000 per month, Net income more than 30000 but less than 60000 per month, Parents are strict, Parents are caring, Student is Emotional, Students tolerance for failures, Student gets anger, Student is patient, Student respects elder, Get in fight with friends, Has bad habit, Has bad habit from adolescence age, Got bad habit from friend, Got bad habit from family, Heavily addicted to bad habits, Everyday routine of bad habits, Sometime or weekly does bad habit, Started realising about bad habits, Want to get rid of bad habits, Gets pocket money, Spends money to buy needed items, Spends money to buy cigarettes and alcohol, Steals money, Has police complaints, Roam around with friends all day, Does social service, Interested in sports, Participates in extra curricular, Has interest in political science, Does part time job or internship, Aptitude practice, Coding practice, Group discussion practice, Personal interview practice	Yes	Support Vector Machine (83.33%)
[31]	Naive Bayes (NB), sequential minimum optimization (SMO), C4.5, multilayer perceptron (MLP), and RIPPER algorithm	Secondary stage, Oral grade of the first semester, Grade of the first test of the first semester, Grade of the second test of the first semester, Grade of the final examination of the first semester, Final grade of the first semester, Oral grade of the second semester, Grade of the first test of the second semester, Grade of the second test of the second semester, Grade of the final examination of the second semester, Final grade of the second semester, Grade in the final examinations		
[32]	Backpropagation (BP), Support Vector Regression (SVM), and Gradient Boosting Classifier (GBC)			BP had the highest accuracy rate of 80.91% and 87.78%
[33]	Decision Tree, Naïve Bayes, Logistic Regression, Support Vector Machine, K-Nearest Neighbour, Sequential Minimal Optimisation, and Neural Network.	Student Number, Study Year, Gender, Birth Year, Registration Course, Employment Activity Point (40), Examination Point (60), Final Point (100), Grade		Logistic Regression (68.7% for passed and 88.8% for failed)
[34]	MLP-ANN, Naïve Bayes(NB), Support Vector Machines(SVM), K Nearest Neighbor (KNN), and LDA.	ID, Age, B.Sc. degree, B.Sc. grade, Course names, Course Grades Instructor Names, Grades (All Courses Grades), Grade (Dissertation Grade)		LDA (79%)
[35]	Decision tree		Yes	92.27%
[36]	Random Forest (RF), Decision Tree	Excellent learner, Good learner, Average learner, Slow learner, Father's educational		Hybrid

	(C5.0), Naïve Bayes (NB) and Support Vector Machine (SVM)	level, Mother's educational level, Father's occupational status, Mother's occupational status, Family's socioeconomic, Parents' attention to students' attitude, Parents' time and money spending, Parents' involvement as education, Parents' feeling responsive and need, Parents' respond to children's attitude, Parents' encouragement, Parents' compliment, Domestic environment for study, Distance from home to school, Number of hours for self-study, Number of hours for private math study, Frequency of doing math homework, Frequency of absence in math class, Frequency of preparing for the math exam, Student' s interest in math, Student's enjoyment in math class, Student's attention in math class, Student's motivation to succeed in math, Student's anxiety in math class, Student's nervous in the math exam, Student's feeling helpless in math, Internet's use at home, Possession of computer, Student's study desk at home, Classroom environment, Content's language in math class, Class session, Teacher mastering in math class, Teacher's absence in math class, Teaching methods in math class, Teacher's involving in education's content, Math teacher's ability, Teacher's encouragement to students, Math teacher's connection with students, Math teacher's help, Adequate number of math teacher, Adequate use of classroom, Adequate use of math handout		Random Forest (99.72%)
[37]	CART, SVM			CART (93.24)
[38]	Multivariate Linear Regression			MLR (70.48%)
[39]	Naive-Bayes, Decision Tree, Random-Forest, JRip, and ZeroR	Name, Sex, Age, Address, Family size, Parent's cohabit, Mother Qualification, Father Qualification, Job type of mother, Job type of the Father, Reason to desire this school, Guardian, Home to school travel time, Weekly study time, No's of past class failures, Extra study support	Yes	JRip (85.31%).
[40]	Neural Network	Class Test Mark, Education, Class Performance, Living Area Hall/Mess, Class Attendance, Social Media Interaction, Assignment, Extra Curricular Activity, Lab Performance, Drug Addiction , Study Time, Affair, Previous Result, Year Final Result		neural network (90%)
[41]	ANN–GSA and ANN–CS models	University entrance examination score, Subject 1, Subject 2, Subject 3, The average overall score of high school graduation examination, Elapsed time between graduating from high school and obtaining university admission, Location of student's high school, Type of high school attended, Student's gender		
[42]	Random Forest, Logistic Regression, Neural Net, XG Boost	Students' ID, Gender, Program Enrolment days after the start of the semester, Days from first to last session, Unique days, Actions taken, Average actions per day, Actions taken from the university network, Actions taken from outside the university network, Tasks count compared to average, Files viewed compared to average, Files uploaded	Yes	Xgboost and Neural Net (90.62%)

