Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

PREDICTION METHODS ON STUDENTS' ACADEMIC PERFORMANCE: A REVIEW

[1]YOUSUF NASSER SAID AL HUSAINI and [2]NUR SYUFIZA AHMAD SHUKOR
[1, 2] Faculty of Communication Visual Art and Computing, Universiti Selangor, Shah Alam, Malaysia.

Abstract

Predicting student academic performance is linked to developing the best educational policies in higher education, which significantly impact economic and financial development. The wealth of readily available educational data makes it possible to address student issues, improve the learning environment, and make decisions based on data through the use of technology-enhanced learning platforms. It is impossible to evaluate a student's standing at a university without considering their academic performance. It allows academic staff, administrators, and decision-makers to evaluate students throughout a semester accurately. It also aids students in assessing their performance and improving it. This paper presents a comprehensive review of related studies on student academic performance. Several techniques have been reviewed, such as Support Vector Machine (SVM), Naïve Bayes (NB), Logistic Regression (LR), Decision Trees (DT), Extreme Learning Machine (ELM), Artificial Neural Network ANN, k-Nearest Neighbors (kNN), and ensemble methods such as Bagging, Random Forest (RF), and Adaptive Boosting (AB). In addition, student factors have been used and compared through different classifiers. Accordingly, the findings confirmed the usefulness of Neural Network as the most competitive classifier, and academic assessment was a prominent factor when predicting students' academic performance.

Key words: Prediction Models, Data Mining, Academic Performance, Deep Learning

1. INTRODUCTION

In the present information era, education is one of the most important factors in determining the level of literacy in a society and the rate at which a country's economy grows. Education not only improves the ability to make decisions but also helps create a generation that is more competitive [1]. One of the most important criteria for evaluating students at a higher education institution is their academic performance. Colleges, educational institutes, and schools are expanding at a faster rate to provide better education to students in this competitive world. The educational institutions place a premium on producing graduates who excel in both academics and extracurricular activities. They keep track of how students perform in a particular field and where they require additional training [2]. The performance of the student shifts throughout the course of an academic year, and the number of students who fail an academic course rises as a direct result of a deterioration in that student's performance for a variety of reasons [3].

There are a number of factors influencing students' academic performance, and those can be used for predictions. Accordingly, the literature revealed that the students' factors that affect the student's academic performance are gender, high school grade,

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

assignment performance, attendance, student parental education, student family status, living location, students' previous marks, financial background, teaching methods, seminar performance, test marks, general proficiency, Interest in a particular course, study behavior, engage time and family Support, previous schools marks, admission type, accommodation type, parent's occupation, parent's qualification [4],[5],[4],[6],[7],[8].

The above factors can assist in predicting students' performance. Next, many Data Mining (DM) techniques and tools are now being used to assist in accurately analyzing prediction results. The Educational Data Mining (EDM) domain is students' academic performance (SAP) prediction domain. EDM tools frequently generate prediction models to aid SAP prediction that monitors students' academic progress and helps students and other education stakeholders identify essential strategies to use.

Academic performance prediction is concerned with determining whether or not a student is likely to be retained in an academic institution because of their academic performance and perhaps other factors. These include socioeconomic, academic, and psychological factors. As a result, education stakeholders need to pay more attention to this group of students if they want to help them stay in the system and achieve their educational goals.

For predicting students' academic performance, several studies in educational data mining have used data mining techniques such as Support Vector Machine (SVM), Naïve Bayes (NB), Logistic Regression (LR), Decision Trees (DT), Extreme Learning Machine (ELM), Artificial Neural Network ANN, k-Nearest Neighbors (kNN),and ensemble methods such as Bagging, Random Forest (RF), and Adaptive Boosting (AB). Other data mining techniques used to predict student academic performance included the use of Deep Learning, Nature-Inspired algorithms, and hybrid methods. These studies demonstrated the ability of these techniques to accurately predict student performance at various levels of study using a variety of student performance factors. However, more research is still needed in order to conduct and produce an improved framework for student academic performance predictions. The purpose of this literature review was to examine the current research on the major factors influencing students' academic performance, as well as the impact of various types of data on various classifiers.

The following is a breakdown of the paper's structure. Section 2 delves the search methodology, while Section 3 delves into the methods used for student academic performance. Finally, Sections 4 and 5 outline the discussion and conclusion of this literature review, respectively.

2. SYSTEMATIC REVIEW METHOD

In this study, the systematic literature review (SLR) methodology that adheres to the suggestions made by Kitchenham et al. (2009) was applied. SLR has a number of benefits over unorganized and unreliable literature reviews because it is more likely to be regarded as reliable and impartial [9]. Information obtained from SLR is also very reliable because it comes from a variety of sources. The three phases of SLR are planning, conducting, and reporting [10].

2.1 Research Questions

Our objective in this paper is to answer the following question:
What methods are used in students' academic performance prediction?

2.2 SEARCH STRATEGY

Multiple databases, including Science Direct, Scopus, IEEE Xplore, ACM Digital Library, and Google Scholar, were used in the search for manuscripts. Between 2014 and 2021, research articles were searched. Search terms used to find articles included "student" AND "predict*" AND "Academic performance" AND "factors" AND "review" OR "survey". In the search, about 3,800 results were found. The papers were thoroughly investigated with the keywords listed as focus of the analysis. In the end, 86 papers have been reviewed.

3. METHODS USED FOR STUDENT ACADEMIC PERFORMANCE PREDICTIONS

There are numerous methods for predicting student academic performance, and this section provides an overview of the data mining classifiers used in each classification. Figure 1 depicts the approach taken to predicting student academic performance from three human endeavor fields: Learning Analytics, Educational Psychology, and Data Mining. It has been hypothesized by Educational Psychology researchers that certain psychological factors have an impact on a student's performance; for example, [11] look at the impact of students' feelings toward their own academic control and academic emotion on their grades.

To put it another way, when compared to educational psychologists' adopted method of analyzing students' performance in learning. The purpose of learning analytics is to better understand and improve learning and the environments in which it takes place through the systematic collection, organization, analysis, and reporting of data about learners and their contexts. An example of one of these methods is the use of statistics and other forms of data analysis such as web analytics and operational research. There have been a number of studies in this area, including [12], who used a method of advanced learning analytics to student academic prediction. Data mining for educational purposes is a relatively new field that is rapidly growing in popularity, Data Mining (DM) has a lot of applications in student academic performance predictions. There are three

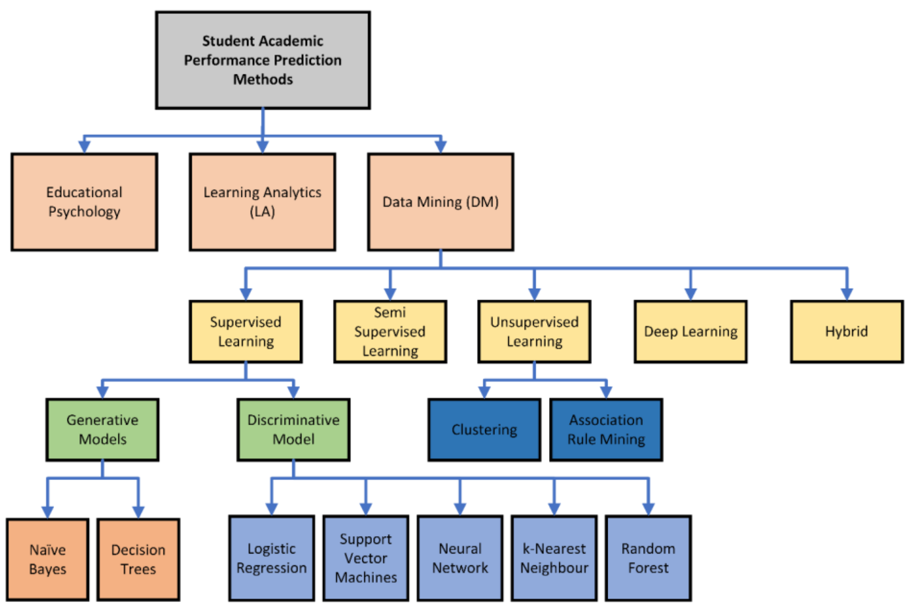
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

types of DM approaches that have been used so far: semi-supervised learning, supervised learning, and unsupervised learning.

