Examining the potential of machine learning for predicting academic achievement: A systematic review

M. Nazir [1], A. Noraziah [2*], M. Rahmah [3], Aditi Sharma[4]

[1,2,3]Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia

[4]IEEE Senior Member, Department of Computer Science and Engineering, Astana IT University, Astana, Kazakhstan

Emails: pcp21002@student.umpsa.edu.my; noraziah@umpsa.edu.my; drrahmah@umpsa.edu.my; aditi.sharma@ieee.org

Abstract

Predicting student academic performance is a critical area of education research. Machine learning (ML) algorithms have gained significant popularity in recent years. The capability to forecast student performance empowers universities to devise an intervention strategy either at the beginning of a program or during a semester, which allows them to tackle any issues that may arise proactively. This systematic literature review provides an overview of the present state of the field under investigation, including the most commonly employed ML techniques, the variables predictive of academic performance, and the limitations and challenges of using ML to predict academic success. Our review of 60 studies published between January 2019 to March 2023 reveals that ML algorithms can be highly effective in predicting student academic performance. ML models can analyse various variables, including demographics, socioeconomic status, and academic history, to identify patterns and relationships that can predict academic performance. However, several limitations need to be addressed, such as the inconsistency in the variables used, small sample sizes, and the failure to consider external factors that may impact academic performance. Future research needs to address these limitations to develop more robust prediction models. Machine learning can fuse data from various sources like test scores like Coursera, edX & Open edX, Udemy, linkedin learning, learn words, and hacker's rank platform etc, attendance, and online activity to help educators better understand student needs and improve teaching, can use for better decision. In conclusion, ML has emerged as a promising approach for predicting student academic performance in online learning environments. Despite the current limitations, the continued refinement of ML techniques, the use of additional variables, and the incorporation of external factors will lead to more robust models and greater accuracy in predicting academic performance.

Keywords: Student Performance Prediction; Online Courses; Machine Learning; Systematic Literature Review; Machine learning for data fusion; Fusion in Decision-making

1. Introduction

The assessment of academic performance is a crucial area of education research. Predicting academic performance is a fundamental topic that always has interested education researchers. Student academic performance is influenced by various factors such as learning skills, peer influence, teacher quality, and infrastructure [Han and Ellis, 2021] [1]. Traditional methods used by educators, such as past academic records, teacher evaluations, and standardized test scores, have limitations in providing a comprehensive picture of a student's potential [Kavitha et al., 2021].

There is a growing need to identify students at risk of academic failure and provide appropriate interventions to ensure future academic success. Exploring different teaching modalities and conducting in-depth research on student perspectives is necessary to improve educational activities. Researchers have consistently demonstrated that various teaching modalities positively correlate with compelling student learning experiences. Moreover, adding computer-mediated components can enhance student achievement [Kavitha et al., 2021] [2].

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

Educational data mining has significantly impacted the education industry [Zhang et al., 2021] [3]. It has identified new possibilities for technologically advanced learning systems that promote a personalized learning environment [Dabhade et al., 2021] [4]. Machine learning is growing as an intriguing strategy to predict academic performance, leveraging large-scale data sets and sophisticated algorithms to identify factors influencing academic success. By analyzing data, machine learning algorithms can identify patterns and trends, which can be used to personalize and improve learning outcomes [Du et al., 2020] [5].

[6] Data analytics refers to methods used to analyze different kinds of data, including identifying hidden patterns, finding relationships, and optimizing with new information [Albreiki et al., 2021]. Exploring features such as student demographic data and behavior can enhance student performance to generate the required results [Huynh-Ly et al., 2021]. Data mining and machine learning advancements in recent years have paved the way for predicting academic performance [7].

In conjunction with offline learning, machine learning algorithms can identify areas where students require improvement and optimize academic performance [Chen and Zhai, 2023] [8]. To predict academic performance, machine learning models can analyze various variables, including demographics, socioeconomic status, and intellectual history [Yang et al., 2021] [9]. Moreover, machine learning algorithms can help educators track student progress over time, providing insight into which interventions are effective and which need to be adjusted [Bao et al., 2020] [10].

This systematic literature review aims to synthesize the existing research on student academic performance prediction using machine learning techniques. The study will examine the field's current state, the most commonly used strategies, the variables that predict academic performance, and the limitations and challenges of using machine learning to predict academic success.

The primary aims of this systematic literature review are as follows:

- To examine the current state of research on predicting student academic performance using machine learning algorithms.
- To identify the most commonly employed machine learning techniques for predicting student academic performance.
- To determine the variables identified as predictive of academic performance using machine learning algorithms.
- To identify the limitations and challenges of using machine learning to predict academic success, including inconsistent variables, small sample sizes, and failure to consider external factors.
- To provide recommendations for future research to address the identified limitations and develop more robust prediction models [Alyahyan and Düştegör, 2020] [11].
- To assess the effectiveness of machine learning algorithms in predicting student academic performance in various learning environments.

To explore the potential impact of the continued refinement of machine learning techniques, the use of additional variables, and the incorporation of external factors on the accuracy of predicting academic performance.

The structure of the remaining systematic literature review is as follows: Part II will present the research questions and methods for reviewing the literature. In contrast, Section III will present the findings of the literature review. Section IV will include a discussion, and Section V will conclude with an outlook for the future.

2. Methodology

The methodology for conducting the review is discussed in this section. This study adhered to the PRISMA guidelines for conducting and presenting a systematic literature review [Page et al., 2021] [12] following the PRISMA declaration. The data retrieval process was also completed in line with these guidelines. Figure 1 depicts the entire systematic literature review (SLR) process, which consisted of two phases: planning and conducting. The planning phase involved defining research objectives, devising a research plan, and identifying relevant literature. In the conducting step, we executed the research plan, which included identifying and selecting high-quality

publications, extracting and synthesizing data, and analyzing the results. The review was conducted using the widely recognized PRISMA statement, and the flow diagram is illustrated in Figure 1. The conducting phase comprised the following steps:

2.1. Inclusion criteria

The Studies that use machine learning algorithms to predict student academic performance.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

- Studies that employ any of the following machine learning techniques: regression analysis, decision trees, random forests, support vector machines, or neural networks.
- Studies that examine the relationship between specific variables and academic performance, such as demographic characteristics, socioeconomic status, intellectual history, or previous academic achievements.
- Studies published in high-quality journals systematic literature review utilized the following inclusion criteria:
- Studies published between 2019 to 2023.
- Studies that provide sufficient details about the dataset used for training and testing the machine learning models.
- Studies that provide evaluation metrics to measure the performance of machine learning models.
- Studies that discuss the limitations and challenges of using machine learning to predict academic success.

