AI-BASED EARLY PREDICTION AND INTERVENTION FOR STUDENT ACADEMIC PERFORMANCE IN HIGHER EDUCATION

Maram Khamis Al-Muharrami, Fatema Said Al-Sharqi, Al-Zahraa Khalid Al-Rumhi and Shamsa Said Al-Mamari, Ijaz Khan

Department of Information Technology, Buraimi University College, Buraimi City, Oman

ABSTRACT

Accurately identifying at-risk students in higher education is crucial for timely interventions. This study presents an AI-based solution for predicting student performance using machine learning classifiers. A dataset of 208 student records from the past two years was preprocessed, and key predictors such as midterm grades, previous semester GPA, and cumulative GPA were selected using information gain evaluation. Multiple classifiers, including Support Vector Machine (SVM), Decision Tree, Naive Bayes, Artificial Neural Networks (ANN), and k-Nearest Neighbors (k-NN), were evaluated through 10-fold cross-validation. SVM demonstrated the highest performance with an accuracy of 85.1% and an F2 score of 94.0%, effectively identifying students scoring below 65% (GPA < 2.0). The model was implemented in a desktop application for educators, providing both class-level and individual-level predictions. This user-friendly tool enables instructors to monitor performance, predict outcomes, and implement timely interventions to support struggling students. The study highlights the effectiveness of machine learning in enhancing academic performance monitoring and offers a scalable approach for AI-driven educational tools.

KEYWORDS

Artificial Intelligence, Machine Learning, Student Performance Prediction, Higher Education, AI-based Application

1. INTRODUCTION

The rapid advancement of Information and Communication Technology (ICT) has significantly impacted various sectors, including education, by reshaping educational systems, prompting the adoption of digital strategies, and highlighting critical gaps and inequalities in digital capacity [1]. In higher educational institutions (HEIs), maintaining high educational standards and ensuring student success have become critical priorities. Governmental and accreditation agencies, such as the Oman Academic Accreditation Authority and Quality Assurance (OAAAQA) is involved in maintaining and ensuring quality in higher education institutions (HEIs) in Oman [2]. Consequently, monitoring student performance has emerged as an essential factor in meeting these standards and providing accountability [3].

Instructors often face an overwhelming number of responsibilities, making it challenging to continuously monitor each student's academic progress and implement timely interventions [4]. Traditional methods of monitoring, which rely on periodic assessments, may not provide the early insights needed to support students at risk of underperforming [5]. The increased workload on instructors underscores the need for technological solutions that integrate psychological theory, research, and statistical methods to assist in tracking and improving student performance within the dynamic processes of classroom environments [6].

Modern educational institutions are increasingly implementing sophisticated systems that continuously collect and analyze data on students' academic activities to improve learning, develop self-regulated learning skills, and support student success [7]. However, this data is often underutilized. Leveraging this data through Machine Learning (ML), a branch of Artificial Intelligence (AI), offers innovative tools for monitoring and predicting student performance [8]. Machine learning algorithms can analyze students' behavioral and academic data to predict their future performance, allowing for early intervention and identifying those who may require additional support [9].

Previous studies have demonstrated the potential of machine learning in predicting student performance. For example, Hashim et al. [10] employed various supervised machine learning algorithms and found that logistic regression was the most accurate in predicting student outcomes. Similarly, Lau et al. [11] utilized an artificial neural network to model and predict student academic performance, achieving effective results. Mondal et al. [12] applied a Recurrent Neural Network to predict student performance, demonstrating higher accuracy compared to traditional neural networks. Pallathadka et al. [13] explored various classifiers, identifying Support Vector Machine (SVM) as the most accurate for classifying and predicting student performance. Sukhbaatar et al. [14] proposed an early prediction scheme using a neural network to identify at-risk students in a blended learning course, which successfully identified failing students early in the semester. Additionally, Alcaraz et al. [15] designed a tailored early warning system for a course, finding that an ensemble classifier with a novel weighted voting strategy was the most effective. These studies highlight the effectiveness of machine learning classifiers in academic contexts, establishing a basis for future exploration.

Despite these progressions, a notable gap persists in creating intuitive and comprehensive applications that incorporate diverse machine learning models for real-time monitoring and intervention within educational environments. Existing studies have demonstrated the efficacy of specific algorithms in predicting academic outcomes, yet practical tools that educators can readily use to continuously track student progress and implement timely interventions are lacking.

This research addresses a significant gap by creating an AI-based application that predicts student performance using diverse machine learning classifiers and provides actionable insights for educators. To develop this AI-based application, we collect and preprocess academic records from host institute over the past two years, creating a training dataset suitable for machine learning. The dataset includes various features such as gender, major, grades in continuous assessments (assignment, midterm exam), and CGPA. Through feature engineering, we identify the most significant predictors of student performance. Our approach applies multiple machine learning classifiers to identify the model with the highest predictive accuracy. The selected model is transformed into a desktop application using JavaScript. This application will analyze students' academic progress after the midterm exam and provide predictions of whether students will achieve satisfactory or unsatisfactory outcomes. The midterm accounts for 30% of the total grade, leaving 70% of the grade for students to improve upon. The evaluation in this study focuses on a single course, allowing for precise and tailored predictions of student performance within this specific context. However, the same methodology can be applied to additional courses by incorporating their respective datasets. The developed software is designed to be flexible and scalable, enabling educators to add and analyze multiple courses seamlessly, extending its applicability across various academic contexts.

The proposed solution will empower instructors to proactively support students, particularly those identified as likely to end the course with unsatisfactory grades. By providing timely consultations and additional resources, instructors can help guide these students towards academic success, thereby enhancing the overall educational standards and accountability of the institution. This research addresses the critical need for effective student performance monitoring in HEIs. By utilizing machine learning to predict academic outcomes, we can provide instructors with valuable tools to improve student support and intervention strategies, ultimately contributing to better educational outcomes and institutional accountability.

The paper is structured as follows: Section 2 reviews relevant literature, Section 3 summarizes related work, Section 4 describes the dataset and tools, Section 5 details the methodology, Section 6 concludes with future directions.

2. LITERATURE REVIEW

Predicting student academic performance has garnered significant interest in the field of education [16]. Traditional approaches relied heavily on periodic assessments such as exams, quizzes, and assignments to gauge student understanding and progress. These methods provided limited insights, often failing to identify at-risk students early enough for timely interventions. With the advent of Information and Communication Technology (ICT), educational institutions have adopted more sophisticated methods, including the integration of digital learning platforms that enhance continuous assessment and data collection [17]. This progression has achieved a new level with the incorporation of machine learning (ML) techniques. Machine learning models can process large volumes of academic, demographic, and behavioral data, enabling more precise and timely predictions of student performance [18]. This shift from static, periodic assessments to dynamic, data-driven approaches represents a significant advancement in educational monitoring, allowing for more personalized and proactive student support.