		compared to average, Exam taken		
[43]	RF, ANN, SVM, and LR	Year of the study cycle, Portuguese citizenship, Portuguese naturality, Gender, Student's age (years), Number of enrolled years in high school, Number of failures in the educational career, Scholarship, Level of financial support received by the government, Availability of a Personal Computer (PC) at home, Internet access, Class size, School size, Economic level of residence area, the Population density of residence area, Rural residence area, Number of unit courses attended in the present academic year		RF (AUROC - 87%)
[44]	Naïve Bayes, Decision Tree, And Neural Network	Gender, Family, Distance, High School, GPA, Entrance exam, Scholarships, Time, Materials, Internet, Grade importance, Earnings		Naïve Bayes (76.65%)
[45]	Gradient Boosting and Random Forest classifiers	Student's status in terms of grades, Other info indicating a student's status, Student's course load, Course's difficulty, Performance / Familiarity with the course's background and department, Information about the prerequisites, and Performance relative to the course's level.	Yes	Random Forest classifiers (AUC - 0.854)
[46]	Ensemble-Based Progressive Regression, Logistic Regression, Random Forest And K-Nearest Neighbors (Knn).			Ensemble-Based Progressive Regression
[47]	Support vector machines (SVM), Decision Tree, and Naïve Bayes (NB)	Attribute Name, School, Sex, Age, Address, Family size, Parental status, Mother Education, Father education		Support vector machine (78.38%)
[48]	Multilayer Perceptron and Generalized Regression Neural Network	Academic factors (students' results)		Generalized Regression Neural Network model (95 %)
[49]	Deep neural network, Decision Tree (C5.0), Naïve Bayes, Random Forest, Support Vector Machine, K-Nearest Neighbor	Gender, Nationality, Place of Birth, Parent Responsible, Stages, Grades, Section ID, Topic, Student Absent day, Semester, Raised hand, Visited Resource, Viewing Announcement, Discussion Group, Parent Answering, Parent Satisfaction		Deep Neural Network (84%).
[50]	K- Nearest Neighbor	Passed Math, Failed Math, Passed the language, Failed the language, Age, Gender, Father's education, Mother's education Father's occupation, Mother's occupation, Socioeconomic Status		
[51]	Adaboost, Bagging, Dagging, and			Adaboost

	Grading			

Conclusion

Predicting students' performance is important in the educational domain because students' status analysis helps improve the performance of institutions and is also crucial for effectively carrying out the necessary pedagogical strategies needed to improve instructional delivery. As educational records are accumulating and increasing rapidly through traditional (demographic, academic background, and behavioural features) and multimedia databases, it is imperative to handle this large volume of data using different data mining techniques. Machine learning techniques in educational data mining aim to develop a model for discovering meaningful hidden patterns and exploring useful information from educational settings. This study proved that the academic performances of the students are primarily dependent on their past performances but also external activities such as social media interaction, living area conditions and so on.

REFERENCES

[1] D. T. Larose and C. D. Larose, *Discovering Knowledge in Data*. John Wiley {\&} Sons, Inc., 2014.

[2] H. A. Madni, Z. Anwar, and M. A. Shah, "Data mining techniques and applications - A decade review," *ICAC 2017 - 2017 23rd IEEE Int. Conf. Autom. Comput. Addressing Glob. Challenges through Autom. Comput.*, 2017, doi: 10.23919/IConAC.2017.8082090.

[3] E. S. Priya and K. Anandhan, "An Overview of Data Mining - A Survey Paper," *Int. J. Mod. Comput. Sci.*, vol. 6, no. 1, pp. 19–21, 2018.

[4] A. V. Manjarres, L. G. M. Sandoval, and M. J. S. Suárez, "Data mining techniques applied in educational environments: Literature review," *Digit. Educ. Rev.*, no. 33, pp. 235–266, 2018, doi: 10.1344/der.2018.33.235-266.