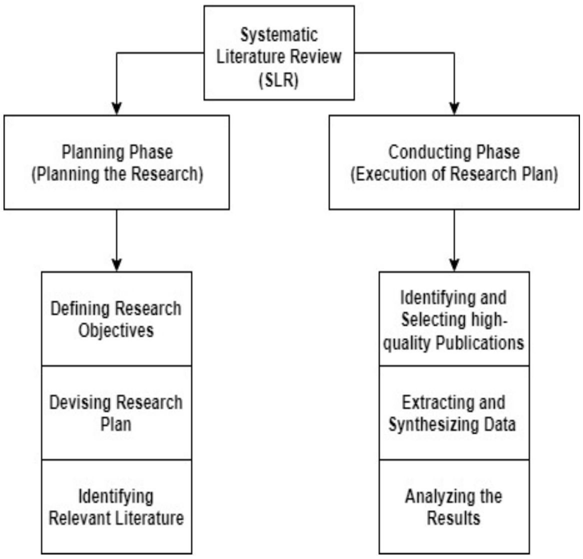

Figure 1: Phases of Systematic Literature Review

2.2. Exclusion Criteria:

- Studies that did not use machine learning techniques for predicting student academic performance.
- Studies that used traditional statistical methods, such as linear or logistic regression, instead of machine learning techniques.
- Studies focused on predicting non-academic outcomes, such as mental health or social behaviour, rather than academic performance.
- Studies that did not report quantitative results or did not provide enough information to calculate effect sizes.
- Studies that were not published in English.
- Studies published before 2019.
- Studies that were not peer-reviewed or were published in non-reputable sources.
- Studies needed to include more information about the machine learning algorithms, such as hyperparameters and feature selection methods.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

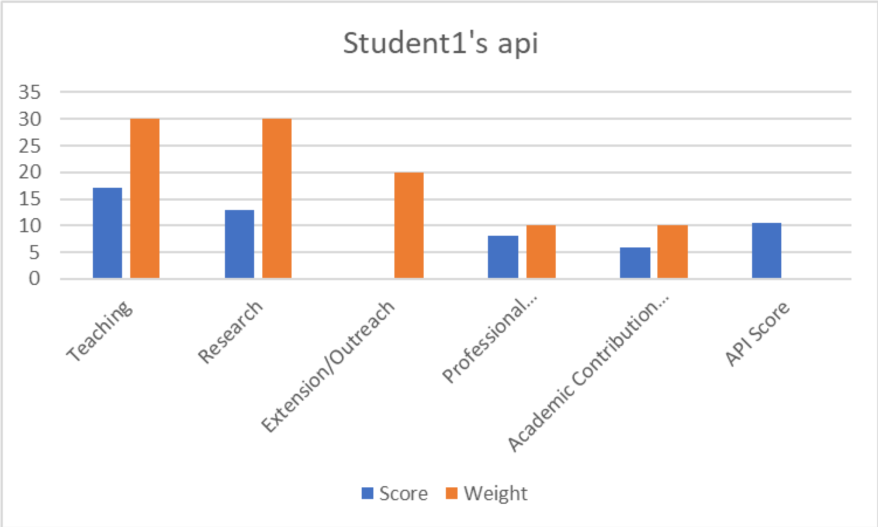

Figure 2: Student1's API

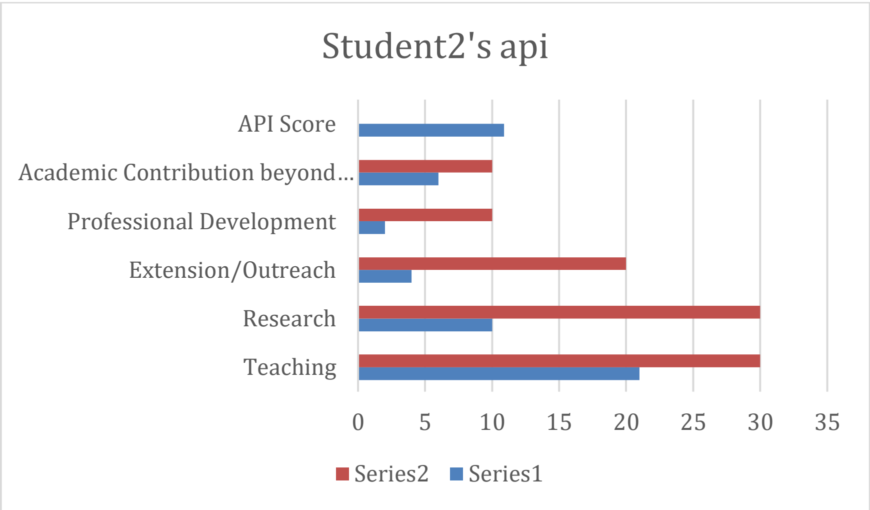

Figure 3: Student2's API

2.3. Data Extraction:

Figure 3 represents the list of journals which are included for our SLR.

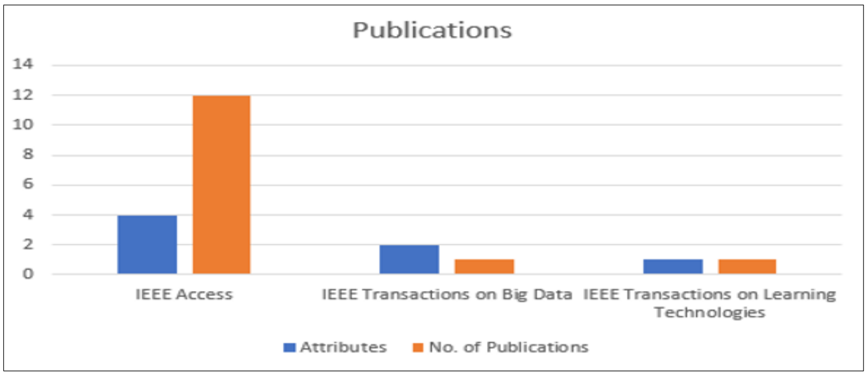

Figure 4: Attributes for performance prediction used by authors in the study period

We extracted the following data from each study: Study ID: Title: Authors: Year of publication: Publication type: Aim/Objective: Study design: Machine learning technique(s) used: Variables used for prediction: Data source(s): Sample size: Performance measures: Key findings: Limitations of the study: Implications for future research.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

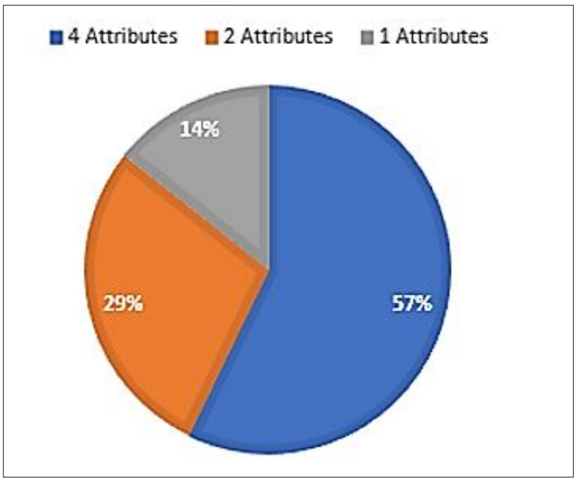

Figure 5: Percentage distribution of studies using one or more student attributes

2.4. Data Analysis:

We created a database to store relevant information of 38006 students extracted from the included studies. Academic Performance Indicator on 38007 students is shown in Figure 5. We extracted data on the author(s), publication year, country, sample size, machine learning algorithm used, variables used for prediction, performance metrics reported, and any limitations or challenges identified.