In recent years, educational technology has rapidly integrated machine learning applications to enhance learning outcomes [19]. Modern trends encompass the use of predictive analytics to anticipate student performance, adaptive learning systems that tailor educational content to individual requirements, and intelligent tutoring systems providing real-time feedback and assistance [20]. Learning management systems (LMS) are increasingly leveraging machine learning algorithms to evaluate student interactions and engagement, offering insights into learning behaviors and highlighting areas for improvement [21]. These technologies enhance personalized education and enable educators to identify and support at-risk students more efficiently. AI-powered chatbots for administrative tasks and the adoption of virtual reality (VR) and augmented reality (AR) for immersive learning are gaining popularity, highlighting the transformative role of machine learning in education [22].

Despite the promising potential of machine learning in educational contexts, several challenges and limitations must be addressed to ensure its effective implementation. A key challenge lies in maintaining data quality and completeness, as machine learning models require reliable datasets to produce accurate predictions [23]. In many educational institutions, data is often fragmented or inconsistent, which can hamper the performance of these models. Additionally, data privacy concerns are paramount; educational institutions must ensure that the use of student data adheres to strict privacy laws and regulations [24]. The complexity of machine learning models presents a challenge, as educators may find it difficult to interpret and utilize the insights they produce [25]. Additionally, there is a risk of algorithmic bias, where models could unintentionally reinforce existing inequalities if not properly managed [26]. Overcoming these challenges necessitates strong data governance policies, continuous educator training, and the creation of transparent, interpretable machine learning models.

Machine learning algorithms are designed to build models from training data to categorize or predict outcomes without being specifically programmed [27]. Machine learning algorithms are utilized across diverse domains, including pattern recognition, object detection, and text analysis, and are developed using the latest data trends [28]. Each classifier operates based on distinct principles and techniques. Logistic regression, for instance, uses an explicit model with a well-understood statistical foundation, making it simple and interpretable for modeling probabilities, even in complex real-world problems [29]. A decision tree divides data into branches according to feature values, enabling clear visualization and simplifying the interpretation of the classification process [30]. Support Vector Machines (SVM) are supervised learning models that use training data to create margins that separate classes, making them effective for classification and regression analysis. It classifies new data points by measuring their distance from the class margins, making it especially effective for complex, small-to-medium-sized datasets [31]. Naive Bayes utilizes Bayes' theorem while assuming conditional independence among features, making it both computationally efficient and highly effective for text classification and high-dimensional data domains [32]. Artificial Neural Networks (ANN) are composed of layers of interconnected neurons that process input data through weighted links, enabling them to model complex and non-linear relationships effectively [33]. Ensemble methods, such as Random Forests, enhance prediction accuracy by combining multiple decision trees and aggregating their outputs through averaging or majority voting [34]. Each type of classifier has distinct strengths and weaknesses, with the selection depending on the characteristics of the data and the specific demands of the prediction task.

The significance and suitability of various machine learning classifiers for predicting student performance lie in their ability to handle diverse data types and uncover intricate patterns that traditional methods might miss. Logistic regression is well-suited for binary classification tasks, such as predicting pass or fail outcomes, because of its simplicity and interpretability.[29]. Decision Trees provide clear visual representations of decision-making processes, making them useful for educators to understand the factors influencing student performance. Support Vector Machines (SVM) are highly effective for handling high-dimensional data and are particularly suitable for datasets with complex but linearly separable patterns, making them ideal for analysing nuanced academic performance data. Naive Bayes classifiers, assuming feature independence, perform effectively in high-dimensional spaces and are especially efficient for real-time prediction tasks, such as tracking on-going student performance. Artificial Neural Networks (ANN) and Recurrent Neural Networks (RNN) are powerful in capturing non-linear relationships and temporal dependencies in student data, respectively, making them ideal for modelling complex student behaviours over time. Ensemble methods, such as Random Forests, improve predictive accuracy and robustness by combining multiple decision trees, minimizing overfitting, and enhancing generalization. The varied strengths of these classifiers enable adaptable and customized approaches to predicting student performance, helping educators deliver timely and effective support to at-risk students and enhance overall educational outcomes.

3. RELATED WORK

This section examines related work on the application of machine learning algorithms for predicting academic performance. Past research has showcased a range of approaches and methodologies, emphasizing the potential of machine learning to deliver actionable insights in educational contexts. The goal of this review is to offer a comprehensive summary of existing literature, highlighting major trends, methods, and gaps that this study aims to address.

Musso et al. [35] design a machine learning model to forecast academic success and dropout rates by analyzing learning strategies, social support, motivation, socio-demographics, health conditions, and academic performance. The study discovered that learning strategies were the strongest predictor of GPA, whereas background information was the key factor in identifying potential dropouts. The study [18] explores educational data mining to predict undergraduate students' final exam grades using their midterm grades. The study compares machine learning algorithms, including Random Forests, SVM, Logistic Regression, Naïve Bayes, and k-NN, using course datasets. Achieving 70-75% accuracy, it highlights the potential of data-driven methods for early identification of at-risk students and informed decision-making in higher education.

Hashim et al. [10] investigates the application of supervised machine learning algorithms to predict student performance in higher education. By analyzing demographic, academic, and behavioral data, they compared multiple algorithms. Logistic Regression proved most effective, achieving approximately 69% accuracy for predicting passes and 89% for failures. This study demonstrates the potential of machine learning in educational data mining to improve institutional decision-making and foster student success. Hussain et al. [36] applies machine learning techniques to predict academic performance at secondary and intermediate levels. Using regression models and decision tree classifiers optimized with genetic algorithms, they forecast grades based on historical data. The study achieved high accuracy and low error rates, validating the potential of these methods for enhancing educational planning and development.

Alhazmi et al. [37] aim to predict students' academic performance in higher education by analyzing various factors such as admission scores, first-level course scores, academic achievement tests, and general aptitude tests. They use both clustering and classification techniques, employing t-SNE for dimensionality reduction and various machine learning algorithms. Their findings suggest that incorporating comprehensive features improves prediction accuracy, helping educational institutions to identify and support at-risk students early on, and thereby enhancing overall educational outcomes. The study [38] examines the relationship between college students' internet usage and academic performance by analysing features such as online duration, traffic volume, and connection frequency. Supervised machine learning algorithms were employed for prediction, revealing that frequent connections positively correlate with success, while high traffic volume negatively impacts performance. Expanding the feature set enhanced prediction accuracy, showcasing the effectiveness of internet usage data in predicting academic outcomes.

Ojajuni et al. [39] aim to predict student academic performance using machine learning by analyzing historical data to identify key factors affecting academic success. The study applies a set of supervised machine learning classifiers to classify student performance.. The study concludes that applying machine learning in education can help educators identify at-risk students early and improve educational outcomes through informed decision-making. Adnan et al. [40] aim to predict at-risk students in online learning environments at various stages of course completion to facilitate timely interventions by instructors. The study utilizes machine learning models to analyze student engagement, demographics, and assessment data. Their findings indicate that early prediction and intervention can significantly improve student retention and performance. Khan et al. [41] aim to monitor student performance and devise preventive measures using artificial intelligence in educational institutions. The study employs machine learning algorithms to predict student performance. Their research finds that decision tree models are the most effective, and highlights CGPA, midterm exam marks, and attendance as key predictors of student outcomes. The model allows instructors to identify struggling students early and implement tailored interventions, leading to improved academic results.