[5] M. Porkizhi, "A Study of Data Mining Techniques And Its Applications," *IInternational J. Sci. Adv. Res. Technol.*, vol. 3, no. 4, pp. 1402–1406, 2017.

[6] R. Tamilselvi and S. Kalaiselvi, "An Overview of Data Mining Techniques and Applications," *Int. J. Sci. Res.*, vol. 2, no. 2, pp. 506–509, 2013.

[7] S. A. Salloum, M. Alshurideh, A. Elnagar, and K. Shaalan, "Mining in Educational Data: Review and Future Directions," in *Advances in Intelligent Systems and Computing*, Springer International Publishing, 2020, pp. 92–102.

[8] N. Naicker, T. Adeliyi, and J. Wing, "Linear Support Vector Machines for Prediction of Student Performance in School-Based Education," *Math. Probl. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/4761468.

[9] T. B. Bithari, S. Thapa, and H. K.C., "Predicting Academic Performance of Engineering Students Using Ensemble Method," *Tech. J.*, vol. 2, no. 1, pp. 89–98, Nov. 2020, doi: 10.3126/tj.v2i1.32845.

[10] E. A. Amrieh, T. Hamtini, and I. Aljarah, "Mining Educational Data to Predict Student's Academic Performance using Ensemble Methods," *Int. J. Database Theory Appl.*, vol. 9, no. 8, pp. 119–136, Aug. 2016, doi: 10.14257/ijdta.2016.9.8.13.

[11] A. Algarni, "Data Mining in Education," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 6, 2016, doi: 10.14569/ijacsa.2016.070659.

[12] A. B. Zorić, "Predicting Students' Success Using Neural Network," *{SSRN} Electron. J.*, 2019, doi: 10.2139/ssrn.3490105.

[13] E. B. Belachew and and Feidu Akmel Gobena, "Student Performance Prediction Model using Machine Learning Approach: The Case of Wolkite University," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 7, no. 2, pp. 46–50, Feb. 2017, doi: 10.23956/ijarcsse/v7i2/01219.

[14] S. Hussain, Z. F. Muhsion, Y. K. Salal, P. Theodoru, F. Kurto\uglu, and G. C. Hazarika, "Prediction Model on Student Performance based on Internal Assessment using Deep Learning," *Int. J. Emerg. Technol. Learn.*, vol. 14, no. 08, p. 4, Apr. 2019, doi: 10.3991/ijet.v14i08.10001.

[15] F. Ofori, E. Maina, and R. Gitonga, "Using Machine Learning Algorithms to Predict Students ' Performance and Improve Learning Outcome : A Literature Based Review Francis Ofori , Dr . Elizaphan Maina and Dr . Rhoda Gitonga ISSN : 2617-3573 Using Machine Learning Algorithms to Predict Students," *J. Inf. Technol.*, vol. 4, no. 1, pp. 33–55, 2020.

[16] H. Waheed, S.-U. Hassan, N. R. Aljohani, J. Hardman, S. Alelyani, and R. Nawaz, "Predicting academic performance of students from {VLE} big data using deep learning models," *Comput. Human Behav.*, vol. 104, p. 106189, Mar. 2020, doi: 10.1016/j.chb.2019.106189.

[17] R. Umer, T. Susnjak, A. Mathrani, and S. Suriadi, "On predicting academic performance with process

mining in learning analytics," *J. Res. Innov. Teach. Learn.*, vol. 10, no. 2, pp. 160–176, 2017, doi: 10.1108/jrit-09-2017-0022.

[18] E. G. Rincon-Flores, E. Lopez-Camacho, J. Mena, and O. O. Lopez, "Predicting academic performance with Artificial Intelligence (AI), a new tool for teachers and students," *IEEE Glob. Eng. Educ. Conf. EDUCON*, vol. 2020-April, pp. 1049–1054, 2020, doi: 10.1109/EDUCON45650.2020.9125141.

[19] M. F. Musso, E. Kyndt, E. C. Cascallar, and F. Dochy, "Predicting general academic performance and identifying the differential contribution of participating variables using artificial neural networks," *Front. Learn. Res.*, vol. 1, no. 1, pp. 42–71, 2013, doi: 10.14786/flr.v1i1.13.

[20] E. A. Amrieh, T. Hamtini, and I. Aljarah, "Mining Educational Data to Predict Student's academic Performance using Ensemble Methods," *Int. J. Database Theory Appl.*, vol. 9, no. 8, pp. 119–136, 2016, doi: 10.14257/ijdta.2016.9.8.13.