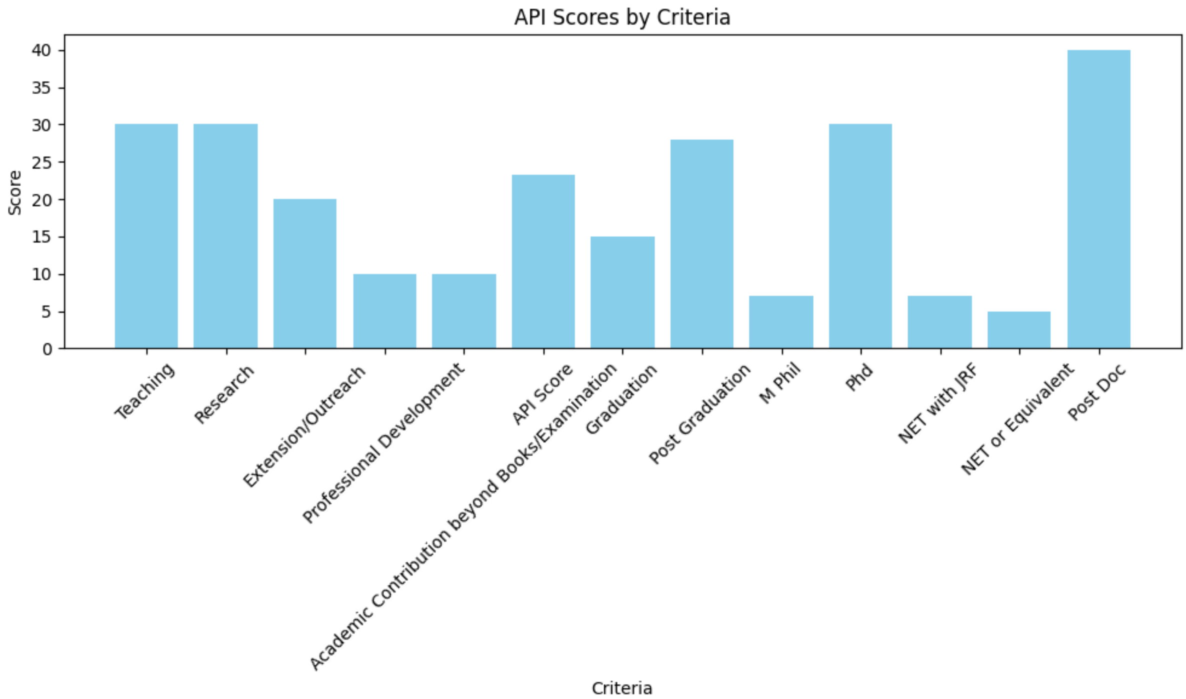

Figure 6: Academic Performance Indicator on 38007 students

We organized the extracted data into a table or spreadsheet format for more accessible analysis.
To assess the quality of the included studies, we used established quality assessment tools or criteria, including the study design, sample size, data collection methods, and potential sources of bias when evaluating study quality. Studies with poor quality or a high risk of bias are excluded from the analysis.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

```
>>>
======= RESTART: G:\research27-09-2023\usgapiscore\confudionmatrixmcc.py =====
Accuracy: 1.0
Accuracy: 0.7662907158267702
Confusion Matrix:
[[     0 13324]
 [     0 43687]]
Matthews Correlation Coefficient (MCC): 0.0
>>>
```

Figure 7: accuracy, confusion-matrix, and MCC

We summarize the characteristics of the included studies, such as the number of studies, sample size, and geographic location, and categorize the studies based on the machine learning algorithm used, variables used for prediction, and performance metrics reported. We analyzed the reported performance metrics, such as accuracy, precision, recall, and F1 score, to determine the effectiveness of different machine learning algorithms and variable selection methods for predicting student performance. We also identified common limitations or challenges across the included studies, such as inconsistent variable selection small sample sizes, or failure to consider external factors. Ultimately, we develop recommendations for future research to address the identified limitations and improve the accuracy of machine learning-based student performance prediction models.

2.5. Research Questions:

The literature review aims to answer the following research questions:

Research Questions	Details
RQ1	What is the current state of research on predicting student academic performance using machine learning algorithms?
RQ2	Which machine learning techniques are commonly employed for predicting student academic performance?
RQ3	What variables have been identified as predictive of academic performance using machine learning algorithms?
RQ4	What are the most commonly used dataset repositories?
RQ5	What are the limitations and challenges of using machine learning to predict academic success, including inconsistent variables, small sample sizes, and failure to consider external factors?
RQ6	What recommendations can be made for future research to address the identified limitations and develop more robust prediction models?

3. Literature Survey analytics

The study by [Alhazmi and Sheneamer, 2023] aimed to predict the academic performance of higher education students. The authors collected data from 5000 students at a university in Saudi Arabia, which included demographic information and academic performance data from their first year. They used multiple machine learning algorithms to develop predictive models for student performance, and they found that the random forest algorithm performed the best, with an accuracy rate of 85%. The study also identified the most critical factors that affect academic performance, such as high school grades, gender, and age. The authors suggest that predictive models can be helpful in identifying students at risk of poor academic performance early on, enabling educational institutions to provide targeted support to help these students succeed [13].

The article by [Nabil et al., 2021] proposes a machine-learning model for predicting and interpreting student performance. The model employs ensemble learning techniques such as Random Forest and AdaBoost algorithms to forecast student performance based on input features. The authors also utilize Shapley Additive Explanations (SHAP) to provide transparency into the model's decision-making process by pinpointing the input features that

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

most impact predicting student performance. The study's results reveal that the proposed model outperforms other models with 98% accuracy and provides meaningful insights into the factors influencing student performance [14].

The authors [Feng et al., 2022] [15] investigate the use of educational data mining techniques to analyze and predict the academic performance of students. They gathered data from a university's information management system and employed data mining techniques to uncover patterns and correlations in the data. The findings revealed that factors such as attendance, past academic performance, and demographic information were significant predictors of future academic performance. The authors created a prediction model using these factors and achieved high levels of accuracy in predicting student grades, with accuracy rates of 95.7%, 98.4%, and 98.0%.

The research conducted by [Rafique et al., 2021] aimed to enhance students' academic performance by combining collaborative learning with learning analytics. The study utilized machine learning algorithms such as LR, KNN, SVM, NB, Decision Tree, and Ensemble models to analyze the personal, socioeconomic, psychological, and academic-related factors of 164 students from COMSATS University. The study assessed the effectiveness of collaborative learning techniques in enhancing student achievement, providing prompt feedback on student performance, and used accuracy, F1-Score, and Recall as performance metrics [16].

The research study by [Alwarthan et al., 2022] characterizes a machine learning model that determines at-risk organization students' responsibility. The study possessed data on different spects such as student statistics, educational acquisition, and online behavior to predict which recruitment is at exposure to poor performance. Results showed that appearance, past academic administration, and accessible learning behavior were powerful predictors of at-risk acceptance. The authors refined an intelligible ML model to arrange observation into the factors leading to student underperformance. This helps educators take precautionary amplification to improve student academic performance [17].