Pallathadka et al. [13] aim to classify and predict student performance using various machine learning algorithms to enhance academic outcomes. The study employs several algorithms on the student performance dataset, focusing on accuracy and error rate. Their findings indicate that SVM outperforms other algorithms in accuracy, demonstrating its effectiveness for educational data mining. This analysis helps institutions identify and support students needing additional focus, ultimately aiming to lower failure rates and improve educational quality. Similarly, several authors [42] [43] [44] [45] [46] [47] have developed student performance prediction models using machine learning algorithms. Table 1 summarizes research studies that highlight the role of machine learning algorithms in predicting students' academic performance.

Table 1 A summary of various related work

Ref.	Algorithm(s) used	Performance Metrics	Dataset Size	Place of Study	Key Findings
[10]	Decision Tree, Naive Bayes, Logistic Regression, Support Vector Machine, K-Nearest Neighbour, Sequential Minimal Optimisation, Neural Network	Accuracy (Logistic Regression: 68.7% for exact final grades, 88.8% for pass/fail prediction)	499 records	Iraq	Logistic Regression most accurate for predicting exact grades and pass/fail status.
[18]	Random Forests, Nearest Neighbour, Support Vector Machines, Logistic Regression, Naïve Bayes, K-Nearest Neighbour	Classification accuracy: 70-75%	1854 records	Turkey	Proposed model achieved 70-75% classification accuracy for predicting final exam grades.
[36]	Decision Tree, K-Nearest Neighbour, Genetic Algorithm	DT accuracy: 94.39%, K-NN accuracy: 85.74%, GA-DT accuracy: 96.64%, GA-KNN accuracy: 89.92%	90,000 records	Pakistan	GA-DT classifier achieved highest accuracy (96.64%) for grade prediction.
[37]	XGBoost, Logistic Regression, Support Vector Machine, K-Nearest Neighbour, Random Forest	Highest accuracy: XGBoost 85%	275,000 records	Saudi Arabia	XGBoost classifier achieved highest accuracy (85%) for early student performance prediction.
[39]	Random Forest, Support Vector Machines, Gradient Boosting, Decision Tree, Logistic Regression, Extreme Gradient Boosting (XGBoost), Deep Learning	Highest accuracy: XGBoost 97.12%	1044 records	USA	XGBoost achieved highest accuracy (97.12%) for predicting student academic performance.
[40]	Random Forest, Deep Feed Forward Neural Network (DFFNN)	Accuracy: Random Forest 91%, DFFNN 89%	32,593 records	Kohat University of Science and Technology, Pakistan	Random Forest outperformed other models with an accuracy of 91% at 100% course length. Earliest identification of at-risk students at 20% course length.
[41]	Decision Tree, k-NN, Naive Bayes, Artificial Neural Network	Decision Tree accuracy: 86%, F-Measure: 0.91, MCC: 0.63	151 records	Buraimi University College, Oman	Decision Tree model achieved highest performance metrics and was transformed into an easily interpretable format for instructors to take preventive measures.
[13]	Naive Bayes, ID3, C4.5, SVM	SVM accuracy: 93%, Naive Bayes accuracy: 89%, ID3 accuracy: 85%, C4.5 accuracy: 87%	649 records	University of Minho, Portugal	SVM achieved highest accuracy for classifying student performance data.

4. MATERIALS USED

4.1. Data Description

The dataset used in this study comprises 208 instances of student performance in a course at host institute. Each instance is characterized by features relevant to predicting final outcomes, with 160 labelled as "High" (indicating likely good performance) and 48 as "Low" (indicating likely poor performance), resulting in an imbalance ratio of approximately 3.33:1. The goal is to accurately identify "Low" performing students after the midterm exam to enable timely interventions. In this study, a "struggling student" is represented by the "Low" label in the dataset. This classification refers to students who score below 65% in their overall marks, which corresponds to a cumulative GPA of less than 2.0 based on the grading policy of the host institution. These students are identified as being at risk of achieving unsatisfactory academic outcomes and in need of timely intervention. All student data is anonymized, and the system complies with data protection regulations, ensuring confidentiality and secure access.

4.2. WEKA

WEKA (Waikato Environment for Knowledge Analysis) is a popular open-source software suite that provides a comprehensive collection of machine learning algorithms and tools for data mining tasks [48]. WEKA, developed by the University of Waikato, is a powerful tool widely adopted in academia and industry due to its intuitive interface and comprehensive features. It offers robust support for data preprocessing, classification, regression, clustering, and visualization, making it highly versatile for analytical tasks. In this study, WEKA was chosen for its reliability and ease of use, enabling efficient experimentation. Classifiers were implemented using WEKA's built-in algorithms with default settings, ensuring a consistent evaluation framework for the imbalanced dataset. Furthermore, WEKA’s advanced functionalities, such as cross-validation and feature selection, were utilized to enhance the accuracy and credibility of the results.

4.3. Confusion Matrix

A confusion matrix is an essential tool for assessing the performance of a classification algorithm. It provides a detailed summary of prediction outcomes, offering valuable insights into the classifier's effectiveness in solving a given classification problem. The matrix is structured in a table format, as shown in table 2, which contrasts the actual class labels against the predicted class labels, typically organizing the data into four categories: True Positives (TP), False Negatives (FN), False Positives (FP), and True Negatives (TN). True positives represent the instances where the model correctly predicted the positive class, whereas false negatives are the positive instances that were incorrectly classified as negative. Similarly, false positives are the negative instances that were incorrectly classified as positive and true negatives are the instances where the model correctly predicted the negative class. This matrix is crucial as it allows for the calculation of various performance metrics such as accuracy, precision, recall, specificity, and the F-Measure, which offer a more nuanced understanding of the classifier's performance beyond just accuracy [49].

Table 2 A standard binary classification confusion matrix

Confusion Matrix		Predicted Results	
		Positive	**Negative**
Actual Values	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

4.4 Evaluation Metrics

The primary objective of predicting student outcomes as "Low" or "High" is to identify students in the "Low" category after their midterm exams. This early identification is critical as it enables instructors to take proactive measures and provide targeted support to help these students improve their performance by the end of the course. Given the imbalance in the dataset, where "High" outcomes are more frequent than "Low" outcomes, the choice of evaluation metrics becomes essential. While accuracy is widely used, it may not provide an accurate representation in such cases. To ensure a more balanced assessment, specificity and the F-measure are utilized, offering deeper insights into the classifier's performance and its ability to handle imbalanced data effectively. Table 3 presents the formulas for calculating the evaluation metrics using the confusion matrix.