There are two types of supervised learning models: generative and discriminative. It's important to note that a generative model defines the joint probability distribution P(x, y) of inputs (x) and outcomes (y) while a discriminative model defines the conditional probability P(y|x) of outputs (y). Classification problems are solved directly rather than through an intermediate step in discriminative models, in contrast to generative models.

Support Vector Machine (SVM), Logistic Regression (LR), k-Nearest Neighbor (kNN), Neural Network (NN), and Random Forest (RF) are examples of commonly used discriminative models, while Decision Trees (DT) and Naïve Bayes (NB) are examples of generative models.

Figure 1: Methods for Student Academic Performance predictions



3.1 NAÏVE BAYES (NB)

This classifier is based on the Bayesian probability theory, which holds that every attribute in a dataset belonging to a particular class is completely unrelated to any other. Strong (naïve) independence assumptions were used in the classification task, assuming that each variable contributed equally. It is a method for describing and predicting whether or not a target tuple belongs to a particular class.

The Naïve Bayes method has been used in several studies to predict student academic performance, the works include Aziz et al. [13] was used to analyze the performance of students from the Student Entry Management Database and Unisza's academic

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

databases. As a dependent variable, Grade Point Average (GPA) is used to predict students' performance based on the other five independent variables, with labels such as "poor," "average," and "good." According to their findings, NBC had the highest accuracy of 57.4 percent, and family income was the most influential predictor of students' performance, with a probability of 56.8 percent when three-fold cross validation was applied. Predictive model outperforms when tested three times on average students, but it fails to do so when tested three times on students who are less average.

Kaur, Singh, and Josan [14] used the Naïve Bayes algorithm, as well as other algorithms such as SMO, MLP, and J48, to prediction student performance and display the results based on classification-based algorithms. Various classification algorithms (Naïve Bayes, Multi-Layer Perceptron, J48, REPTree, and SMO) were used to test a dataset from a high school. FP Rate, TP Rate, recall, precision, ROC area, F-measure, and accuracy were used to test and validate the models. There were no other classifiers that performed better than the Multi-Layer Perceptron in the study, with an accuracy of 75%.

On the dataset of undergraduate students, Mueen, Zafar, and Manzoor [15] used three different classification algorithms: Neural Network, Decision Tree and Naïve Bayes. Naïve Bayes was found to be superior to the other two algorithms, with an accuracy rate of 86%.

Applying this method to fictitious information containing nine attributes, Kaur and Singh [16], used Naïve Bayes alongside a variant of Decision Tree, J48. The outcomes of the classification revealed that the Naïve Bayes algorithm, with a prediction accuracy of 63.5%, outperformed the J48 algorithm, which had a prediction accuracy of 61.5%. The classifiers' prediction accuracy is used to make comparisons between the techniques.

Agrawal, Vishwakarma, and Sharma [17], Naïve Bayes, random forest, rule induction, and decision tree were used to classify student records from two Portuguese schools. At cross validation folds of 10, 20, 30, 40, and 50, the decision tree outperformed the other classifiers, but all of the classifiers employed, performed within the same range and were extremely close including the Naïve Bayes.

3.2 DECISION TREES (DT)

Use of data mining methods like decision trees is commonplace. A root node (the first node) starts with an attribute, which is then divided into leaves, which can be further divided into other leaf nodes (internal nodes with an incoming edge and/or several outgoing edges) according to certain criteria. Tree termination occurs when the last node in a tree has no outgoing edges. The attributes in a dataset are used to create a top-down tree-like model in a decision tree, which is a supervised classification technique. Decision trees use leaf nodes to represent the predicted class label for each instance [18].

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

There are two phases in every decision tree classifier: the building phase and the pruning phase. First, the tree is constructed by recursively splitting the training set according to local optimal criteria until all or nearly all of the records in each partition are labeled as belonging to the same class as the previous one. Nevertheless, overfitting is a possibility and is dealt with in the second phase of pruning. This phase enhances classification accuracy by removing noise and outliers from the tree. Decision trees also included the C4.5 or J48, C5.0, the CART, and ID3 (the Interactive Dichotomizer 3).

Ogunde and Ajibade [19] used the ID3 decision tree algorithm to link students' final graduation academic results to their entry grades, and many other studies have used similar algorithms to predict student academic performance. The academic department of Redeemer's University provides data on gender, student entrance exam scores, graduation grades (B.Sc), and entry grades in secondary school. The ID3 algorithm was applied to the dataset using WEKA, and the rules generated by the application were incorporated into the knowledge base for the Java prediction system. Even before they started college, the developed prediction system was able to aid in predicting student final grade.

Joseph and Devadas [20] used 56 students from the first batch of CSE students at the College of Engineering Munnar to develop the weighted modified ID3 algorithm to predict students' performance. ID3 (52.08 %), C4.5 (45.83 %), and CART (56.25 %) are all outperformed by the modified weighted ID3 (76 %) prediction accuracy. End Semester mark was used as well as previous semester mark and class test mark. Attendance in class and lab work were also taken into consideration. The purpose of the study was to see how well the algorithm had been modified.

Al-barrak and Al-razgan [4] using WEKA, the Computer Sciences College students at King Saud University of the year 2012 dataset was used to predict the students' final GPA using the J48 decision tree algorithm. There was a stronger correlation between students' final GPA and Java2 than Java1 in this study, indicating that students' final grades were heavily influenced by Software Engineering. Using the J48 algorithm, the resulting tree revealed that Java1 has the greatest impact on a student's final grade point average (GPA).

Afeni, Oloyede, and Okurinboye [21] predicted student performance in six academic departments at Joseph Ayo Babalola University using ID3 and C4.5 algorithms. Using four performance metrics, the model's accuracy, precision, false alarm rate, and sensitivity were all verified to be accurate, and reliable. With a 61 % accuracy rate, the ID3 algorithm outperformed the C4.5 algorithm.

3.3 LOGISTIC REGRESSION (LR)

This is a linear model for categorical response variables. Probability of a certain value for the dependent variable is calculated. In fact, logistic regression can only be used with a categorical binary output variable, even if the inputs are quantitative. The final

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

probability value of the output can be used to draw a conclusion about which categorical value, 0 or 1, is more likely to occur after the logistic regression probability for all inputs has been calculated [22].