[21] E. J. Phua and N. K. Batcha, "Comparative analysis of ensemble algorithms'prediction accuracies in education data mining," *J. Crit. Rev.*, vol. 7, no. 3, pp. 37–40, 2020, doi: 10.31838/jcr.07.03.06.

[22] A. K. Hamoud and A. M. Humadi, "Student's Success Prediction Model Based on Artificial Neural Networks (ANN) and A Combination of Feature Selection Methods," *J. Southwest Jiaotong Univ.*, vol. 54, no. 3, 2019, doi: 10.35741/issn.0258-2724.54.3.25.

[23] L. S. Affendey, I. H. M. Paris, N. Mustapha, M. N. Sulaiman, and Z. Muda, "Ranking of influencing factors in predicting students' academic performance," *Information Technology Journal*, vol. 9, no. 4. pp. 832–837, 2010, doi: 10.3923/itj.2010.832.837.

[24] F. D. Pereira, S. C. Fonseca, E. H. T. Oliveira, D. B. F. Oliveira, A. I. Cristea, and L. S. G. Carvalho, "Deep learning for early performance prediction of introductory programming students: a comparative and explanatory study," *Rev. Bras. Informática na Educ.*, vol. 28, pp. 723–748, 2020, doi: 10.5753/rbie.2020.28.0.723.

[25] Y. K. Saheed, T. O. Oladele, A. O. Akanni, and W. M. Ibrahim, "Student performance prediction based on data mining classification techniques," *Niger. J. Technol.*, vol. 37, no. 4, p. 1087, 2018, doi: 10.4314/njt.v37i4.31.

[26] F. Ünal, "Data Mining for Student Performance Prediction in Education," in *Data Mining - Methods, Applications and Systems*, IntechOpen, 2021.

[27] S. Vairachilai and Vamshidharreddy, "Student's Academic Performance Prediction Using Machine Learning Approach," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 9, pp. 6731–6737, 2020.

[28] M. Imran, S. Latif, D. Mehmood, and M. S. Shah, "Student academic performance prediction using supervised learning techniques," *Int. J. Emerg. Technol. Learn.*, vol. 14, no. 14, pp. 92–104, 2019, doi: 10.3991/ijet.v14i14.10310.

[29] H. Altabrawee, O. A. J. Ali, and S. Q. Ajmi, "Predicting Students' Performance Using Machine Learning Techniques," *J. Univ. BABYLON Pure Appl. Sci.*, vol. 27, no. 1, pp. 194–205, 2019, doi: 10.29196/jubpas.v27i1.2108.

[30] A. Soni, V. Kumar, R. Kaur, and D. Hemavathi, "Predicting Student Performance Using Data Mining Techniques," *Int. J. Pure Appl. Math.*, vol. 119, no. 12, pp. 221–227, 2018.

[31] I. E. Livieris, K. Drakopoulou, V. T. Tampakas, T. A. Mikropoulos, and P. Pintelas, "Predicting Secondary School Students' Performance Utilizing a Semi-supervised Learning Approach," *J. Educ. Comput. Res.*, vol. 57, no. 2, pp. 448–470, 2019, doi: 10.1177/0735633117752614.

[32] B. Sekeroglu, K. Dimililer, and K. Tuncal, "Student performance prediction and classification using machine learning algorithms," *ACM Int. Conf. Proceeding Ser.*, vol. Part F1481, pp. 7–11, 2019, doi: 10.1145/3318396.3318419.

[33] A. S. Hashim, W. A. Awadh, and A. K. Hamoud, "Student Performance Prediction Model based on Supervised Machine Learning Algorithms," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 928, no. 3, 2020, doi: 10.1088/1757-899X/928/3/032019.

[34] L. M. A. Zohair, "Prediction of Student's performance by modelling small dataset size," *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 1, 2019, doi: 10.1186/s41239-019-0160-3.

[35] G. Okereke, "A Machine Learning Based Framework for Predicting Student's Academic Performance," *Phys. Sci. Biophys. J.*, vol. 4, no. 2, 2020, doi: 10.23880/psbj-16000145.

[36] P. Sokkhey and T. Okazaki, "Hybrid machine learning algorithms for predicting academic performance," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 1, pp. 32–41, 2020, doi: 10.14569/ijacsa.2020.0110104.

