A MACHINE LEARNING FRAMEWORK FOR STUDENT PERFORMANCE FORECASTING: BEST FIRST SEARCH WITH LOGISTIC REGRESSION AND DECISION TABLE IN EDUCATIONAL DATA MINING

Anosha[1], Irfana Shahab[2], Awrang Zaib[3], Malak Roman[4*], Jawad Hussain Shah[5], Abdul Qadir Khan[6]

[1,2&5]*BS-Computer Science, Department of Computer Science, University of Chitral, Pakistan.*
[3&4]*Lecturer, Department of Computer Science, University of Chitral, Pakistan.*
[6]*Department of Computer Science, GDC Wari Dir-Upper, Pakistan.*
[*4]*malak_5116@uoch.edu.pk*

DOI: https://doi.org/10.5281/zenodo.17926225

Keywords:
Educational Data Mining, Machine Learning, Best First Search, Decision Table, Logistic Regression.

Article History
Received on 10 Nov, 2025
Accepted on 09 Dec 2025
Published on 12 Dec 2025

Corresponding Author:
Malak Roman

Abstract

Predicting student academic performance is a critical priority for educational institutions striving to enhance learning outcomes and provide timely support for students at risk of failure. Machine learning offers robust techniques for analyzing student data and identifying influential factors that shape academic achievement. This study conducts a comparative evaluation of two machine learning classification algorithms Logistic Regression (LR) and Decision Table (DT) to predict students' final grades (G3). Using the Students' Academic Performance dataset from the UCI Machine Learning Repository, which includes 395 records and 33 attributes, the research employs the Best-First Search (BFS) method for feature selection. BFS effectively reduces the dataset to 13 key attributes, encompassing demographic details, personal factors, and prior academic performance indicators (G1 and G2). Both LR and DT models were developed and assessed using evaluation metrics such as correlation coefficient, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). The results demonstrate that the Logistic Regression model outperforms in terms of RMSE, showing a stronger correlation and lower error rates. Overall, the study contributes to the field of Educational Data Mining (EDM) by providing insights into predictive modeling approaches that can support early identification and intervention for students requiring academic assistance.

INTRODUCTION

Artificial Intelligence (AI), commonly referred to as machine learning (ML), enabling computers to develop intelligent behavior and improve process performance without explicit programming [1]. Unlike traditional rule-based systems, ML offers a more adaptable approach by learning from example data and making predictions based on the patterns it discovers. In educational settings, machine learning has a significant impact on examining and forecasting students' academic performance via utilizing past records, social indicators, and knowledge patterns [2]. Through predictive models, institutions can identify at-risk students early and provide timely interferences such as tutoring, curriculum adjustments, or counseling to support their success. Data mining is termed as acquiring useful and valuable information from complex data using a combination of statistical and computational tools. With the help of these tools, organizations can make predictions about future developments in complex data and facilitate intelligent multidimensional decision-making in numerous fields such as telecommunications, finance, transport, and insurance [3].

Academic performance refers to the information, abilities, and competencies that pupils develop with in a particular subject or academic discipline [4]. It is evaluated through systematic impost and finding of students' development over a specified period [5]. Academic performance serves as a key pointer of how effectively an educational system is achieving its intended goals [6,7]. Despite substantial government investment in higher education, many students particularly in developing countries continue to struggle academically and fail to meet expected performance standards [8]. A cross-sectional study conducted in Malaysia reported that an increasing number of higher education students fail to graduate on time, indicating inadequate academic achievement [9]. Similarly, a correlative study at Arba Minch University in southern Ethiopia found that the number of students graduating was not in line with enrollment trends, and an increase in readmissions due to poor academic performance was observed [7]. Although accurate dropout statistics are not available for Iran, limited reports from university officials estimate that approximately 1,500 students drop out of public universities each year [10].