Logistic regression was usen in many studies for student academic performance predictions such as [23] as they used data from the university database of the University of Washington in the United States to model student dropout. The dataset contains the data of 32,538 students, and k-nearest neighbor, Random Forest and regularized logistic regression were used to predict binary drop-out based on 784 additional factors such as gender, race, GPA, resident status, and so on. Their findings show the strongest individual predictors of attrition were GPA in English, elementary, psychology and Math courses.

3.4 SUPPORT VECTOR MACHINE (SVM)

Support Vector Machine (SVM) is an algorithm with a set of distinct characteristics. Nonlinear separation problems can be solved by mapping nonlinearly separable points into a higher-dimensional space, where a linear classifier can then be applied, according to the "kernel trick" [24]. In contrast to other learning algorithms, SVM doesn't suffer from the problems of computational complexity or overfitting. It is based on a linear model that uses an algorithm to find the maximum margin hyperplane. With linearly distinct classes, the maximum margin hyperplane is most effective at separating them. Multivariate and categorical variables can be handled by SVM, which can perform classification and regression tasks at the same time [25], [26].

Another way of saying this: the SVM is a classifier that works by creating multidimensional hyperplanes to separate cases with different class labels. Examples are added to the database and given a category based on where they fit into the overall scheme.

Bhagvatula et al. [27] used SVM, NB, and J48algorithms to predict student performance and investigate academic performance. While, Lakkaraju et al. [28] developed a machine learning framework using four other data mining algorithms additional to SVM to identify high school students at risk of not graduating.

Asogbon, Samuel, Omisore, & Ojokoh (2016) developed a multi-class SVM for performance prediction using a University of Lagos dataset. Using 7-fold cross validation, multi-class SVM accurately predicted student performance [29].

To classify students' performances, Burman & Som (2019) using RBF and linear basis functions (RBF) as kernels for multi-classifier SVM. The two SVM variants were applied to students' data psychological parameters gathered through questionnaires. A prediction accuracy of 90.97% was found for RBF kernel SVM over its linear counterpart [30].

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

3.5 NEURAL NETWORK (NN)

An input/output unit in a neural network has a weighted combination of inputs and outputs. Neurons found in human brains were used as a basis for it. Simulated neurons are connected in a network similar to the brain's actual neurons. It is possible for the network to learn by changing (or being changed by a learning algorithm) the strength of neural connections in response to an input stimulus or an output. The network adjusts its weights during the learning phase so that it can correctly predict the class labels of the input tuples [31].

Neural Network has been used in a number of research studies, Osofisan, Adeyemo, and Oluwasusi (2014) used MLP, a variant of Neural Network, and the J48 algorithm to investigate the behavior of student performance data. The researchers tested the two algorithms on M.Sc students' data from the University of Ibadan's Computer Science department to see how well they performed in mining educational data. Despite the fact that the J48 algorithm (0.25secs) took less time to build than the MLP (2.7secs) on the training dataset, the MLP algorithm (98.3 %) outperformed the J48 algorithm (85.4 %) in prediction accuracy. The same trend was seen in the test dataset, with MLP achieving 60.2 % accuracy in 5.93 seconds and the J48 model achieving 52.8 % accuracy in 0.04 seconds. As a result, the study concluded that in mining educational data, Artificial Neural Networks provide the best classification and prediction results [32].

Ruby and David [33] predict pupils' educational performance and evaluate factors that influence student performance predictions using MLP. The Multi-Layer Perceptron (MLP) algorithm was used to model student performance using data from a PG Computer Application course. WEKA was used to generate MLP models for the two datasets with high influencing factors (7 attributes each), as well as the entire dataset of 12 attributes. Research shows that models based on highly influential factors perform better than models based on any one or more of the other 12 attributes.

3.6 K-NEAREST NEIGHBOR (KNN)

This algorithm selects a value for K, the number of neighbors, searches the training set for K observations that are close to the target, and uses the most popular response value from the K closest neighbors as the predicted response. It starts with K=1 for parameter tuning, then KNN searches for one closest observation at a time until the best value for K is found, at which point it increases by one [34]. This technique was used in studies like Asif, Merceron, et al., [35] to analyze the students' performance using a variety of methods.

Another study by Verma, Singh & Verma [36], Students' academic performance was predicted using kNN and other mining methods. To analyze the dataset of SPSU University students, the researchers used six algorithms, including two Bayesian classifiers – Naïve Bayes and BayesNet, two rule learners – OneR and JRip, J48 (C4.5), and an instance-based learner (k-NN). These include birth date, gender, current

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

semester's total university score, and location and so forth. Researchers found that a student's university entrance score and a number of first-year exam failures were among the most critical factors in classification.

3.7 RANDOM FOREST (RF)

This is a method that combines a number of different decision tree classifiers into a single forest. The split is determined by selecting random attributes at each node of the decision tree. In the forest, the value of each tree is determined by the value of a random vector sampled independently from the others and distributed uniformly across all trees. When classifying trees with RF, each tree gets to cast a vote, and the class with the most popularity returned [37].

In addition, it is a supervised learning algorithm with the advantage of being able to be used for both classification and regression. It was created to address the problem of data overfitting in decision trees, which is addressed through pruning and ensemble learning. The main difference between the decision forest and the random forest is that the decision forest constructs its models using either the sample subset method or the feature subset method, whereas the random forest employs both [38].

Gilbert [39] used Random Forest to predict student outcomes. Over 31,000 freshmen and transfer students at California State University were studied using RF and the genetic algorithm (GA) from Fall 2000 to Fall 2010. Non-linear ensemble methods, such as RF and GA, were used to uncover interactions between variables that would have gone unnoticed in a linear system, as well as to improve the feature selection process and achieve a healthy balance between precision and recall. Researchers found that retention and graduation rates after one and two years of attendance could be accurately predicted.

Mishra, Kumar, and Gupta [40] used Random Forest in conjunction with the J48 algorithm for student performance prediction based on their academic and social integration factors. The study used data mining techniques to investigate the relationship between students' emotional skills and socioeconomic and previous performance parameters in order to predict the performance. Standard Emotional Skill Assessment Process was used to assess emotional skills such as assertiveness, leadership, and stress management. The study's findings revealed, among others, that among all of the students' leadership, drive, and emotional attributes were found to have an impact on their performance.

3.8 CLUSTERING

In the process of clustering, a set of data objects is divided into numerous groups or clusters, where each cluster has a high similarity but is distinct from the other clusters. The label is unknown, unlike the classification. Clustering algorithms such as k-medoid and the k-means algorithms are just two examples.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

In the work of Asif et al., [35], Other well-known data mining techniques were used in conjunction with this one. In order to categorize the students in the datasets, they applied the X-means algorithm, a variant of the means clustering algorithm.

3.9 Association Rule Mining

The process of discovering reliable rules for the correlation of various items in a large dataset is called association rule mining. As a result, two steps are required: first, find out how many transactions are supported by each item and then use that information to create a set of database association rules with a level of certainty (or confidence coefficient) that is higher than the predetermined minimum.

Association rules are frequently generated using the Apriori algorithm. The Apriori algorithm was used in the study by Oladipupo et al. [41] to generate strong rules for predicting the impact of student attendance on educational outcomes. To find out whether students' attendance had any bearing on their grades, an association rule mining analysis was performed. According to the research, this shows that students' academic performance is not solely dependent on their ability to show up to class.