Table 3 A list of evaluation metrics used in this research

Metric	Formula
Accuracy	(TP+TN)/(TP+FN+FP+TN)
Specificity	TN/(TN+FP)
F-beta Measure	((1 + beta^2) * Precision * Recall) / (beta^2 * Precision + Recall)

Accuracy measures the proportion of correctly classified instances out of the total instances. It indicates how often the classifier is correct overall: In this context, accuracy helps provide a general sense of the classifier's performance. However, in imbalanced datasets, a high accuracy can be achieved by correctly predicting the majority class ("High") most of the time, while failing to identify the minority class ("Low"). Despite its limitations, accuracy is included as it is one of the most widely used metrics and offers a baseline for comparison. Specificity measures the proportion of actual negatives correctly identified. In this context, specificity is crucial as it reflects how well the classifier identifies students who are likely to perform "Low." High specificity means that the classifier is effective at minimizing false positives, ensuring that students predicted to perform poorly genuinely need support. This focus on accurately identifying the "Low" performers is essential for targeted interventions and support.

In this study, the F2 Measure takes precedence. The F-Measure combines precision and recall into a single score, with the F1 score being their harmonic mean. Variants like F0.5 and F2 adjust the balance between precision and recall based on specific needs. F0.5 places greater emphasis on precision, making it ideal when minimizing false positives is critical, ensuring highly accurate positive predictions. On the other hand, the F2 Measure prioritizes recall, making it better suited when reducing false negatives is more important. In this context, the F2 Measure is crucial as it ensures that students likely to perform "Low" are correctly identified, minimizing the risk of overlooking those in need of assistance.

5. METHODOLOGY

The methodology for this study involved several phases to ensure robust model performance. Initially, data preprocessing was conducted, including feature removal, handling missing values, and noise reduction to prepare a clean and comprehensive dataset. Feature selection utilized Information Gain to identify and retain the most relevant attributes, enhancing model accuracy and efficiency. Subsequently, various classifiers—Naive Bayes, Support Vector Machine (SVM), Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), and Decision Tree—were evaluated using 10-fold cross-validation [50] on the refined dataset to determine the most effective model. Model evaluation involved validating predictions on a prediction dataset to assess accuracy and identify misclassification patterns. Finally, the chosen model was integrated into a desktop application using JavaScript, allowing for efficient and user-friendly prediction of student outcomes based on input data.

5.1. Data Preprocessing

In this phase, several operations were conducted to ensure the dataset was suitable for analysis and model training. Initially, irrelevant features were removed to simplify the model and enhance its performance. Features that did not contribute significantly to the predictive model were identified and eliminated. Missing values were addressed by removing records with missing data, thereby ensuring the dataset's completeness. Noise reduction techniques were applied to identify and remove outliers or noisy data that could potentially distort the model's performance. The table 4 presents the features of the dataset.

Table 4 The training dataset with all features

Feature	Description	Data Type
Gender	Gender of the student	nominal
Degree	Degree program	nominal
Major	Major subject	nominal
Year	Year of study	nominal
MidExam	Marks obtained in the Midterm Exam	numeric
Registered_Course	Number of registered courses	numeric
Prev_Sem_GPA	GPA obtained in the previous semester	numeric
CGPA	Cumulative GPA	numeric
Hostel	Whether the student resides in a hostel	nominal
Grade	Binary classification of the grade	High, Low

The subsequent step in this phase involved feature selection, with the primary objective of identifying and retaining the most impactful features in the dataset that play a significant role in enhancing the predictive model's performance [51]. Feature selection is essential for reducing model complexity, which in turn helps to minimize the risk of overfitting [52]. Additionally, feature selection improves computational efficiency by lowering the dataset's dimensionality, resulting in quicker training and prediction processes. The technique used for feature selection in this study was InfoGain (Information Gain) Attribute Evaluation.

Information Gain is a feature selection metric that identifies and retains significant features, reducing data dimensionality and improving classification performance [53]. It quantifies the contribution of each feature in predicting the target variable. In WEKA, the Information Gain attribute evaluator ranks features by assessing the information they contribute to the class label, enhancing classification accuracy by eliminating redundant and irrelevant features [54]. Features that contribute more to reducing uncertainty in the target variable are ranked higher. The table 5 provides the features ranked by the Information Gain attribute evaluator. Information Gain was selected as our feature selection method because it effectively identifies features that significantly contribute to reducing uncertainty about the target variable, thus improving the model's predictive accuracy. It is particularly well-suited for handling categorical data, which is prevalent in our dataset. Information Gain streamlines the feature selection process by ranking features according to their contribution to class separation, reducing the risk of overfitting and promoting a more interpretable model. Its effectiveness across various classification problems further validates its role in improving model performance.

Table 5 Ranked Attributes Based on Information Gain

Rank	Attribute	InfoGain
1	Prev_Sem_GPA	0.29612
2	MidExam	0.212228
3	CGPA	0.18578
4	Year	0.046815
5	Registered_Course	0.030815
6	Major	0.023433
7	Gender	0.005324
8	Degree	0.00519
9	Hostel	0.000194

The final step focused on selecting key features based on their information gain values. This feature selection process is vital to ensure that only the most significant features are incorporated into the final model. The main reasons for selecting a few features are to maintain model simplicity, improve performance, and enhance computational efficiency. Reducing data dimensions and selecting appropriate feature sets simplifies the model, enhances interpretability and maintainability, and leads to better generalization on unseen data by improving classification accuracy and removing redundant and irrelevant features [55]. Eliminating irrelevant or redundant features reduces data noise, resulting in improved model performance and increased accuracy [56]. Additionally, fewer features mean less computational overhead, translating to faster training and prediction times. We specifically chose features with computable values, as these are easier for students to control and influence [57]. For instance, features like MidExam and CGPA scores are directly influenced by the students' efforts and performance, unlike incomputable features such as Gender. The table 6 provides the final set of features along with their descriptions.

Table 6 Final Set of features and their descriptions

Feature	Description
MidExam	Marks obtained in the Midterm Exam
Prev_Sem_GPA	GPA obtained in the previous semester
CGPA	Cumulative GPA
Grade	High, Low

This comprehensive data processing and feature selection phase highlights the rigorous steps taken to ensure the dataset's suitability for analysis and the strategic selection of features to enhance model performance.

5.2. Model Selection

This experimental evaluation aimed to assess the performance of various machine learning classifiers on the training dataset to identify the most suitable model for the application. The output of the previous phase was a training dataset that had undergone preprocessing and feature selection to ensure it was suitable for analysis and model training. In this phase, we aimed to apply various classifiers to this refined dataset to identify which model delivered the best performance metrics.

We convert the training dataset into ARFF file format, as it is a crucial step for utilizing WEKA effectively, which supports various file formats, including CSV and ARFF, essential for using the software suite efficiently [58]. To ensure a thorough evaluation, we used 10-fold cross-validation, a method that splits the dataset into ten segments, trains the model on nine segments, and tests it on the remaining one, repeating this process ten times [59]. This approach provided a more reliable assessment of the model's performance by minimizing the variance linked to a single train-test split.