The study by [Alshanqiti and Namoun, 2020] recommended a cross-ML approach to conclude student academic achievement and classify the component that influences it. The study possessed student demographics, scholarly background, and accessible cultural behavior. Regression and classification techniques were utilized to predict students' grades and identify factors contributing to academic success [18]. The hybrid machine learning model developed by the authors accurately predicted students' grades and identified the factors contributing to academic success. The study highlights the potential of a hybrid machine learning approach in predicting students' academic performance and identifying the factors that influence it. The study's findings could assist educators in designing targeted interventions and support systems to improve students' academic outcomes.

The study by [Hung et al., 2019] used LR, decision tree, RF, and Gradient boosting algorithms to predict the performance of at-risk students. The study analyzed data from 11,000 students who participated in an online learning platform between 2012 and 2016 [19]. The dataset included demographic information, such as gender and ethnicity; academic histories, such as GPA and credits earned; and course-specific variables, such as attendance, participation, time spent on the online learning platform, and the number of posts and responses made in discussion forums. The study used evaluation metrics, including sensitivity, specificity, precision, recall, F1 score, and the area under the receiver operating characteristic curve. The study found that academic variables, such as GPA, course grades, and attendance, were more predictive of at-risk status than demographic variables, such as age, gender, and ethnicity. The study achieved an accuracy of 86.3% for predicting student performance.

The study by [Adnan et al., 2021] used six artificial intelligence models to identify vulnerable pupils based on socioeconomic, academic performance, and program-specific characteristics [20]. These models were DT, RF, SVM, KNN, LR, and MLP. The study's dataset included 15,200 individuals enrolled in 24 distance learning courses across 4 subjects. The models' performance was assessed using a variety of criteria, including accuracy, precision, recall, the F1- score, and AUC-ROC. The results indicate that the suggested model employs machine learning models and can properly predict at-risk students at multiple points throughout the course. The study also emphasized the need to choose proper evaluation measures and establish acceptable response standards to implement successful early interventions to help at-risk pupils.

The paper by [Ahajjam et al., 2022] [21] proposed a novel approach to forecast students' final performance by applying Artificial Neural Networks (ANNs). The model's efficacy was tested using a dataset containing information on 494 students. The ANNs were trained using features such as the frequency of absences, course page visits, forum messages, and quiz attempts. The model's performance was evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) as metrics. The results indicated that the ANNs outperformed traditional regression models regarding lower MAE and RMSE values. Hence, the study concluded that ANNs could serve as an effective tool to predict students' final performance and offer timely interventions to support struggling students.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

The study by [Pek et al., 2022] utilized six machine learning algorithms, including decision tree, logistic regression, random forest, K-nearest neighbors, Naïve Bayes, and support vector machines [22]. The dataset used in the study consisted of information on 555 students and included 38 features related to demographic and academic information. These features included school name, kindergarten attendance status, travel time to school, university plans, GPA of semesters, gender, age, city, district, number of siblings, and free time. The study evaluated the performance of the models using metrics such as accuracy, precision, and recall. The results showed that the hybrid model trained on academic data achieved an accuracy of 98.8%, indicating that academic data is valuable for training such models.
The paper by [Bujang et al., 2021] [23] proposes a multiclass prediction model for student grade prediction using machine learning. The study aimed to predict students' grades in multiple classes using six machine learning algorithms: Decision Tree, Random Forest, Support Vector Machine, Naive Bayes, K-Nearest Neighbor, and Artificial Neural Network. The dataset included demographic variables like age, gender, and race, as well as academic variables such as attendance, homework scores, and exam scores. The models' performance was evaluated using accuracy (99.5%), precision, recall, and F1-score. According to the study's results, the proposed multiclass prediction model using machine learning accurately predicted student grades. This model can be beneficial in identifying students who require additional support and intervention.

The paper by [Liu et al., 2020] proposed a new deep knowledge tracing model for predicting student performance by fusing multiple features with an attention mechanism [24]. The suggested model merges sequential modeling of student performance and feature modeling by integrating previous knowledge states and features like exercise types, difficulty levels, and student behaviors in the learning process. A vast dataset from an online learning platform was used to evaluate the model. The findings revealed that the proposed model surpassed the conventional deep knowledge tracing model regarding forecasting accuracy. The study also emphasized integrating appropriate features and attention mechanisms for precise and efficient student performance prediction in e-learning settings.

The article by [Moreno-Marcos et al., 2020] analyzes the factors influencing the accuracy of learners' performance prediction using learning analytics [25]. The study concentrates on two courses utilizing data from the online learning platform edX. Various data mining and machine learning methods, including correlation analysis, decision trees, and random forests, were utilized by the authors to scrutinize the data. The study recognizes several influential factors in forecasting learners' performance, such as engagement, effort, and prior knowledge Additionally, the authors emphasize the significance of feature selection in enhancing the precision of performance prediction models. In summary, the study furnishes insights into the factors that can enhance the accuracy of performance prediction models in learning environments.

[3] The paper by [Zhang et al., 2021] proposes a multisource sparse attention convolutional neural network (MSCNN) model for predicting and understanding student learning performance. The model uses course-level, video-level, and interaction-level data as input features. The data utilized in the study are classified into course-level, video-level, and interaction-level data. Course-level data includes course codes, instructors, and enrollment figures. Video-level data encompasses video watch time, length, and the number of video segments watched. Interaction-level data involves information on students' interactions with the learning management system, including login frequency, forum posts, and quizzes taken. The model's effectiveness was evaluated on a dataset of 21,479 students and yielded an accuracy rate of 85.7% in predicting student performance. Additionally, the paper investigates the significance of various features and concludes that interaction-level features are the most critical in predicting student performance. Lastly, the study provides insights into the learning habits of high- and low-performing students, aiding instructors in identifying at-risk students and developing appropriate interventions.

The research paper by [Zhao et al., 2021] explores the use of decision trees, LSTM, and Gradient Boosting Decision Tree (GBDT) machine learning algorithms for academic performance prediction [27]. The study used a dataset comprising demographic information, academic records, course information, assignment submissions, online activity, and library usage data from the Freshman Seminar in 2018-2019. Accuracy served as the evaluation metric. The results indicated that combining behavioral and demographic features could enhance the precision of theoretical performance predictions, surpassing the usage of behavioral or demographic features alone.