3.10 Nature-Inspired Algorithms

Nature-Inspired Algorithm is a type of artificial intelligence that is based on observation of biological systems in a specific context. To solve a variety of optimization problems, this observed cooperation and foraging among natural systems is adapted to find the optimal solution for a given situation, whether it be globally or locally [42]. There are many of these algorithms, but some of the most common ones include Genetic Programming (GP), Genetic Algorithm, Cuckoo Optimization Algorithm (COA), Particle Swarm Optimisation (PSO), and Ant Colony Optimization (ACO), among others.

These techniques are used in studies that predict students' academic performance include Chen, Hsieh, and Do (2014) which trained the feed forward neural network for student performance prediction using the standard COA and CS. Utilizing CS and COA, the weights between layers and biases of the neural network were optimized. RMSE, MAPE, and R obtained from simulations indicate that the ANN-COA algorithm outperforms the ANN-CS algorithm in terms of MAPE, R and RMSE [43].

Marquez-Vera et al., [44] predictions of student were made using an Interpretable Classification Rule Mining (ICRM) that relies on a GP variant, Grammar Based Genetic Programming (GBGP), as its foundation. The results of their various iterations of experiments revealed that the proposed method outperforms other traditional data mining techniques such as SMO, Naïve Bayes, and others in terms of prediction.

Hasheminijad & Sarvmili [45] used the Particle Swarm Optimization algorithm to propose a rule-based method called S3PSO to predict students' academic performances. The method extracts rules that are used to predict student performance using Association Rule Mining. When compared to other rule-based classification

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

algorithms such as CART, ID3, and C4.5, the results showed that S3PSO improved the fitness function by 31%. When compared to other data mining techniques such as SVM, KNN, and Naïve Bayes, it also improved accuracy by 9%.

3.11 Deep Learning

Deep Learning is a data mining technique that uses a 'deep' architecture to approximate complex functions to the same accuracy. ‘Deep’ is defined as using multiple layers with a smaller total number of neurons. The auto-encoders, which built in Deep Neural Network, are the basic building blocks of most deep learning models (DNN). Restricted Boltzmann Machines (RBM) are used as building blocks in Deep Generative models such as Deep Boltzmann Machine (DBM) and Deep Belief Network (DBN) [46].

According to [47] , used Convolutional Neural Network (CNN) which is a common deep learning model. Studies that used these techniques to predict students' performance are still scarce in the literature, but some examples include [48] , [49] , [50] , [51] , [52] , [53] , [54] , who used deep neural network to student prediction.

3.12 Hybrid Methods

The methods that combine two or more traditional data mining techniques to achieve a better result are known as hybrid methods of predicting student academic performance. Altaher & Barukab (2018) proposed a hybrid model for the prediction of SAP that combines an Adaptive Neuro-Fuzzy Inference System (ANFIS) with a Genetic Algorithm (GA). The ANFIS was fed a dataset of 100 records, and the training and testing results were then optimized using GA. The hybrid model, HGANFIS, was especially in comparison to Neural Network (NN) and ANFIS approaches using the RMSE as a performance metric. The HGANFIS model outperformed the other two algorithms, with RMSE values of 0.101 and 0.104 for training and testing data, respectively [55].

Chen, Feng, Sun, Wu, Yang, and Chen [56] In order to predict MOOC (Massive Open Online Course) performance, researchers combined the Decision Tree (DT) algorithm with the Extreme Learning Machine (ELM). The DT algorithm is used to extract important characteristics from MOOC students' learning behavior records. A total of eight algorithms were tested, including the DT, GA-ELM model, LR, SVM, BP (Backward Propagation Neural Network), ELM, LSTM and EN (Entropy-Net). The DT-ELM model significantly outperforms all others in terms of Accuracy (0.9642), F1-score (0.9667), and AUC (0.9412).

Francis & Babu [57] proposed a hybrid model to select the best characteristics from real-life datasets of students in different programs at universities and colleges using four algorithms (Naïve Bayes, SVM, NN and DT) in a wrapper feature selection mode. Students' educational performance was then predicted using the K-means clustering algorithm, which was based on a majority vote. A combination of behavioral, academic, and extra characteristics produced the best accuracy (0.7547) during model feature

ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

selection, and the model outperformed NN and DT in terms of recall, precision, F-scores, and accuracy values of 0.6415, respectively. The table below summarizes the common methods and factors used in student performance prediction.

Table 1 : Common methods and factors used in student performance prediction

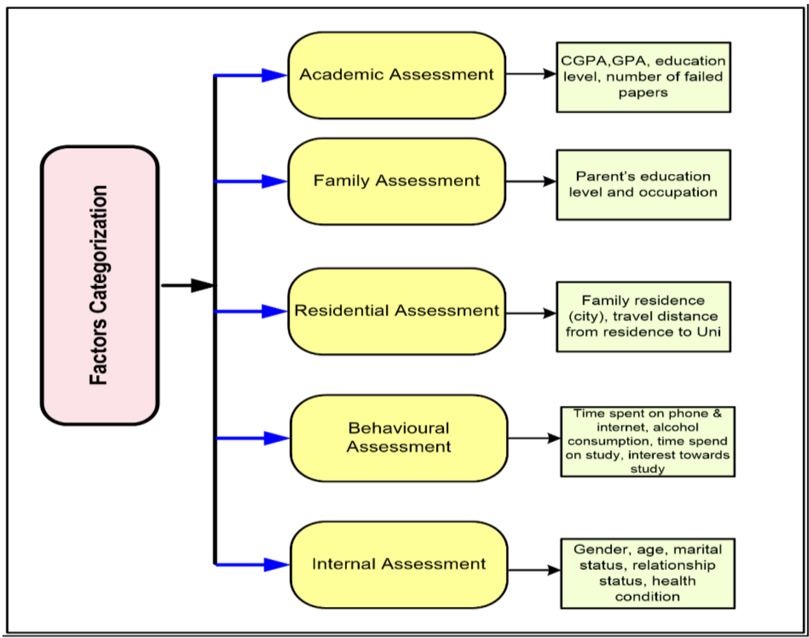
No.	METHODS	FACTORS
1	Decision Tree Neural Network K-Nearest Neighbor	Internal assessments
2	Support Vector Machine	Internal assessments, CGPA
3	Decision Tree Naïve Bayes K-Nearest Neighbor Support Vector Machine	Internal assessments, CGPA, Extra-curricular activities
4	Decision Tree Naïve Bayes K-Nearest Neighbor	Internal assessments, CGPA, Student Demographic
5	Neural Network	Internal assessments, External assessment
6	Neural Network Naïve Bayes	External assessment, Student Demographic, High school background
7	Decision Tree Neural Network K-Nearest Neighbor Support Vector Machine	Psychometric factors
8	Decision Tree Naïve Bayes Neural Network	External assessments
9	Decision Tree Neural Network Naïve Bayes	CGPA
10	Decision Tree Neural Network Naïve Bayes	CGPA, Student Demographic, High school background, Scholarship, Social network interaction
11	Decision Tree	CGPA, Student Demographic, High school background, Scholarship, Social network interaction, Internal assessments, Extra-curricular activities
12	Neural Network	Student Demographic, High school background
13	Decision Tree	External assessment, CGPA, Student Demographic, Extracurricular activities
14	Decision Tree	Psychometric factors, Extra-curricular activities, soft skills
15	Decision Tree	Student Demographic, High school background, Internal assessment, Student Demographic, Extra-curricular activities
16	Decision Tree Neural Network	Internal assessments, External assessment, Demographic, Extracurricular activities

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

4. DISCUSSION

The relationships between factors and classifiers were discussed in this section. Aside from that, the effects of various factors on various classifiers were investigated and discussed. The figure below is a simplified map of factor categorization.