We utilized five classifiers on the dataset—Decision Tree, Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), and Naive Bayes—each selected for its specific strengths in managing various data types and patterns. For each classifier, we obtained a confusion matrix, a table that summarized the performance of the model. This confusion matrix was then used to evaluate the overall effectiveness of the classifiers in classifying the data. We focus on metrics such as accuracy, Specificity, and F-Measure to evaluate and compare the classifiers, as these metrics are more meaningful in the context of correctly identifying students likely to perform "Low." Table 7 provides the confusion matrixes for the classifiers:

Table 7 The confusion matrixes of the classifiers

Classifier	TP	FN	FP	TN
Naive Bayes	135	25	13	35
Support Vector Machines (SVM)	154	6	25	23
Artificial Neural Networks (ANN)	148	12	22	26
K-Nearest Neighbors (KNN)	139	21	24	24
Decision Tree	148	12	18	30

Figure 1 demonstrates the accuracy achieved by the classifiers. The comparative analysis of classifier accuracy reveals distinct differences in performance across the five classifiers evaluated. Decision Tree emerged as the most accurate classifier, achieving an accuracy of 85.6%. This indicates a strong ability to correctly classify both "High" and "Low" performing students. SVM followed closely with an accuracy of 81.1%, demonstrating robust performance as well. Naive Bayes and ANN exhibited similar accuracies, with Naive Bayes at 81.7% and ANN at 83.7%, showing they are effective but slightly less reliable compared to Decision Tree and SVM. KNN had the lowest accuracy at 78.4%, suggesting it struggles more with accurately predicting student performance. These results highlight that while Decision Tree and SVM are highly effective in handling this classification task, KNN's lower accuracy indicates it may not be the best choice for predicting student outcomes in this context. Despite accuracy being a commonly used metric, it is important to consider other evaluation metrics, especially in the presence of class imbalance, to ensure a comprehensive assessment of classifier performance.

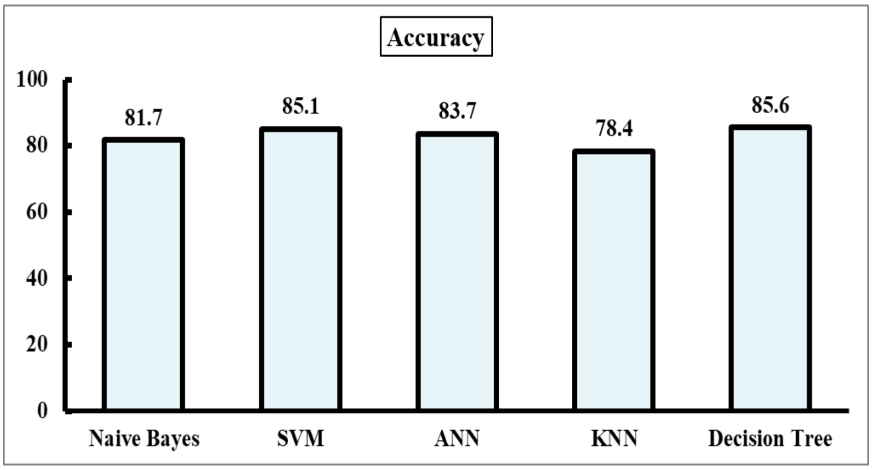

Figure 1 A comparison of the classifier's accuracy

Figure 2 shows the specificity comparison of the five classifiers, highlighting their effectiveness in correctly identifying "Low" performing students. Naive Bayes demonstrated the highest specificity at 72.9%, showcasing its effectiveness in accurately identifying true negatives while minimizing false positives. This makes it particularly effective for targeted interventions to support at-risk students. Decision Tree also performed well with a specificity of 62.5%, demonstrating its reliability in distinguishing between the classes. The ANN classifier achieved a specificity of 54.2%, showing moderate performance but still being a viable option for recognizing "Low" performers. SVM and KNN, however, exhibited lower specificities of 47.9% and 50.0%, respectively, suggesting that these models struggle more with accurately identifying students likely to perform poorly. These findings highlight the significance of choosing a classifier with high specificity in scenarios where accurately identifying "Low" performing students is essential for timely and effective educational interventions.

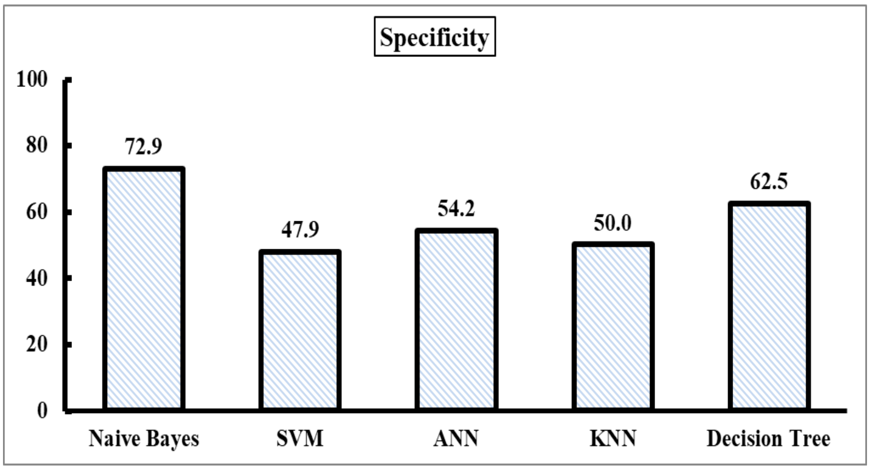

Figure 2 Comparison of classifiers' specificity

Figure 3 compares the F2 measure of the five classifiers, highlighting their effectiveness in prioritizing recall over precision, which is crucial for ensuring that most "Low" performing students are identified. The SVM has highest F2 measure of 94% demonstrating their strong ability to capture the majority of at-risk students while maintaining a reasonable level of precision, making them highly effective for early identification and intervention purposes. It is followed by Decision Tree and ANN with 91.8% and 94% respectively. Naive Bayes and KNN both achieved an F2 measure of 82.7% and 86.6% respectively, which, while slightly lower than the others, still shows a good balance between identifying true positives and minimizing false negatives. These results underscore the importance of the F2 measure in educational contexts where the primary goal is to ensure that as many "Low" performing students as possible are identified for timely support. The higher F2 scores of SVM and Decision Tree indicate that these classifiers are particularly well-suited for identifying at-risk students, effectively minimizing the risk of overlooking those who require intervention. The slightly lower F2 measures for Naive Bayes and KNN indicate that while they are still effective, they may not be as comprehensive in capturing all students who are at risk. Overall, the F2 measure provides a valuable perspective on the classifiers' performance in ensuring broad coverage of at-risk students, which is essential for effective educational interventions.