[37] V. A. Sungar, P. D. Shinde, and M. V. Rupnar, "Predicting Student's Performance using Machine Learning," *Commun. Appl. Electron.*, vol. 11, no. 7, 2017.

[38] H. Agrawal and H. Mavani, "Student Performance Prediction using Machine Learning," *Int. J. Eng. Res.*, vol. V4, no. 03, Mar. 2020, doi: 10.17577/ijertv4is030127.

[39] N. Walia, M. Kumar, N. Nayar, and G. Mehta, "Student's Academic Performance Prediction in Academic using Data Mining Techniques," *SSRN Electron. J.*, pp. 1–5, 2020, doi:

10.2139/ssrn.3565874.
[40] M. F. Sikder, M. J. Uddin, and S. Halder, “Predicting students yearly performance using neural network: A case study of BSMRSTU,” *2016 5th Int. Conf. Informatics, Electron. Vision, ICIEV 2016*, pp. 524–529, 2016, doi: 10.1109/ICIEV.2016.7760058.
[41] J.-F. Chen and Q. H. Do, “Training Neural Networks To Predict Student Academic Performance: A Comparison Of Cuckoo Search And Gravitational Search Algorithms,” *Int. J. Comput. Intell. Appl.*, vol. 13, no. 01, p. 1450005, Mar. 2014, doi: 10.1142/s1469026814500059.
[42] A. Yanka, “Predicting Students Performance In Moodle Platforms Using Machine Learning Algorithms,” 2019.
[43] F. Cruz-Jesus *et al.*, “Using artificial intelligence methods to assess academic achievement in public high schools of a European Union country,” *Heliyon*, vol. 6, no. 6, p. e04081, 2020, doi: 10.1016/j.heliyon.2020.e04081.
[44] E. Osmanbegovic and M. Suljic, “Data Mining Approach for Predicting Student Performance,” *J. Econ. Bus.*, vol. 10, no. 1, pp. 3–12, 2012.
[45] A. Polyzou and G. Karypis, “Feature extraction for classifying students based on their academic performance,” *Proc. 11th Int. Conf. Educ. Data Mining, EDM 2018*, pp. 356–362, 2018.
[46] J. Xu, K. H. Moon, and M. Van Der Schaar, “A Machine Learning Approach for Tracking and Predicting Student Performance in Degree Programs,” *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 5, pp. 742–753, 2017, doi: 10.1109/JSTSP.2017.2692560.
[47] S. Urkude and K. Gupta, “Student intervention system using machine learning techniques,” *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6 Special Issue 3, pp. 2061–2065, 2019, doi: 10.35940/ijeat.F1392.0986S319.
[48] A. R. Iyanda, O. D. Ninan, A. O. Ajayi, and O. G. Anyabolu, “Predicting Student Academic Performance in Computer Science Courses: A Comparison of Neural Network Models,” *Int. J. Mod. Educ. Comput. Sci.*, vol. 10, no. 6, pp. 1–9, 2018, doi: 10.5815/ijmecs.2018.06.01.
[49] V. Vijayalakshmi and K. Venkatachalapathy, “Comparison of Predicting Student‘s Performance using Machine Learning Algorithms,” *Int. J. Intell. Syst. Appl.*, vol. 11, no. 12, pp. 34–45, 2019, doi: 10.5815/ijisa.2019.12.04.
[50] M. A. Rasheed *et al.*, “Use of artificial intelligence on Electroencephalogram (EEG) waveforms to predict failure in early school grades in children from a rural cohort in Pakistan,” *PLoS One*, vol. 16, no. 2 February, pp. 1–12, 2021, doi: 10.1371/journal.pone.0246236.
[51] A. Shanthini, G. Vinodhini, and R. M. Chandrasekaran, “Predicting students’ academic performance in the University using meta decision tree classifiers,” *J. Comput. Sci.*, vol. 14, no. 5, pp. 654–662, 2018, doi: 10.3844/jcssp.2018.654.662.
[52] Z. Iqbal, J. Qadir, A. N. Mian, and F. Kamiran, “Machine Learning Based Student Grade Prediction: A Case Study,” pp. 1–22, 2017, [Online]. Available: http://arxiv.org/abs/1708.08744.
[53] H. Liu and L. Yu, “Toward integrating feature selection algorithms for classification and clustering,” *{IEEE} Trans. Knowl. Data Eng.*, vol. 17, no. 4, pp. 491–502, Apr. 2005, doi: 10.1109/tkde.2005.66.