Figure 2 : Simplified Map of Factor Categorization



The classification of the factors is shown in Table 2. The classification simplified the task of analyzing the impact of various factors on various classifiers. The findings are compiled in the table below based on the analysis of various articles from 2014 to 2021.

Table 2 : Classifier accuracy based on factors

Classifier	Author/Title	Factors	Accuracy
Naïve Bayes	[58]	Demographical Factors Academic, Academic Background Factors, Behavioral Factors, Parents Participation on learning	73%
	[59]	Family Factors, Student personal information	84.8%
	[60]	Activity log: Web page visited	63.8%
	[61]	Student profile data, GPA, Senior High School, and residence of student	69.07%
	[62]	The national exam score, the average exam score, the presence or absence, and the number of books read each month, GPA	70%
	[63]	The grade, period of study, school score and student score in elementary school	75.90%

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

	[64]	CGPA	75%
	[65]	Internal assessment, CGPA, Extra-curricular activities	73%
	[66]	Academic assessment	83.65%
Decision Tree	[58]	Demographical Factors Academic, Academic Background Factors, Behavioral Factors, Parents Participation on learning	69%
	[67]	Demographical Factors	83.14%
	[68]	Student's Academic Information ((CGPA), High Risk (student having high failure rate in the same module), Term Exceed at Risk, At Risk (student failed 2 or more modules previously), Student Success Center (SSC), Coursework 1 (CW1), Coursework 2 (CW2), End Semester Examination (ESE) and Plagiarism Count) Students Activity (the time spent by the student on Moodle in minutes)	63.63%
	[69]	Behavioral assessment, internal assessment, residential assessment, family assessment, academic assessment	73.92%
	[70]	Internal assessment, academic assessment, family assessment, residential assessment	72.50%
	[64]	CGPA	91%
	[65]	Internal assessment, CGPA, Extra-curricular activities	66%
	[19]	Courses Grades	79.50%
	[71]	Psychometric factors	65%
	[72]	Internal assessment, Student Demographic, Extra-curricular activities	90%
	[73]	External assessment, CGPA, Student Demographic, Extra-curricular activities	90%
	[40]	Psychometric factors, Extra-curricular activities, soft skills	88%
Neural Network	[48]	demographic and geographic Factors	79.82%
	[74]	Student's information, Assessments Marks	72.04%
	[53]	Behavioral	82.5%
	[52]	Internal assessment marks	95.34%
	[54]	Online learning activities	73.51%
	[75]	Student behavioral data, student demographic data, and student discussion posts	86.8%
	[76]	Activity Logs	80%
	[50]	Students' activity from log data	85%
	[49]	Internal Assessment Marks	95.34%
	[51]	Students' activity from log data	91.07%
	[58]	Demographical Factors Academic, Academic Background Factors, Behavioral Factors, Parents Participation on learning	84%

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

	[77]	Behavioral and student absent in class	78.60%
	[78]	learning logs of videos	87.2%
	[79]	Activity logs, and demographic Factors	85.4%
	[80]	Activity logs	90%
	[81]	Activity logs	84.6%
	[63]	The grade, period of study, school score and student score in elementary school	95%
	[82]	High school score, score of subjects, credits passed, GPA, types of high school attended and gender	84.60%
	[64]	CGPA	75%
	[71]	Psychometric factors	69%
Support Vector Machine	[48]	demographic and geographic Factors	79.95%
	[59]	Family Factors, Student personal information	86.7%
	[58]	Demographical Factors Academic, Academic Background Factors, Behavioral Factors, Parents Participation on learning	75%
	[83]	Internal assessment, academic assessment, family assessment, behavioural assessment	93.90%
	[84]	Academic assessment	87.50%
	[65]	Internal assessment, CGPA, Extra-curricular activities	80%

From the summary shown in Table 2, evidently NN has the highest prediction accuracy of 95%. The accuracy of the prediction result is highly dependent on the features that were used during the prediction process. Because of the influence of a particular factor during the running process, NN had the highest prediction accuracy. In this case, the determining factor is an academic assessment. Academic assessment is the main factor that has been used with NN. The prediction accuracy increased dramatically when only academic assessment was used. When behavioral and academic assessments were combined as inputs to make a prediction, however, the accuracy score dropped. The findings of [85] have increased the confidence in the findings by highlighting that the behavioral assessment decreased NN's accuracy level. This occurred because NN has trouble predicting qualitative data, which is non-numerical in nature. The second highest prediction accuracy was achieved by SVM, which was 93.95 percent. The success of SVM is due to its ability to solve high-dimensional data, which includes data with many attributes and features [86]. According to the findings, SVM can also perform well with only minor tweaking. Internal, academic, family, and behavioral assessments are the relevant factors that the classifier has used to make a prediction. In SVM, the accuracy level decreased when the academic assessment was used alone, contrary to NN. This suggests that academic assessment must be combined with other factors for SVM to achieve high accuracy.

Decision Tree and Naïve Bayes, on the other hand, are thought to be able to work with a wide range of factors and produce an average level of accuracy. According to the table above, Decision Tree and Naïve Bayes can work with both categorical and numerical data because the accuracy score of this classifier did not show any significant differences. When the academic assessment is the only factor used as input, however, the accuracy score rises. As a result, while Naïve Bayes and Decision Tree performed well with academic assessment factors, they produced an average accuracy score when used with other factors. Overall, the findings show that NN and SVM are the most competitive classifiers when compared to Naïve Bayes and Decision Tree and that academic assessment is the most significant factor in predicting students' performance.

5. CONCLUSION

This paper identified the data mining techniques used by previous studies in student prediction, and the factors used in the predication methods Students who do well academically in college are more likely to remain enrolled and ultimately earn a degree than their underperforming counterparts. If the predicition holds, university administrators will have a better idea of which students are unlikely to persist on their current course and can focus on helping those students boost their grades and eventually graduate. When taken as a whole, the results of this study revealed that Neural Networks could be used to predict how well college students would do in their courses. Many factors are used as inputs to these models, and they correlate well with criteria used in formal education settings. Students' past and present academic performance assessment factors were found to be strong predictors of their future academic performance. Last but not least, this work is significant because of its goal of aiding and helping other researchers in determining the factors and developing a genuine model that can easily and accurately predict students' academic performance. In addition, it will help teachers single out students who need extra help in class, allowing for more methodical and accurate prediction of student achievement outcomes.

References

1. J. R. Hillman and E. Baydoun, The Future of Universities in the Arab Region : A Review. 2018.
2. A. Soni, V. Kumar, R. Kaur, and D. Hemavathi, “PREDICTING STUDENT PERFORMANCE USING DATA MINING TECHNIQUES,” Int. J. Pure Appl. Math. Vol., vol. 119, no. 12, pp. 221–227, 2018.
3. D. Shingari, Isha Kumar, “International Journal of Computer Sciences and Engineering Open Access,” vol. 6, no. July 2018, 2020, doi: 10.26438/ijcse/v6i7.4348.
4. M. A. Al-Barrak and M. Al-Razgan, “Predicting Students Final GPA Using Decision Trees: A Case Study,” Int. J. Inf. Educ. Technol., vol. 6, no. 7, pp. 528–533, 2016, doi: 10.7763/ijiet.2016.v6.745.
5. F. Ahmad, N. H. Ismail, and A. A. Aziz, “The prediction of students’ academic performance using classification data mining techniques,” Appl. Math. Sci., vol. 9, no. 129, pp. 6415–6426, 2015, doi: 10.12988/ams.2015.53289.
6. K. David Kolo, S. A. Adepoju, and J. Kolo Alhassan, “A Decision Tree Approach for Predicting Students Academic Performance,” Int. J. Educ. Manag. Eng., vol. 5, no. 5, pp. 12–19, Oct. 2015, doi: 10.5815/ijeme.2015.05.02.