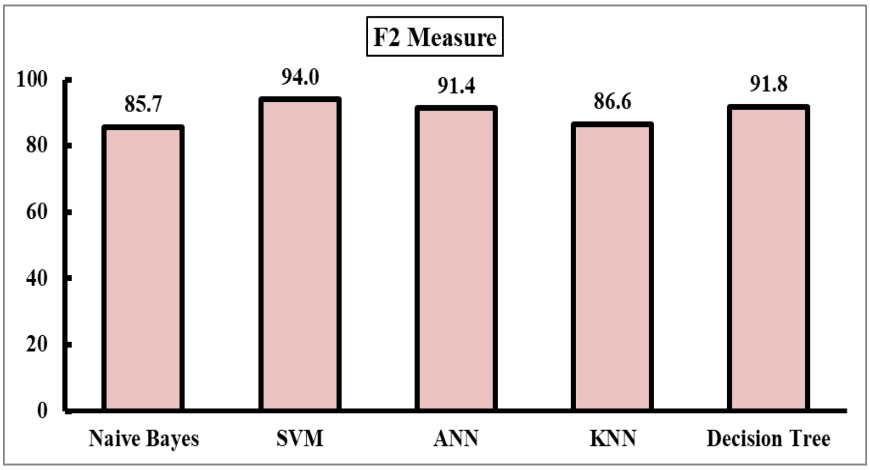

Figure 3 Comparative analysis of classifiers' F2 measures

From the comparative analysis of accuracy, specificity, and the F2 measure, the Support Vector Machine (SVM) stands out as the most effective classifier for identifying students at risk of failing the course. With an accuracy of around 85%, SVM demonstrates a robust ability to accurately classify both "High" and "Low" performing students. It also attained the F2 measure at 94.0%, showcasing demonstrating its strong ability to capture the majority of at-risk students while maintaining a reasonable level of precision. These metrics collectively underscore SVM robustness in identifying "Low" performing students, making it particularly effective for targeted interventions. Overall, the SVM classifier stands out as the most reliable and comprehensive model for predicting student outcomes and providing the necessary support to those at risk, ensuring timely and effective educational interventions.

5.3. Model Evaluation

We conducted a rigorous evaluation of the selected machine learning model using the WEKA platform to validate its predictive performance. The evaluation process involved preparing a validation dataset that was identical to the training dataset, except for the final outcome/class feature, which was intentionally left unspecified (denoted as "?"). The model was required to predict these unspecified outcomes, and the predictions were subsequently compared against the actual outcomes to assess the model's accuracy and reliability in a real-world context. The confusion matrix generated from this evaluation, presented in table 8, provides a detailed breakdown of true positives, true negatives, false positives, and false negatives.

Table 8 Confusion matrix of model evaluation phase

	Predicted High	Predicted Low
Actual High	26	2
Actual Low	7	10

The model achieved an accuracy of 80%, a key metric that represents the proportion of correctly predicted instances out of the total instances. This high level of accuracy demonstrates the model's robust ability to classify instances correctly in a majority of cases. When considering the total number of instances (45), this specific misclassification rate translates to approximately 13.3%, indicating that out of all instances, 13. 3% were incorrectly predicted as "High" when they were actually "Low."

It is crucial to understand the impact of different types of misclassifications. In this context, predicting "High" as "Low" is not as critical as predicting "Low" as "High." Misclassifying a "Low" instance as "High" can be more detrimental because it may lead to overestimating an outcome's importance or potential, which can have significant consequences, depending on the application. Conversely, predicting "High" as "Low" is less severe in most cases, as it might only lead to underestimating the potential, which can be managed with additional assessments or conservative approaches.

The table 9 lists the records of students who were wrongly identified as "High" but were actually "Low." This analysis aims to identify possible reasons for the misclassification by the model.

Table 9 The actual data of student to investigate False Positive

Record	MidExam	Prev_Sem_GPA	CGPA
1	18	1.98	2.12
2	13	1.9	1.67
3	18	2.03	2.46
4	16	1.9	2.52
5	17	2.4	1.88
6	18	2.30	2.84
7	19	2.36	2.71

The model's misclassification of these "Low" students as "High" appears to be influenced by relatively high exam and assignment scores in certain instances. The midterm and final exam scores for these students are not exceptionally low, with some students having moderately high scores, which may have influenced the model to predict "High." For example, the student in record 5 has a very high final exam score of 33, which likely skewed the model's prediction towards "High." Additionally, the GPAs (both previous semester GPA and CGPA) for these students are relatively low but not consistently so across all records. For instance, records 3 and 4 have CGPAs above 2.4, which might have contributed to the model predicting "High." Furthermore, assignment scores are generally high for these students, suggesting that the model might heavily weigh assignment performance when predicting the final classification. These factors collectively indicate areas where the model could be adjusted or improved to better handle such cases and reduce the rate of misclassification.

In conclusion, while the model shows strong overall accuracy at approximately 87%, attention is needed to address the misclassification of "Low" instances as "High." With a misclassification rate of about 13%, these errors should be closely monitored and reduced to minimize their impact. Enhancing the model's precision in differentiating between "High" and "Low" classes will be crucial for boosting its reliability and effectiveness.

5.4. Model Integration

The final phase of this research is to transform the chosen model into a desktop application. The integration of the WEKA model into the software is a critical aspect of its functionality. WEKA was utilized to develop and train the prediction model. This trained model is then integrated into a desktop application using JavaScript. The desktop application is designed to call the WEKA model to perform prediction tasks. By leveraging the capabilities of the WEKA model, the software can efficiently process the input data and generate accurate predictions about student outcomes. This seamless integration ensures that the predictive power of the WEKA model is harnessed effectively within the user-friendly interface of the desktop application.

The software features two primary forms to facilitate predictions. The first form is designed to predict the final outcome for an entire class. Figure 4 illustrates this operation. Instructors begin by selecting the relevant course and then upload a prediction dataset in CSV format. This dataset must have identical features to the training dataset used for building the model. The form offers the flexibility to view the report on screen, download it as a CSV file, or both.

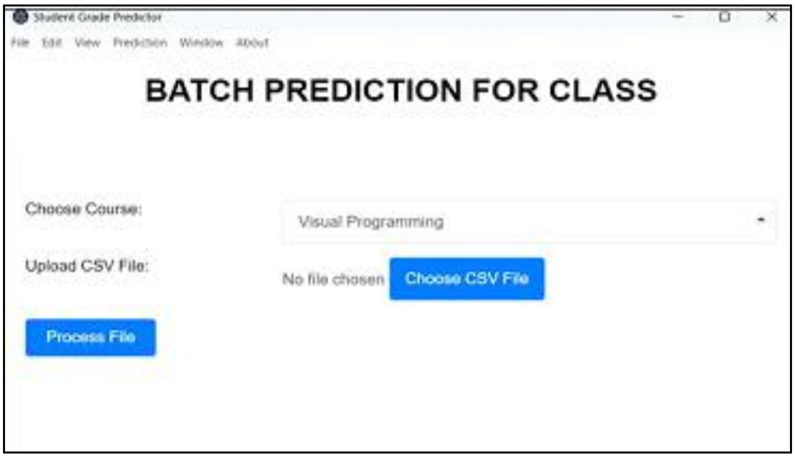

Figure 4 The interface to predict the final outcome of whole class

The second form, as demonstrated in figure 5, focuses on individual student predictions. Instructors input specific values for prediction features such as midterm exam grades, assignment grades, previous semester GPA, and CGPA. the application execute the prediction model for a file containing single instance and displays the student's final outcome as either "High" or "Low." The application is designed to be user-friendly, providing instructors with immediate, actionable insights. The software also includes additional features, such as the ability for instructors to view the technical details of the prediction models and perform various other supportive operations.