The article by [Wang et al., 2023] presents a comprehensive and high-performance system called ProbSAP for student academic performance prediction [28]. The system leverages probabilistic graphical models and deep learning techniques to forecast student performance, drawing on demographic information, academic records, course details, and behavioral data to ensure precise predictions. The system's effectiveness was assessed using an extensive dataset comprising information on thousands of students. The results show that ProbSAP outperforms with 90.4% accuracy from state-of-the-art methods regarding accuracy and efficiency. The article concludes that ProbSAP has the potential to be used in real-world educational settings to improve student outcomes.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

The research paper by [Zeineddine et al., 2021] presents an approach for enhancing the prediction of student success using automated machine learning (AutoML) [29]. The method put forward in the research employs a dataset containing student demographic details, prior academic achievements, and additional factors related to their participation in online learning systems. The AutoML Approach automatically selects machine learning algorithms and feature engineering techniques to construct a model for predicting student success. The study tested the suggested method using data from two different universities. The results showed that the AutoML Approach outperformed several baseline models regarding accuracy, precision, recall, and F1-score. The study also compared the proposed approach with a traditional machine-learning approach that used manual feature selection and hyperparameter tuning. The results showed that the AutoML Approach achieved higher accuracy with fewer features.

The paper by [Asselman et al., 2020] [33] discusses the impact of prior required scaffolding items on improving student performance prediction. The research utilized information from an online learning platform that monitored student engagement with the platform, and machine learning algorithms was employed to predict their academic performance. The dataset contained information such as the student's age, gender, previous academic performance, and results from a pre-course test. The authors observed that incorporating prior required scaffolding items improved the accuracy of models by 71.51%, 65.86%, and 72.29%, respectively, for student performance prediction. The study suggests that including prior required scaffolding items as a feature in performance prediction models can help educators to identify struggling students earlier and provide appropriate interventions to improve their learning outcomes.
The research study by [Bilal et al., 2022] [30] aimed to predict student performance based on demographic and academic features. The study used six supervised machine learning models: RG, RF, DT, KNN, SVM, and ANN. The dataset comprised 1,075 undergraduate students from a Pakistani institution, and demographic features including age, gender, country, and high school type. In contrast, academic features included CGPA, the field of study, semester pressure, and the number of classes taken. The evaluation metrics used were 15-fold cross-validation, accuracy, precision, and recall. The results indicated that SVM had the highest accuracy, achieving 92% in predicting student performance.

The article by [Wang et al., 2022] utilized a Complexity-based attentive, interactive neural network (CAINN) to predict the academic performance of undergraduate students based on several factors, including student academic records, course records, course characteristics, and student characteristics. The dataset consisted of information from 580 undergraduate students at a Chinese university [31]. The evaluation metrics used in the study were MSE and MAE, and the proposed approach demonstrated superior performance compared to conventional machine learning techniques, achieving a precision rate of 91.8%. Furthermore, the authors identified significant factors influencing student achievements, such as previous academic performance, course grades, and difficulty.

The study by [Arashpour et al., 2023] utilizes a hybrid approach of machine learning and teaching-learning-based optimization (TLBO) to forecast individual learning performance. The models employed in this study are Support Vector Machines (SVM) and Artificial Neural Networks (ANN) [32]. At the same time, the predictive variables encompass the gender, age, disability, prior qualifications, and the Index of Multiple Deprivations. The Open University Analytics dataset, with 15,931 records, is used in the study. The evaluation metrics adopted for the research include accuracy, precision, recall, F1-Score, MCC, FM, confusion matrix, and ROC-AUC. The study concludes that combining machine learning and TLBO can enhance the justification of student test performance. Furthermore, the SVM model's accuracy was superior to the ANN model, with almost 92.27% accuracy on all data.

4. Discussion

Table 1 provides an overview of various research studies focused on predicting student academic performance using a range of features, techniques, and datasets. Reading from Table 2 above, it can be seen that each study is identified by its author(s) and publication year, making it easy to reference and locate the original research. The objectives of the studies vary, from predicting academic performance to identifying at-risk students, understanding learning behaviours, and more. This diversity in objectives reflects the multifaceted nature of academic performance prediction research. The features used in these studies encompass a wide range, including academic data, demographic information, behavioural factors, and exercise-related data.

Table 1: Parametrial Performance Findings for various Literature

Ref.	**Title**			**Sample Source**		**Findings**

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

		Machine learning technique(s)	Variables for prediction	and Size	Performance measuring metrics	
[Alhazmi and Sheneamer, 2023] [13]	"Early Predicting of Student's Performance in Higher Education"	Xgboost, Logistic Regression, SVM KNN and RF	"Admission scores, Admission scores and gender, Admission scores with all first-level courses scores."	Five thousand students as the sample size.	Prediction accuracy, Precision, Recall and F1-score	The random forest algorithm performed the best, achieving an accuracy of 85%. High school grades, gender, and age were critical in predicting student performance.
[Nabil et al., 2021] [14]	"Predicting and Interpreting Student Performance Using Ensemble Models and Shapley Additive Explanations"	RF, BC, KNC, NB Decision Tree, XGB and Neural Networks	"Nationality, Gender, Place of birth, Relation, Stage ID, Grade ID, Section ID, Semester, Topic, Student absence days, Parent answering survey, Parent school satisfaction, and Discussion"	The sample consists of 480 students' data with 17 features from Kaggle	Accuracy, Precision, and Recall	Findings show the effectiveness of ensemble learning with 98% of accuracy. SHAP for equal measuring of student performance in educational institutions.
[Feng et al., 2022] [15]	"Analysis and Prediction of Students' Academic Performance Based on Educational Data Mining"	K-means and CNN	Good, Excellence, and Poor ranks,	Three types of datasets consistin52,66 and 51 sample	Shuffle 5-fold cross-validation, precision, and recall	Help educators provide targeted interventions to students. Using the top percentile will accelerate general growth.
[Rafique et al., 2021] [16]	"Integrating Learning Analytics and Collaborative	LR, KNN, SVM, NB, Decision Tree, and Ensemble model	Personal, socioeconomic, psychological, and	164 from COMSATS University	Accuracy, F1-Score and Recall	The effectiveness of the collaborative learning technique shows improvements in

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

	Learning for Improving Students Academic Performance”		academic-related variables.			student achievement. It provided real-time feedback on students' performance.
[Alwarthan et al., 2022] [17]	“An Explainable Model for Identifying At-Risk Students in Higher Education”	SVM, RF, and ANN	“Demographic data, pre-university data, courses grades, CGPA, and assessments marks (detailed grades) for each course”.	Student Information System (SIS), Learning Management System (Blackboard), and Deanship of Preparatory Year and Supporting Studies. Students at the humanity track study six courses in the first semester and seven courses in the second semester.	Testing Accuracy, Precision, Recall and F1-score	Identifying at-risk students and providing personalized interventions to improve their academic outcomes. The findings could help educators design targeted interventions and support systems to improve success.
[Alshanqiti and Namoun, 2020] [18]	“Predicting Student Performance and Its Influential Factors Using Hybrid	Hybrid Regression Models and Neural Networks	Academic grades, GPA, gender, age, and exam scores.	Dataset of 3820 students' demographic information, academi	Mean Absolute Error, Mean Squared Error, Root Mean Squared Error, Precision,	A hybrid machine learning approach could be useful for predicting students' academic performance.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