7. M. Pandey and S. Taruna, “Towards the integration of multiple classifier pertaining to the Student’s performance prediction,” Perspect. Sci., vol. 8, pp. 364–366, Sep. 2016, doi: 10.1016/j.pisc.2016.04.076.
8. M. P. G. Martins, V. L. Migueis, and D. S. B. Fonseca, “Uma Metodogia de Data Mining para Prever o Desempenho de Estudantes de Licenciatura,” Iber. Conf. Inf. Syst. Technol. Cist., vol. 2018-June, pp. 1–7, 2018, doi: 10.23919/CISTI.2018.8399175.
9. N. Al-Qaysi, N. Mohamad-Nordin, and M. Al-Emran, “A Systematic Review of Social Media Acceptance From the Perspective of Educational and Information Systems Theories and Models,” J. Educ. Comput. Res., vol. 57, no. 8, pp. 2085–2109, Jan. 2020, doi: 10.1177/0735633118817879.
10. B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, “Systematic literature reviews in software engineering - A systematic literature review,” Information and Software Technology, vol. 51, no. 1. pp. 7–15, Jan. 2009. doi: 10.1016/j.infsof.2008.09.009.
11. L. Respondek, T. Seufert, R. Stupnisky, and U. E. Nett, “Perceived academic control and academic emotions predict undergraduate university student success: Examining effects on dropout intention and achievement,” Front. Psychol., vol. 8, no. MAR, pp. 1–18, 2017, doi: 10.3389/fpsyg.2017.00243.
12. A. Daud, M. D. Lytras, N. R. Aljohani, F. Abbas, R. A. Abbasi, and J. S. Alowibdi, “Predicting student performance using advanced learning analytics,” 26th Int. World Wide Web Conf. 2017, WWW 2017 Companion, pp. 415–421, 2019, doi: 10.1145/3041021.3054164.
13. A. A. Aziz, N. H. Ismail, F. Ahmad, and H. Hassan, “A framework for students’ academic performance analysis using naïve bayes classifier,” J. Teknol., vol. 75, no. 3, pp. 13–19, 2015, doi: 10.11113/jt.v75.5037.
14. P. Kaur, M. Singh, and G. S. Josan, “Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector,” Procedia Comput. Sci., vol. 57, pp. 500–508, 2015, doi: 10.1016/j.procs.2015.07.372.
15. A. Mueen, B. Zafar, and U. Manzoor, “Modeling and Predicting Students’ Academic Performance Using Data Mining Techniques,” Int. J. Mod. Educ. Comput. Sci., vol. 8, no. 11, pp. 36–42, Nov. 2016, doi: 10.5815/ijmecs.2016.11.05.
16. G. Kaur and W. Singh, “Prediction Of Student Performance Using Weka Tool,” Res. Cell An Int. J. Eng. Sci., vol. 17, no. January, pp. 2229–6913, 2016.
17. S. Agrawal, S. K., and A. K., “Using Data Mining Classifier for Predicting Student’s Performance in UG Level,” Int. J. Comput. Appl., vol. 172, no. 8, pp. 39–44, Aug. 2017, doi: 10.5120/ijca2017915201.
18. S. Shalev-Shwartz and S. Ben-David, Understanding Machine Learning. Cambridge: Cambridge University Press, 2014. doi: 10.1017/CBO9781107298019.
19. Ogunde and Ajibade, “A Data Mining System for Predicting University Students’ Graduation Grades Using ID3 Decision Tree Algorithm Ogunde A. O 1 . and Ajibade D. A 1 .,” Comput. Sci. Inf. Technol., vol. 2, no. 1, pp. 21–46, 2014.
20. S. Joseph and L. Devadas, “Student’s Performance Prediction Using Weighted Modified ID3 Algorithm,” Int. J. Sci. Res. Eng. Technol., vol. 4, no. 5, pp. 2278–882, 2015, [Online]. Available: www.ijsret.org
21. B. O. Afeni, I. A. Oloyede, and D. Okurinboye, “Students’ Performance Prediction Using Classsification Algorithms,” J. Adv. Math. Comput. Sci., vol. 30, no. 2, pp. 1–9, Jan. 2019, doi: 10.9734/JAMCS/2019/45438.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

22. C. Burgos, M. L. Campanario, D. de la Peña, J. A. Lara, D. Lizcano, and M. A. Martínez, “Data mining for modeling students’ performance: A tutoring action plan to prevent academic dropout,” Comput. Electr. Eng., vol. 66, pp. 541–556, Feb. 2018, doi: 10.1016/j.compeleceng.2017.03.005.
23. L. Aulck and J. Blumenstock, “Predicting Student Dropout in Higher Education,” 2016, doi: https://doi.org/10.48550/arXiv.1606.06364.
24. A. S. Subaira and P. Anitha, “Efficient classification mechanism for network intrusion detection system based on data mining techniques: A survey,” 2014 IEEE 8th Int. Conf. Intell. Syst. Control Green Challenges Smart Solut. ISCO 2014 - Proc., pp. 274–280, 2014, doi: 10.1109/ISCO.2014.7103959.
25. S. S. Nikam, “A Comparative Study of Classification Techniques in Data Mining Algorithms,” Int. J. Mod. Trends Eng. Res., vol. 4, no. 7, pp. 58–63, Jul. 2017, doi: 10.21884/IJMTER.2017.4211.VXAYK.
26. B. Trstenjak and D. Donko, “Determining the impact of demographic features in predicting student success in Croatia,” in 2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), May 2014, no. May, pp. 1222–1227. doi: 10.1109/MIPRO.2014.6859754.
27. S. Bhagvatula, K. J. Shreyas, S. Gupta, and M. Singh, “Contribution of Students’ Academic Efforts on Prediction of Students’ Performance Using WEKA,” Int. J. Curr. Eng. Sci. Res., vol. 2, no. 11, pp. 83–89, 2015, [Online]. Available: troindia.in/journal/ijcesr/vol2iss11/83-89.pdf
28. H. Lakkaraju et al., “A machine learning framework to identify students at risk of adverse academic outcomes,” in Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Aug. 2015, vol. 2015-Augus, pp. 1909–1918. doi: 10.1145/2783258.2788620.
29. A. G. Mojisola, O. W. Samuel, M. O. Omisore, and B. A. Ojokoh, “A Multi-class Support Vector Machine Approach for Students Academic Performance Prediction,” Research Gate, vol. 4, no. March. pp. 210–215, 2016.
30. I. Burman and S. Som, “Predicting Students Academic Performance Using Support Vector Machine,” Proc. - 2019 Amity Int. Conf. Artif. Intell. AICAI 2019, pp. 756–759, 2019, doi: 10.1109/AICAI.2019.8701260.
31. A. Hernández-Blanco, B. Herrera-Flores, D. Tomás, and B. Navarro-Colorado, “A Systematic Review of Deep Learning Approaches to Educational Data Mining,” Complexity, vol. 2019, pp. 1–22, May 2019, doi: 10.1155/2019/1306039.
32. A. O. Osofisan, O. O. Adeyemo, and S. T. Oluwasusi, “Empirical Study of Decision Tree and Artificial Neural Network Algorithm for Mining Educational Database,” vol. 7, no. 2, pp. 187–196, 2014.
33. J. Ruby and K. David, “Analysis of Influencing Factors in Predicting Students Performance Using MLP -A Comparative Study,” Int. J. Innov. Res. Comput. Commun. Eng. (An ISO Certif. Organ., vol. 3297, no. 2, pp. 1085–1092, 2015, doi: 10.15680/ijircce.2015.0302070.
34. M. Mohammadi, M. Dawodi, W. Tomohisa, and N. Ahmadi, “Comparative study of supervised learning algorithms for student performance prediction,” 1st Int. Conf. Artif. Intell. Inf. Commun. ICAIIC 2019, pp. 124–127, 2019, doi: 10.1109/ICAIIC.2019.8669085.
35. R. Asif, A. Merceron, S. A. Ali, and N. G. Haider, “Analyzing undergraduate students’ performance using educational data mining,” Comput. Educ., vol. 113, pp. 177–194, Oct. 2017, doi: 10.1016/j.compedu.2017.05.007.