Declining academic performance has become a significant challenge for educational systems, leading to wasted resources, student frustration, and broader negative consequences for families, institutions, and society [7]. Evidence indicates that many factors influence educational outcomes, including poor educational choices, particularly inappropriate selection of academic fields lack motivation, and inadequate academic engagement. These factors are among the primary contributors to low academic performance and high dropout rates in higher education [5, 7, 11,12]. In addition to predicting outcomes and probabilities, modern Intelligent Tutoring Systems (ITS) and learning analytics dashboards support self-regulated learning by providing constructive feedback. Nevertheless, issues such as data privacy, algorithmic bias, and interpretability of models must be addressed to ensure ethical and responsible implementation. Advances in natural language processing (NLP) have further expanded the role of ML [13], enabling automated assessment of essay responses and analysis of student participation in discussion forums to predict academic performance. As educational technologies advance, machine learning is closing the gap between students' ability and achievement, thereby fostering more personalized, equitable, and efficient learning environments [14]. Accurate rating calculation and dropout-risk estimation require high-quality datasets, as real-world educational data is often noisy, imprecise, or imbalanced [15].

In this study, we apply Best-First Search (BFS) as a preprocessing method alongside classifiers such as Logistic Regression and Decision Tables. By identifying the most relevant features in the dataset, BFS helps reduce processing time and minimize data dimensionality without compromising the integrity or accuracy of analytical results.

LITERATURE REVIEW

Roman, M. et al., (2025) [2] Performed explore work on students' academic performance using Decision Tree and Random Forest machine learning algorithms. The dataset was obtained from the UCI Machine Learning Repository, containing 395 student records with 33 attributes. They implemented a Best First search algorithm for feature selection, which reduced the attributes from 33 to 11. After data preprocessing, the Decision Tree algorithm yielded 1.92 RMSE, the best results compared to the Random Forest algorithm 0.90.

Yusoff, S. M. et al., (2024) [16] Explore academic performance prediction using various machine learning approaches. The dataset was taken from Kaggle and consisted of 33 attributes. In the preprocessing phase, they used Correlation attributes Eval and Gain Ratio for attribute selection. Out of five machine learning algorithms used, the algorithms are OneR, Attribute Selected Classifier, J48, Naive Bayes, MLP (Multilayer Perceptron). Among the classifiers OneR achieves