	Regression and Multi-Label Classification"			c history, and online learning behavior.	Recall, and F1-score.	The findings could help educators design targeted interventions and support systems to improve students' academic outcomes.
[Hung et al., 2019] [19]	"Improving Predictive Modeling for At-Risk Student Identification: A Multistage Approach"	LR, Decision tree, RF, and Gradient boosting algorithms	Demographic information such as gender and ethnicity, academic histories such as GPA and credits earned, and course-specific variables such as attendance and participation. Time spent on the online learning platform and the number of posts and responses made in discussion forums.	Data of 11000 students from the online platform between 2012 to 2016.	Sensitivity, specificity, precision, recall, F1 score, and area under the receiver operating characteristic curve.	Achieving an accuracy of 86.3%.for predicting student performance. Demographic variables such as age, gender, and ethnicity were less predictive of at-risk status than academic variables such as GPA, course grades, and attendance.
[Adnan et al., 2021] [20]	"Predicting at-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models"	DT, RF, SVM, KNN, LR, and Multi-Layer Perceptron (MLP)	Demographic variables such as age, gender, and program of study, Historical academic performance variables such as cumulative grade point average (CGPA) and the number	Twenty-four online courses covering 4 disciplines and a total of 15,200 students.	Accuracy Precision Recall F1-score Area Under the Receiver Operating Characteristic Curve	The study also emphasized the importance of selecting appropriate evaluation metrics and setting appropriate intervention thresholds to improve the effectiveness of early interventions for at-risk students.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

			of credits earned Course-specific variables such as course grade, number of attempts at a course, and the number of forum posts			
[Ahajjam et al., 2022] [21]	"Predicting Students' Final Performance Using Artificial Neural Networks"	Artificial Neural Networks (ANNs)	Academic performance variables such as mid-term exam scores, final exam scores, attendance rate, and participation rate as the predicting variables.	High school cycle common core, first-year baccalaureate, and second-year baccalaureate.	Mean squared error (MSE) and the coefficient of determination (R-squared)	Previous academic performance, including their cumulative grade point average (CGPA), the number of credits earned, and their attendance rate.
[Pek et al., 2022] [22]	"The Role of Machine Learning in Identifying Students At-Risk and Minimizing Failure"	Decision tree, LR,RF, KNN, Naïve Bayes and SVM	Academic features include the school's name, kindergarten attendance status, Travel time to school, University plan, and GPA of semesters. Demographic features such as Gender, Age, City, District, No of siblings, and free time.	The dataset contains 555 student information with 38 features.	Accuracy, Precision, and Recall.	The finding shows that academic data is useful for hybrid model training. The accuracy was 98.8%.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

[Bujang et al., 2021] [23]	"Multiclass Prediction Model for Student Grade Prediction Using Machine Learning"	Decision tree, LR, RF, KNN, NB, ANN, and SVM	"Demographic data (age and gender), socioeconomic data (family income), and academic data (previous semester grade, attendance rate, and study hours)." The target variable was the final grade, which was classified into five categories: A, B, C, D, and E.	The dataset includes 6,832 records, with 2,277 students and 43 features—six courses at a public university in Malaysia.	Confusion Matrics, Accuracy, Precision, Recall, F1-score, and ROC	The proposed multiclass prediction model using ML models, including KNN, SVM, Naïve Bayes, Decision Tree, and Random Forest, can accurately predict student grades. Academic performance features such as previous grades, attendance, and participation in activities predict student grades, followed by demographic features such as gender, race, and parent's education level.
[Liu et al., 2020] [24]	"Multiple Features Fusion Attention Mechanism Enhanced Deep Knowledge Tracing for Student Performance Prediction"	Deep knowledge tracing (DKT) model.	Log data, question data, The log data included variables such as time spent on a question, time spent on a bundle, question type, difficulty level, and question tag. And the number of attempts.	The student interaction dataset logs 10,674 students and 19,593 unique interactions collected from an online educational platform.	Accuracy, Precision, Recall, F1-score, and ROC.	The proposed model, MFAME-DKT, outperformed other state-of-the-art models regarding the accuracy, AUC, and F1 score for predicting student performance.
[Moreno-Marcos et al., 2020] [25]	"Analysis of the Factors Influencing Learners' Performance Prediction With	RG,RF, KNN, SVM and MLP	Gender, age, nationality, and educational factors such as previous academic performance, type of high school	Four online courses were offered on the edX platform in 2017,	MAE, RMSE, coefficient of determination (R^2), precision, recall, and F1-score	Performance prediction of learners using learning analytics is highly influenced by various factors such as the learning

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

	Learning Analytics”		attended, the degree program, the number of courses taken and time spent on the course, the number of discussion forum messages posted, and the number of clicks on the platform	including data from 22,564 learners.		management system used, course metadata, and the characteristics of learners. A combination of various factors and ML models could improve the accuracy of performance prediction models.
[Zhang et al., 2021] [3]	“Predicting and Understanding Student Learning Performance Using Multisource Sparse Attention Convolutional Neural Networks”	Multisource Sparse Attention Convolutional Neural Network (MSCNN	Gender and age, academic backgrounds such as high school scores, university entrance scores, cumulative GPA, and learning behavior features such as course attendance, quiz scores, and assignment scores. Student information system, learning management system, and academic records	Dataset from a massive open online course (MOOC) platform containing the log data of 16,417 learners from 12 different courses, totaling 1,512,556 log entries.	Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Pearson Correlation Coefficient (PCC).	The attention mechanism used in the MS-SACNN model helped identify the most important features from multiple sources. The authors also revealed the MS-SACNN model improved when the dataset included more student data.
[Zhao et al., 2021] [27]	“Academic Performance Prediction Based on Multisource, Multifeature Behavioral Data”	Decision tree, LSTM, and Gradient boosting decision tree (GBDT).	Demographic information, academic history, course information, assignment submission, online behavior,	Data from Freshman Seminar from 2018-2019 based on Multisource,	Accuracy	The results suggest that a combination of behavioral and demographic features can improve the accuracy of theoretical performance prediction compared to

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

			and library access.	Multifeature behavioral dataset.		using only behavioral or demographic features alone.
[Wang et al., 2023] [28]	"ProbSAP: A comprehensive and high-performance system for student academic performance prediction"	XGBoost, CNN, RFR and SVR	Demographic information (e.g., gender, age), academic history (e.g., previous GPA), course-related information (e.g., course load, difficulty level), and behavioral data (e.g., online learning activities, library usage).	The dataset contains information on over 10,000 students and 40,000 courses from universities in China.	Precision, Accuracy, Recall, F1-Score, and AUC-ROC	The collaborative component can provide a new information material to address educational data discrepancies.
[Zeineddine et al., 2021] [29]	"Enhancing prediction of student success: Automated machine learning approach."	AutoMLs such as Generalized linear models, Gradient boosting machines, and NN	Age, gender, nationality, high school grades, standardized test scores, and University-related factors such as major and previous academic performance. Student engagement, such as class attendance and participation, Student behavior, such as study habits and time management skills; and Student mental	The dataset contains 10,235 records with 17 input features and one target variable from a UAE university.	The evaluation metrics used in this study include 10-fold cross-validation and accuracy, precision, recall , and F1-score.	The proposed model outperformed traditional machine learning models and touched an accuracy of 94.7% in predicting student success. They revealed using automated ML models for educational data analysis and provides insights into the factors most strongly associated with student success.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