36. K. Verma, A. Singh, and P. Verma, “A review on predicting student performance using data mining method,” MMW Fortschr. Med., vol. 3, no. 1, pp. 127–132, 2016, [Online]. Available: http://queens.ezp1.qub.ac.uk/login?url=http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=medl&AN=23270045 http://resolver.ebscohost.com/openurl?sid=OVID:medline&id=pmid:23270045&id=doi:&issn=1438-3276&isbn=&volume=154&issue=21&spage=5&d

37. R. Gondane and V. S. Devi, “Classification using Probabilistic Random Forest,” 2015, doi: 10.1109/SSCI.2015.35.

38. A. Jain, K. Shah, P. Chaturvedi, and A. Tambe, “Prediction and Analysis of Student Performance using Hybrid Model of Multilayer Perceptron and Random Forest,” 2018 Int. Conf. Adv. Comput. Telecommun. ICACAT 2018, pp. 1–7, 2018, doi: 10.1109/ICACAT.2018.8933580.

39. N. Gilbert, “P REDICTING S UCCESS : A N A PPLICATION O F D ATA M INING T ECHNIQUES T O S TUDENT O UTCOMES,” vol. 7, no. 2, pp. 1–20, 2017, doi: 10.5121/ijdkp.2017.7201.

40. T. Mishra, “Mining Students ’ Data for Performance Prediction,” pp. 255–262, 2014, doi: 10.1109/ACCT.2014.105.

41. O. Oladipupo, O. Daramola, J. Oyelade, and ..., “assessing the impact of class attendance on student’s academic performance using data mining,” Asian J. Comput. …, 2016.

42. K.-L. Du and M. N. S. Swamy, Search and Optimization by Metaheuristics. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-41192-7.

43. J. F. Chen, H. N. Hsieh, and Q. H. Do, “Predicting student academic performance: A comparison of two meta-heuristic algorithms inspired by cuckoo birds for training neural networks,” Algorithms, vol. 7, no. 4, pp. 538–553, 2014, doi: 10.3390/a7040538.

44. C. Márquez-Vera, A. Cano, C. Romero, A. Y. M. Noaman, H. Mousa Fardoun, and S. Ventura, “Early dropout prediction using data mining: A case study with high school students,” Expert Syst., vol. 33, no. 1, pp. 107–124, 2016, doi: 10.1111/exsy.12135.

45. S. M. Hasheminejad and M. Sarvmili, “S3PSO: Students’ Performance Prediction Based on Particle Swarm Optimization,” J. Artif. Intell. Data Min., vol. 7, no. 1, pp. 77–96, 2018, doi: 10.22044/JADM.2018.5506.1662.

46. S. K. Zhou, H. Greenspan, and D. Shen, Deep Learning for Medical Image Analysis. 2017. doi: 10.4103/jpi.jpi_27_18.

47. C. C. Aggarwal, Neural Networks and Deep Learning. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-94463-0.

48. 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, 2020, doi: 10.1016/j.chb.2019.106189.

49. S. Hussain, Z. F. Muhsion, Y. K. Salal, P. Theodoru, F. Kurtoğlu, 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.

50. A. Yunita, H. B. Santoso, and Z. A. Hasibuan, “Deep Learning for Predicting Students’ Academic Performance,” in 2019 Fourth International Conference on Informatics and Computing (ICIC), Oct. 2019, pp. 1–6. doi: 10.1109/ICIC47613.2019.8985721.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

51. R. Raga and J. Raga, “Early prediction of student performance in blended learning courses using deep neural networks,” Proc. - 2019 Int. Symp. Educ. Technol. ISET 2019, pp. 39–43, 2019, doi: 10.1109/ISET.2019.00018.
52. S. Hussain, S. Gaftandzhieva, M. Maniruzzaman, R. Doneva, and Z. F. Muhsin, “Regression analysis of student academic performance using deep learning,” Educ. Inf. Technol., 2020, doi: 10.1007/s10639-020-10241-0.
53. 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.
54. X. Li, X. Zhu, X. Zhu, Y. Ji, and X. Tang, “Student Academic Performance Prediction Using Deep Multi-source Behavior Sequential Network,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12084 LNAI, H. W. Lauw, R. C.-W. Wong, A. Ntoulas, E.-P. Lim, S.-K. Ng, and S. J. Pan, Eds. Cham: Springer International Publishing, 2020, pp. 567–579. doi: 10.1007/978-3-030-47426-3_44.
55. A. Altaher and O. M. Barukab, “An Intelligent Hybrid Approach for Predicting the Academic Performance of Students using Genetic Algorithms and Neuro Fuzzy System,” IJCSNS Int. J. Comput. Sci. Netw. Secur., vol. 18, no. 10, p. 64, 2018.
56. J. Chen, J. Feng, X. Sun, N. Wu, Z. Yang, and S. Chen, “MOOC Dropout Prediction Using a Hybrid Algorithm Based on Decision Tree and Extreme Learning Machine,” Math. Probl. Eng., vol. 2019, pp. 1–11, Mar. 2019, doi: 10.1155/2019/8404653.
57. B. K. Francis and S. S. Babu, “Predicting Academic Performance of Students Using a Hybrid Data Mining Approach,” J. Med. Syst., vol. 43, no. 6, 2019, doi: 10.1007/s10916-019-1295-4.
58. 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.
59. N. R. Aljohani, A. Fayoumi, and S. U. Hassan, “Predicting at-risk students using clickstream data in the virtual learning environment,” Sustain., vol. 11, no. 24, pp. 1–12, 2019, doi: 10.3390/su11247238.
60. E. N. Azizah, U. Pujianto, and E. Nugraha, “Comparative performance between C4 . 5 and Naive Bayes classifiers in predicting student academic performance in a Virtual Learning Environment,” 2018 4th Int. Conf. Educ. Technol., no. 1, pp. 18–22, 2018.
61. A. Kesumawati and D. T. Utari, “Predicting patterns of student graduation rates using Naïve bayes classifier and support vector machine,” AIP Conf. Proc., vol. 2021, no. October 2018, 2018, doi: 10.1063/1.5062769.
62. U. Pujianto, E. N. Azizah, and A. S. Damayanti, “Naive Bayes using to predict students’ academic performance at faculty of literature,” Proceeding - 2017 5th Int. Conf. Electr. Electron. Inf. Eng. Smart Innov. Bridg. Futur. Technol. ICEEIE 2017, vol. 2018-Janua, pp. 163–169, 2018, doi: 10.1109/ICEEIE.2017.8328782.
63. R. M. De Albuquerque, A. A. Bezerra, D. A. De Souza, L. B. P. Do Nascimento, J. J. De Mesquita Sá, and J. C. Do Nascimento, “Using neural networks to predict the future performance of students,” 2015 Int. Symp. Comput. Educ. SIIE 2015, pp. 109–113, 2016, doi: 10.1109/SIIE.2015.7451658.