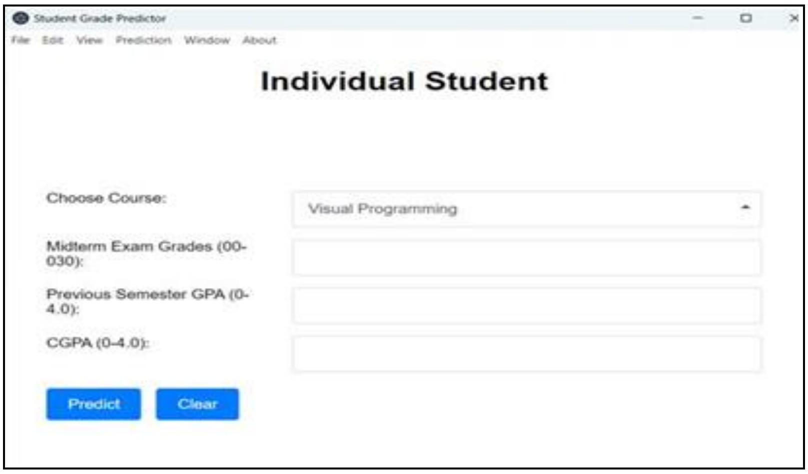

Figure 5 The interface to predict the outcome of single student

6. Conclusion

This study tackles the pressing need for efficient student performance monitoring in higher education through the application of machine learning. Our AI-based application, utilizing the Support Vector Machine (SVM) classifier, accurately predicts student outcomes based on academic data. This enables timely interventions, particularly for students likely to achieve unsatisfactory grades. The application's user-friendly interface ensures practical use for educators, enhancing their ability to support at-risk students and improve overall educational standards. Future work will focus on refining the model and expanding its application across diverse educational contexts to further validate its efficacy.

While this study adopts a supervised classification approach to identify struggling students, the problem could alternatively be framed as an anomaly detection task, treating underperforming students as rare deviations from the norm. This approach may be particularly useful for imbalanced datasets or scenarios with limited labeled data. Future work could explore the use of anomaly detection methods, such as Isolation Forests or One-Class SVM, to complement or enhance the current methodology.

Acknowledgement

This research was funded by the Ministry of Higher Education, Research, and Innovation (MoHERI) of the Sultanate of Oman under the Block Funding Program (BFP), Agreement No. MOHERI/BFP/URG/ICT/23/004.

AUTHORS

Maram Khamis Khalfan Al-Muharrami is a Bachelor of Software Engineering student with a strong interest in software development and Artificial Intelligence research. She consistently performs at a high level and is on the college honors list for her academic excellence.

Fatema Said Ali Al-Sharqi is a Bachelor of Software Engineering student with a strong interest in software development and research.

Al-Zahraa Khalid Said Al Rumhi is a Bachelor of Computer Science student with a keen interest in researching Artificial Intelligence and its applications in education.

Shamsa Said Mohammed Al-Mamari is a Bachelor of Computer Science student, recognized on the College Honor List, and a young researcher in Artificial Intelligence.

Dr. Ijaz Khan, a PhD in Artificial Intelligence, focuses on AI-driven teaching tools, supervised learning techniques, Wireless Sensor Networks, and Machine Learning. He is currently working on various projects to enhance AI applications in education. He is also an experienced educator and writer.

REFERENCES

[1] Timotheou, S., et al., *Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review.* Education and information technologies, 2023. **28**(6): p. 6695-6726.

[2] Al-Saadi, Z.T., *A review of graduate attributes in the Oman Authority for Academic Accreditation and Quality Assurance of Education (OAAAQAE's) quality audit reports.* Gulf Education and Social Policy Review (GESPR), 2023: p. 107–124-107–124.

[3] Liubchenko, V., N. Komleva, and S. Zinovatna. *Monitoring Student Performance Based on Educational Measurements*. in *International Conference on Interactive Collaborative Learning*. 2023. Springer.

[4] Khan, I., et al., *An artificial intelligence approach to monitor student performance and devise preventive measures.* Smart Learning Environments, 2021. **8**(1): p. 1-18.

[5] Silberglitt, B., D. Parker, and P. Muyskens, *Assessment: Periodic assessment to monitor progress*, in *Handbook of response to intervention: The science and practice of multi-tiered systems of support*. 2015, Springer. p. 271-291.

[6] Brown, G.T. *The past, present and future of educational assessment: A transdisciplinary perspective*. in *Frontiers in Education*. 2022. Frontiers Media SA.

[7] Macfadyen, L.P., et al., *Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge.* Research & Practice in Assessment, 2014. **9**: p. 17-28.

[8] Deo, R.C., et al., *Modern artificial intelligence model development for undergraduate student performance prediction: An investigation on engineering mathematics courses.* IEEE Access, 2020. **8**: p. 136697-136724.

[9] Zeineddine, H., U. Braendle, and A. Farah, *Enhancing prediction of student success: Automated machine learning approach.* Computers & Electrical Engineering, 2021. **89**: p. 106903.

[10] Hashim, A.S., W.A. Awadh, and A.K. Hamoud. *Student performance prediction model based on supervised machine learning algorithms*. in *IOP conference series: materials science and engineering*. 2020. IOP Publishing.

[11] Lau, E.T., L. Sun, and Q. Yang, *Modelling, prediction and classification of student academic performance using artificial neural networks.* SN Applied Sciences, 2019. **1**(9): p. 982.

[12] Mondal, A. and J. Mukherjee, *An Approach to predict a student's academic performance using Recurrent Neural Network (RNN).* Int. J. Comput. Appl, 2018. **181**(6): p. 1-5.

[13] Pallathadka, H., et al., *Classification and prediction of student performance data using various machine learning algorithms.* Materials today: proceedings, 2023. **80**: p. 3782-3785.

[14] Sukhbaatar, O., T. Usagawa, and L. Choimaa, *An artificial neural network based early prediction of failure-prone students in blended learning course.* International Journal of Emerging Technologies in Learning (iJET), 2019. **14**(19): p. 77-92.

[15] Alcaraz, R., et al., *Early prediction of students at risk of failing a face-to-face course in power electronic systems.* IEEE Transactions on Learning Technologies, 2021. **14**(5): p. 590-603.

[16] Namoun, A. and A. Alshanqiti, *Predicting student performance using data mining and learning analytics techniques: A systematic literature review.* Applied Sciences, 2020. **11**(1): p. 237.

[17] Sayaf, A.M., et al., *Information and communications technology used in higher education: An empirical study on digital learning as sustainability.* Sustainability, 2021. **13**(13): p. 7074.

[18] Yağcı, M., *Educational data mining: prediction of students' academic performance using machine learning algorithms.* Smart Learning Environments, 2022. **9**(1): p. 11.

[19] Villegas-Ch, W., M. Román-Cañizares, and X. Palacios-Pacheco, *Improvement of an online education model with the integration of machine learning and data analysis in an LMS.* Applied Sciences, 2020. **10**(15): p. 5371.

[20] Gligorea, I., et al., *Adaptive learning using artificial intelligence in e-learning: a literature review.* Education Sciences, 2023. **13**(12): p. 1216.

[21] Ismail, S.N., et al., *Exploring students engagement towards the learning management system (LMS) using learning analytics.* Computer systems science & engineering, 2021. **37**(1).