			health, such as stress levels and emotional well-being.			
[Bilal et al., 2022] [30]	"The Role of Demographic and academic features in a student performance prediction"	SuprvisedRG, RF, DT, KNN, SVM, and ANN	Age, gender, country, and high school type are all demographic characteristics, as are CGPA, the field of study, semester pressure, and the number of classes taken.	One thousand seventy-five undergraduate students from a Pakistani institution provided data.	15-Fold cross-validation, Accuracy, Precision, and Recall.	The results showed that SVM performed the best, with an accuracy of 92%.
[Wang et al., 2022] [31]	"Complexity-based attentive, interactive student performance prediction for personalized course study planning"	Complexity-based attentive, interactive neural network (CAINN)	Student academic records, course records, course characteristics, and student characteristics. Previous academic performance, attendance records, learning behavior records, course workload, course difficulty	The paper included academic and demographic information from 580 undergraduate students at a Chinese institution.	MSE and MAE	The suggested approach outperformed typical machine learning methods, obtaining a precision of 91.8%, according to the research. In addition, the authors identified the most important factors of student achievement, such as past academic performance, course grades, and course difficulty.
[Arashpour et al., 2023] [32]	"Predicting individual learning performance using machine learning hybridized with the teaching-learning-	SVM and ANN using TLBO	Gender, Age, Disability, Prior qualification, and Idex of multiple deprivations.	The open university analytics dataset has 15,931 records.	Accuracy, Precision, Recall, F1-Score, MCC, FM, Confusion matrix, and ROC-AUC.	This research delivers scientific effectiveness by improving anticipates of student test performance by constructing hybridized ML

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

	based optimizatio n”					models and TLBO. The accuracy of model SVM is better than ANN, almost 92.27% on All data.

This highlights the importance of considering various aspects when predicting student performance. The studies are conducted at different educational levels, including undergraduate, high school, and preparatory year. This diversity in educational levels allows for a broader understanding of academic performance prediction across different contexts. Furthermore, the size of the datasets varies significantly, ranging from as small. Larger datasets can potentially yield more accurate models, but smaller datasets may have limitations in terms of generalizability.

Regarding employed techniques, various machine learning techniques and algorithms are employed in these studies, including Support Vector Machines (SVM), Random Forest (RF), k-Nearest Neighbours (KNN), Neural Networks (NN), Decision Trees (DT), and more. The choice of technique often depends on the specific research objectives and dataset characteristics. These have wider implication on the performance of the employed model. Each study reports the best-performing model or technique providing ground for understanding which methods tend to be more effective in predicting academic performance.

In general, this table shows the range and intricate nature of research, on predicting performance. It highlights the importance of considering factors and using machine learning methods to create reliable models. Researchers in this field face challenges such, as obtaining more inclusive datasets addressing biases and improving how well predictive models can be understood and applied in contexts.

ML possesses significant potential to accelerate progress in the field of education, leading to noticeable improvements in its efficiency. When ML techniques are appropriately and efficiently applied in the educational domain, they have the power to revolutionize teaching, learning, and research. Educators who utilize ML will gain greater insight into their students' learning progress, enabling them to intervene early and take measures to enhance success and retention rates for struggling students. Our literature survey also revealed some critical aspects that are rquired to address. We can categorize those aspects as follows:

- Dataset Requirement
- Static and Dynamic Feature Requirement

4.1 Dataset Requirement

Developing ML-based student prediction models/techniques requires careful consideration of various dataset requirement issues that may arise during the development process. One of the most critical factors is the quality of the data used to train the models, as it can significantly impact the accuracy and reliability of the predictions. Insufficient data quantity can result in underfit models that perform poorly on new data, while irrelevant data may not contribute to solving the problem effectively. Additionally, data imbalance, where the number of samples in each class is unequal, can introduce bias towards the majority class, leading to poor performance on the minority class. Data privacy and security regulations must also be considered, as well as potential biases that may be introduced during data collection from specific populations or groups. By addressing these issues, ML-based student prediction models can be developed that are accurate, reliable, and fair, and can effectively predict student performance and progress.

4.2 Static and Dynamic Feature Requirement

There are several challenges related to the requirements of static and dynamic features in predicting student performance using ML. While static features like demographic data, socioeconomic status, and previous academic achievements are critical for the accuracy and relevance of the model, relying solely on them may not provide sufficient information to make accurate predictions as they do not account for changes that occur over time. On the other hand, dynamic features such as attendance, test scores, and study habits are essential for predicting student performance, but their collection and processing can be challenging and time-consuming, leading to overfitting or underfitting of the model. The selection of features is also crucial, as including irrelevant or redundant features can lead to overfitting, while omitting important features may result in underfitting. Additionally, feature extraction is a critical process that can affect the accuracy and reliability of the model and

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

may lead to biased predictions if not correctly addressed. Finally, the availability of data is a significant challenge, as some features may not be available or may be difficult to collect, limiting the scope of the model and leading to biased predictions. To address these challenges, it is crucial to carefully select appropriate features, use reliable feature extraction techniques, and consider the limitations of available data to develop accurate and reliable ML models for predicting student performance.

5. Conclusion

In conclusion, the systematic literature review on ML techniques to predict student performance has highlighted several issues that must be addressed for accurate and reliable predictions. The use of appropriate datasets is crucial and combining multiple datasets can improve the accuracy and comprehensiveness of the model. The comprehensive literature analysis shows that ML approaches may predict student performance as well as underlines the important issues that must be addressed to construct accurate and reliable models. UCI, Kaggle, and MOOCs can give varying and massive data sets to construct more robust and trustworthy models but still massive and comprehensive domain specific datasets are lacking. Overfitting, bias, scope, interpretability, and ethics are all areas that need to be addressed in order to ensure the reliability and accuracy of models. In conclusion, we highlight the possible avenues for future research that will use machine learning approaches to forecast the performance of students. One of our future goals is to put some of the outstanding works into practise to develop new state of the art Student performance prediction techniques. As a consequence of this, instructors could pick up additional indications to assist them in developing appropriate solutions for their students to accomplish precise learning goals.

6. Acknowledgment

Appreciation is also conveyed to University Malaysia Pahang Al-Sultan Abdullah for project financing under UMP Short Term Grant RDU1903122.

References

[1] F. Han and R. A. Ellis, "Predicting Students' Academic Performance by Their Online Learning Patterns in a Blended Course To What Extent Is a Theory driven Approach and a Data driven Approach Consistent?," Educational Technology & Society, vol. 24, no. 1, pp. 191-204, 2021.

[2] R. K. Kavitha, W. Jaisingh, and S. K. K. Devi, "Applying Learning Analytics to Study the Influence of Fundamental Computer Courses on Project Work and Student Performance Prediction using Machine Learning Techniques," in 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), 2021, pp. 1-5.

[3] Y. Zhang, Y. Yun, R. An, J. Cui, H. Dai, and X. J. F. i. p. Shang, "Educational data mining techniques for student performance prediction: method review and comparison analysis," vol. 12, p. 698490, 2021.