64. S. T. Jishan, R. I. Rashu, N. Haque, and R. M. Rahman, “Improving accuracy of students ’ final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique,” pp. 1–25, 2015, doi: 10.1186/s40165-014-0010-2.

65. M. Mayilvaganan and D. Kalpanadevi, “Comparison of classification techniques for predicting the cognitive skill of students in education environment,” 2014 IEEE Int. Conf. Comput. Intell. Comput. Res. IEEE ICCIC 2014, pp. 113–118, 2015, doi: 10.1109/ICCIC.2014.7238346.

66. R. Asif, A. Merceron, and M. K. Pathan, “Predicting Student Academic Performance at Degree Level: A Case Study,” Int. J. Intell. Syst. Appl., vol. 7, no. 1, pp. 49–61, 2014, doi: 10.5815/ijisa.2015.01.05.

67. S. Rizvi, B. Rienties, and S. Ahmed, “Computers & Education The role of demographics in online learning ; A decision tree based approach,” Comput. Educ., vol. 137, no. January, pp. 32–47, 2019, doi: 10.1016/j.compedu.2019.04.001.

68. R. Hasan, S. Palaniappan, A. R. A. Raziff, S. Mahmood, and K. U. Sarker, “Student Academic Performance Prediction by using Decision Tree Algorithm,” 2018 4th Int. Conf. Comput. Inf. Sci. Revolutionising Digit. Landsc. Sustain. Smart Soc. ICCOINS 2018 - Proc., pp. 1–5, 2018, doi: 10.1109/ICCOINS.2018.8510600.

69. S. Roy and A. Garg, “Predicting academic performance of student using classification techniques,” 2017 4th IEEE Uttar Pradesh Sect. Int. Conf. Electr. Comput. Electron. UPCON 2017, vol. 2018-Janua, pp. 568–572, 2017, doi: 10.1109/UPCON.2017.8251112.

70. C. Anuradha and T. Velmurugan, “A comparative analysis on the evaluation of classification algorithms in the prediction of students performance,” Indian J. Sci. Technol., vol. 8, no. 15, 2015, doi: 10.17485/ijst/2015/v8i15/74555.

71. G. Gray, C. McGuinness, and P. Owende, “An application of classification models to predict learner progression in tertiary education,” Souvenir 2014 IEEE Int. Adv. Comput. Conf. IACC 2014, pp. 549–554, 2014, doi: 10.1109/IAdCC.2014.6779384.

72. Elakia, Gayathri, Aarthi, and Naren J, “Application of Data Mining in Educational Database for Predicting Behavioural Patterns of the Students,” no. August, 2014.

73. S. Natek and M. Zwilling, “Student data mining solution–knowledge management system related to higher education institutions,” Expert Syst. Appl., vol. 41, no. 14, pp. 6400–6407, Oct. 2014, doi: 10.1016/j.eswa.2014.04.024.

74. F. Xiao, “Student Academic Performance Prediction with Recurrent Neural Network,” 2020.

75. J.-L. Hung, K. Rice, J. Kepka, and J. Yang, “Improving predictive power through deep learning analysis of K-12 online student behaviors and discussion board content,” Inf. Discov. Deliv., vol. 48, no. 4, pp. 199–212, May 2020, doi: 10.1108/IDD-02-2020-0019.

76. Y. He et al., “Online at-risk student identification using RNN-GRU joint neural networks,” Inf., vol. 11, no. 10, pp. 1–11, 2020, doi: 10.3390/info11100474.

77. M. Hasibur Rahman and M. Rabiul Islam, “Predict Student’s Academic Performance and Evaluate the Impact of Different Attributes on the Performance Using Data Mining Techniques,” 2nd Int. Conf. Electr. Electron. Eng. ICEEE 2017, no. December, pp. 1–4, 2018, doi: 10.1109/CEEE.2017.8412892.

78. N. F. Huang, C. C. Chen, J. W. Tzeng, T. Te Fang, and C. A. Lee, “Concept Assessment System Integrated with a Knowledge Map Using Deep Learning,” Proc. 2018 Learn. With MOOCS, LWMOOCS 2018, no. September, pp. 113–116, 2018, doi: 10.1109/LWMOOCS.2018.8534674.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 09-2022
DOI 10.17605/OSF.IO/CHJF2

79. A. Mondal and J. Mukherjee, "An Approach to Predict a Student's Academic Performance using Recurrent Neural Network (RNN)," Int. J. Comput. Appl., vol. 181, no. 6, pp. 1–5, 2018, doi: 10.5120/ijca2018917352.

80. F. Okubo, A. Shimada, T. Yamashita, and H. Ogata, "A neural network approach for students' performance prediction," ACM Int. Conf. Proceeding Ser., pp. 598–599, 2017, doi: 10.1145/3027385.3029479.

81. F. Okubo, T. Yamashita, A. Shimada, and S. Konomi, "Students' performance prediction using data of multiple courses by recurrent neural network," Proc. 25th Int. Conf. Comput. Educ. ICCE 2017 - Main Conf. Proc., pp. 439–444, 2017.

82. S. A. Naser, I. Zaqout, M. A. Ghosh, R. Atallah, and E. Alajrami, "Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology," Int. J. Hybrid Inf. Technol., vol. 8, no. 2, pp. 221–228, 2015, doi: 10.14257/ijhit.2015.8.2.20.

83. V. L. Miguéis, A. Freitas, P. J. V. Garcia, and A. Silva, "Early segmentation of students according to their academic performance: A predictive modelling approach," Decis. Support Syst., vol. 115, pp. 36–51, 2018, doi: 10.1016/j.dss.2018.09.001.

84. S. K. Pushpa, T. N. Manjunath, T. V. Mrunal, A. Singh, and C. Suhas, "Class result prediction using machine learning," Proc. 2017 Int. Conf. Smart Technol. Smart Nation, SmartTechCon 2017, vol. 6, pp. 1208–1212, 2018, doi: 10.1109/SmartTechCon.2017.8358559.

85. A. M. Shahiri, W. Husain, and N. A. Rashid, "A Review on Predicting Student's Performance Using Data Mining Techniques," Procedia Comput. Sci., vol. 72, pp. 414–422, 2015, doi: 10.1016/j.procs.2015.12.157.

86. X. Ma and Z. Zhou, "Student pass rates prediction using optimized support vector machine and decision tree," in 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Jan. 2018, vol. 2018-Janua, pp. 209–215. doi: 10.1109/CCWC.2018.8301756.