[22] Wongwatkit, C., et al., *The Future of Connectivist Learning with the Potential of Emerging Technologies and AI in Thailand: Trends, Applications, and Challenges in Shaping Education.* Journal of Learning Sciences and Education, 2023. **2**(1): p. 122-154.

[23] Liang, W., et al., *Advances, challenges and opportunities in creating data for trustworthy AI.* Nature Machine Intelligence, 2022. **4**(8): p. 669-677.

[24] Saura, J.R., D. Ribeiro-Soriano, and D. Palacios-Marqués, *Assessing behavioral data science privacy issues in government artificial intelligence deployment.* Government Information Quarterly, 2022. **39**(4): p. 101679.

[25] Watson, J., et al., *Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: what can we learn from US academic medical centers?* JAMIA open, 2020. **3**(2): p. 167-172.

[26] Ferrara, E., *Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies.* Sci, 2023. **6**(1): p. 3.

[27] Charbuty, B. and A. Abdulazeez, *Classification based on decision tree algorithm for machine learning.* Journal of Applied Science and Technology Trends, 2021. **2**(01): p. 20-28.

[28] Dhall, D., R. Kaur, and M. Juneja, *Machine learning: a review of the algorithms and its applications.* Proceedings of ICRIC 2019: Recent innovations in computing, 2020: p. 47-63.

[29] Kost, S., O. Rheinbach, and H. Schaeben, *Using logistic regression model selection towards interpretable machine learning in mineral prospectivity modeling.* Geochemistry, 2021. **81**(4): p. 125826.

[30] Stiglic, G., et al., *Comprehensive decision tree models in bioinformatics.* PloS one, 2012. **7**(3): p. e33812.

[31] Abbas, N., Y. Nasser, and K.E. Ahmad, *Recent advances on artificial intelligence and learning techniques in cognitive radio networks.* EURASIP Journal on Wireless Communications and Networking, 2015. **2015**: p. 1-20.

[32] Feng, G., et al., *Feature subset selection using naive Bayes for text classification.* Pattern Recognition Letters, 2015. **65**: p. 109-115.

[33] Dongare, A., R. Kharde, and A.D. Kachare, *Introduction to artificial neural network.* International Journal of Engineering and Innovative Technology (IJEIT), 2012. **2**(1): p. 189-194.

[34] Seni, G. and J. Elder, *Ensemble methods in data mining: improving accuracy through combining predictions*. 2010: Morgan & Claypool Publishers.

[35] Musso, M.F., C.F.R. Hernández, and E.C. Cascallar, *Predicting key educational outcomes in academic trajectories: a machine-learning approach.* Higher education, 2020. **80**(5): p. 875-894.

[36] Hussain, S. and M.Q. Khan, *Student-performulator: Predicting students' academic performance at secondary and intermediate level using machine learning.* Annals of data science, 2023. **10**(3): p. 637-655.

[37] Alhazmi, E. and A. Sheneamer, *Early predicting of students performance in higher education.* IEEE Access, 2023. **11**: p. 27579-27589.

[38] Xu, X., et al., *Prediction of academic performance associated with internet usage behaviors using machine learning algorithms.* Computers in Human Behavior, 2019. **98**: p. 166-173.

[39] Ojajuni, O., et al. *Predicting student academic performance using machine learning*. in *Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part IX 21*. 2021. Springer.
[40] Adnan, M., et al., *Predicting at-risk students at different percentages of course length for early intervention using machine learning models.* Ieee Access, 2021. **9**: p. 7519-7539.
[41] Khan, I., et al., *An artificial intelligence approach to monitor student performance and devise preventive measures.* Smart Learning Environments, 2021. **8**: p. 1-18.
[42] Costa, E.B., et al., *Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses.* Computers in human behavior, 2017. **73**: p. 247-256.
[43] Cano, A. and J.D. Leonard, *Interpretable multiview early warning system adapted to underrepresented student populations.* IEEE Transactions on Learning Technologies, 2019. **12**(2): p. 198-211.
[44] Gupta, S. and A.S. Sabitha, *Deciphering the attributes of student retention in massive open online courses using data mining techniques.* Education and Information Technologies, 2019. **24**(3): p. 1973-1994.
[45] Khan, I., et al. *Tracking student performance in introductory programming by means of machine learning*. in *2019 4th mec international conference on big data and smart city (icbdsc)*. 2019. IEEE.
[46] Sharma, P., et al. *Student engagement detection using emotion analysis, eye tracking and head movement with machine learning*. in *International Conference on Technology and Innovation in Learning, Teaching and Education*. 2022. Springer.
[47] Alam, A. and A. Mohanty. *Predicting students' performance employing educational data mining techniques, machine learning, and learning analytics*. in *International Conference on Communication, Networks and Computing*. 2022. Springer.
[48] Kiranmai, S.A. and A.J. Laxmi, *Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy.* Protection and Control of Modern Power Systems, 2018. **3**(1): p. 1-12.
[49] Tharwat, A., *Classification assessment methods.* Applied Computing and Informatics, 2018.
[50] Arlot, S. and A. Celisse, *A survey of cross-validation procedures for model selection.* Statistics surveys, 2010. **4**: p. 40-79.
[51] Theng, D. and K.K. Bhoyar, *Feature selection techniques for machine learning: a survey of more than two decades of research.* Knowledge and Information Systems, 2024. **66**(3): p. 1575-1637.
[52] Xie, J., M. Sage, and Y.F. Zhao, *Feature selection and feature learning in machine learning applications for gas turbines: A review.* Engineering Applications of Artificial Intelligence, 2023. **117**: p. 105591.
[53] Omuya, E.O., G.O. Okeyo, and M.W. Kimwele, *Feature selection for classification using principal component analysis and information gain.* Expert Systems with Applications, 2021. **174**: p. 114765.
[54] Mishra, S., et al., *Performance evaluation of a proposed machine learning model for chronic disease datasets using an integrated attribute evaluator and an improved decision tree classifier.* Applied Sciences, 2020. **10**(22): p. 8137.
[55] Wang, J., et al., *Generalizing to unseen domains: A survey on domain generalization.* IEEE transactions on knowledge and data engineering, 2022. **35**(8): p. 8052-8072.
[56] Ahuja, A., L. Al-Zogbi, and A. Krieger, *Application of noise-reduction techniques to machine learning algorithms for breast cancer tumor identification.* Computers in Biology and Medicine, 2021. **135**: p. 104576.
[57] Khan, I., et al., *A Conceptual Framework to Aid Attribute Selection in Machine Learning Student Performance Prediction Models.* International Journal of Interactive Mobile Technologies, 2021. **15**(15).
[58] Attwal, K.P.S. and A.S. Dhiman, *Exploring data mining tool-Weka and using Weka to build and evaluate predictive models.* Adv. Appl. Math. Sci, 2020. **19**(6): p. 451-469.
[59] Khan, I., et al., *Machine Learning Prediction and Recommendation Framework to Support Introductory Programming Course.* International Journal of Emerging Technologies in Learning, 2021. **16**(17).