[4] P. Dabhade, R. Agarwal, K. P. Alameen, A. T. Fathima, R. Sridharan, and G. Gopakumar, "Educational data mining for predicting students' academic performance using machine learning algorithms," Materials Today: Proceedings, vol. 47, pp. 5260-5267, 2021/01/01/ 2021.

[5] X. Du, J. Yang, and J. L. Hung, "An Integrated Framework Based on Latent Variational Autoencoder for Providing Early Warning of At-Risk Students," IEEE Access, vol. 8, pp. 10110-10122, 2020.

[6] B. Albreiki, N. Zaki, and H. Alashwal, "A Systematic Literature Review of Student' Performance Prediction Using Machine Learning Techniques," Education Sciences, vol. 11, p. 552, 09/16 2021.

[7] T.-N. Huynh-Ly, H.-T. Le, and N. Thai-Nghe, "Integrating Deep Learning Architecture into Matrix Factorization for Student Performance Prediction," in Future Data and Security Engineering, Cham, 2021, pp. 408-423: Springer International Publishing.

[8] Y. Chen and L. Zhai, "A comparative study on student performance prediction using machine learning," Education and Information Technologies, pp. 1-19, 03/04 2023.

[9] T.-C. Yang, Y.-L. Liu, and L.-C. Wang, "Using an Institutional Research Perspective to Predict Undergraduate Students Career Decisions in the Practice of Precision Education," Educational Technology & Society, vol. 24, no. 1, pp. 280-296, 2021.

[10] Y. Bao, F. Lu, Y. Wang, Q. Zeng, and C. Liu, "Student Performance Prediction Based on Behavior Process Similarity," Chinese Journal of Electronics, https://doi.org/10.1049/cje.2020.02.012 vol. 29, no. 6, pp. 1110-1118, 2020/11/01 2020.

[11] E. Alyahyan and D. J. I. J. o. E. T. i. H. E. Düştegör, "Predicting academic success in higher education: literature review and best practices," vol. 17, pp. 1-21, 2020.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023

[12] M. J. Page et al., "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," vol. 88, p. 105906, 2021.
[13] E. Alhazmi and A. Sheneamer, "Early Predicting of Students Performance in Higher Education," IEEE Access, vol. 11, pp. 27579-27589, 2023.
[14] A. Nabil, M. Seyam, and A. Abou-Elfetouh, "Prediction of Students' Academic Performance Based on Courses' Grades Using Deep Neural Networks," IEEE Access, vol. 9, pp. 140731-140746, 2021.
[15] G. Feng, M. Fan, and Y. Chen, "Analysis and Prediction of Students' Academic Performance Based on Educational Data Mining," IEEE Access, vol. 10, pp. 19558-19571, 2022.
[16] A. Rafique et al., "Integrating Learning Analytics and Collaborative Learning for Improving Student's Academic Performance," IEEE Access, vol. 9, pp. 167812-167826, 2021.
[17] S. Alwarthan, N. Aslam, and I. U. Khan, "An Explainable Model for Identifying At-Risk Student at Higher Education," IEEE Access, vol. 10, pp. 107649-107668, 2022.
[18] A. Alshanqiti and A. Namoun, "Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification," IEEE Access, vol. 8, pp. 203827-203844, 2020.
[19] J. L. Hung, B. E. Shelton, J. Yang, and X. Du, "Improving Predictive Modeling for At-Risk Student Identification: A Multistage Approach," IEEE Transactions on Learning Technologies, vol. 12, no. 2, pp. 148-157, 2019.
[20] M. Adnan et al., "Predicting at-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models," IEEE Access, vol. 9, pp. 7519-7539, 2021.
[21] T. Ahajjam, M. Moutaib, H. Aissa, M. Azrour, Y. Farhaoui, and M. Fattah, "Predicting Students' Final Performance Using Artificial Neural Networks," Big Data Mining and Analytics, vol. 5, no. 4, pp. 294-301, 2022.
[22] R. Z. Pek, S. T. Özyer, T. Elhage, Ö. T, and R. Alhajj, "The Role of Machine Learning in Identifying Students At-Risk and Minimizing Failure," IEEE Access, vol. 11, pp. 1224-1243, 2023.
[23] S. D. A. Bujang et al., "Multiclass Prediction Model for Student Grade Prediction Using Machine Learning," IEEE Access, vol. 9, pp. 95608-95621, 2021.
[24] D. Liu, Y. Zhang, J. Zhang, Q. Li, C. Zhang, and Y. Yin, "Multiple Features Fusion Attention Mechanism Enhanced Deep Knowledge Tracing for Student Performance Prediction," IEEE Access, vol. 8, pp. 194894-194903, 2020.
[25] P. M. Moreno-Marcos, T. C. Pong, P. J. Muñoz-Merino, and C. D. Kloos, "Analysis of the Factors Influencing Learners' Performance Prediction With Learning Analytics," IEEE Access, vol. 8, pp. 5264-5282, 2020.
[26] Y. Zhang, R. An, S. Liu, J. Cui, and X. Shang, "Predicting and Understanding Student Learning Performance Using Multi-Source Sparse Attention Convolutional Neural Networks," IEEE Transactions on Big Data, vol. 9, no. 1, pp. 118-132, 2023.
[27] L. Zhao et al., "Academic Performance Prediction Based on Multisource, Multifeature Behavioral Data," IEEE Access, vol. 9, pp. 5453-5465, 2021.
[28] X. Wang, Y. Zhao, C. Li, and P. Ren, "ProbSAP: A comprehensive and high-performance system for student academic performance prediction," Pattern Recognition, vol. 137, p. 109309, 2023/05/01/ 2023.
[29] H. Zeineddine, U. Braendle, and A. Farah, "Enhancing prediction of student success: Automated machine learning approach," Computers & Electrical Engineering, vol. 89, p. 106903, 2021/01/01/ 2021.
[30] M. Bilal, M. Omar, W. Anwar, R. H. Bokhari, and G. S. Choi, "The role of demographic and academic features in a student performance prediction," Scientific Reports, vol. 12, no. 1, p. 12508, 2022/07/22 2022.
[31] P. Wang, L. Li, R. Wang, Y. Xie, and J. Zhang, "Complexity-based attentive, interactive student performance prediction for personalized course study planning," Education and Information Technologies, vol. 27, no. 4, pp. 5471-5493, 2022/05/01 2022.
[32] M. Arashpour et al., "Predicting individual learning performance using machine-learning hybridized with the teaching-learning-based optimization," Computer Applications in Engineering Education, https://doi.org/10.1002/cae.22572 vol. 31, no. 1, pp. 83-99, 2023/01/01 2023.
[33] A. Asselman, M. Khaldi, and S. Aammou, "Evaluating the impact of prior required scaffolding items on improving student performance prediction," Education and Information Technologies, vol. 25, no. 4, pp. 3227-3249, 2020/07/01 2020.

Doi: https://doi.org/10.54216/FPA.130207
Received: April 23, 2023 Revised: July 19, 2023 Accepted: September 19, 2023