higher accuracy of 55.2% while MLP stands worst with 31.03%.
Pan, J. et al., (2025) [17] works on academic performance prediction, the datasets were collected from the Open University Learning Analytics Dataset (OULAD) and the Chinese University dataset (CUD). Preprocessing techniques included Digitization Direct Discretization, Removal of Duplicate data, and Balancing Samples. Multiple algorithms were implemented on both datasets with the following classifiers Gradient Boosted Tree (GBT), multiclass logistic regression (MLG), Neural network (NN), Naive Bayes (NB), K-Nearest Neighbors (KNN), Random Forest (RF), Decision tree (DT). For the CUD, the best algorithm was Gradient Boosted Tree 43%, and the worst was Random Forest 31%. For the OULAD, the best algorithm was Neural Network 41%, and the worst was Naïve Bayes 81%.
Ahmed, E. (2024) [18] Collected dataset from Wollo University and the Kombolcha Institute of Technology, containing 8 attributes. After data cleaning, data categorization, data reduction, and feature selection (RF Algorithm), the following classifiers were implemented SVM, Decision tree (DT), Naive Bayes (NB) and K-Nearest neighbors (KNN). Among the classifiers Support Vector Machine shows best results with 96% accuracy.
Sandeepa, A. G. R., & Mohottala, S. (2025) [19] Conducted research focused on the assessment and comparative performance of machine learning models for predicting secondary students' academic outcomes. The data was assembled from Jordan High School Students using a LMS. Preprocessing techniques included Label Encoding, Standardization, and Target Variable Encoding. The multi-layers perceptron classifier (MCPc), Support vector machine (SVM), Decision tree (DT), Naive Bayes (NB) and K- Nearest neighbors (KNN), Logistic regression (LR), were used as a machine learning algorithm The best result was given by MLPC7 95.8%, and the LR 67% result was depraved as compared to other algorithms.
Kalita, E. et al., (2025) [20] Carried out research focusing on predicting academic scores using Bi-LSTM, a deep learning framework with SHAP-based interpretability and statistical validation. The dataset was downloaded from a Middle Western university in the USA. Preprocessing techniques used were Data Augmentation, Synthetic Minority Over-Sampling, Various machine learning algorithm vertex tested such a Catboost, XGBoost, HistGradient Boosting, lightGBM, and Bi-LSTM were used as a deep learning model. Bi-LSTM progressed accuracy of 88%, while Light GBM attained 86%.
Agyemang, E. F. et al., (2024) [21] steered work on predicting students' performance. The dataset was obtained from XAPI-Edu Kaggle repository, containing 16 attributes. Preprocessing techniques included hyperparameter tuning and standard scalar. Among the following machine learning algorithms Support vector machine (SVM), Decision tree (DT), K- Nearest neighbors (KNN), Random Forest (RF) and logistic regression (LR), Random Forest shows good result with 85% and Logistic Regression attains lowest results with 81% accuracy.
Jović, J. et al., (2022) [22] Piloted their research on predicting student hypothetical performance using machine learning algorithms. The dataset was obtained from the Learning Management System (LMS) and Educational Management System (EMS) at Belgrade Metropolitan University, having totally 7 attributes. Preprocessing techniques used were Min Max Normalization, Integration, Data cleaning, and Feature selection. The support vector machine (SVM), Decision tree (DT), Naive Bayes (NB), Linear Discriminant Analysis (LDA) and K- Nearest neighbors (KNN) were used as classifier models. Support Vector Machine (SVM) achieved 88% accuracy and stands best among all the classifiers.
Yadav, N. R., & Deshmukh, S. S. (2023) [23] performed their research work on the prediction of student performance using Support vector machine (SVM), Naive Bayes (NB), C4.5 and ID3 classifiers as machine learning models. The dataset was obtained from UCI repository and processed with the Synthetic Minority Oversampling Technique (SMOTE) algorithm. Support Vector Machine (SVM) achieves 88% accuracy, and the wickedest was Decision Tree (DT) scoring 60%.
Yang, Y. et al., (2025) [24] performed research work on student academic performance prediction constructing the following ML models Tab Net, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Ad boost, Decision Tree (DT), Random Forest (RF), Hyper graph Neural Network (HGNN) Grip convolutional network (GCN) and Hyper Tab. The dataset was taken from a University Information System and preprocessed with Anomaly removal, Missing value handling, Sample imbalance correction, Feature engineering. Hyper Tab scored 93% of accuracy among all models while Decision Tree (DT) scored lowest i.e. 47%.
Taher Mazandarani, M. et al., (2025) [25] performed research work on predicting student performance using a machine learning approach and feature analysis with

ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 12, 2025

Support vector machine (SVM), Decision tree (DT), K-Nearest neighbors (KNN), Random Forest (RF) and cat Boost. The data was collected from Iranian schools through multiple surveys, containing 3701 student records and 197 features. Cat Boost Regress achieves 87% accuracy and again Decision Tree (DT) scores lowest with 75%.

RESEARCH METHODOLOGY

The dataset titled Students' Academic Performance and Grade Prediction was sourced from the UCI Machine Learning Repository and consisted of 33 attributes and 395 instances. Following data preprocessing, the dataset was divided into two subsets: one for model training and the other for testing [26]. For optimal feature selection, the Best-First Search (BFS) algorithm implemented using a greedy heuristic search strategy was applied. After feature selection, the refined dataset was used to train two models: Decision Table (DT) and Logistic Regression (LR) as illustrated in Figure 1. Each classification algorithm was evaluated independently to measure and compare model accuracy.

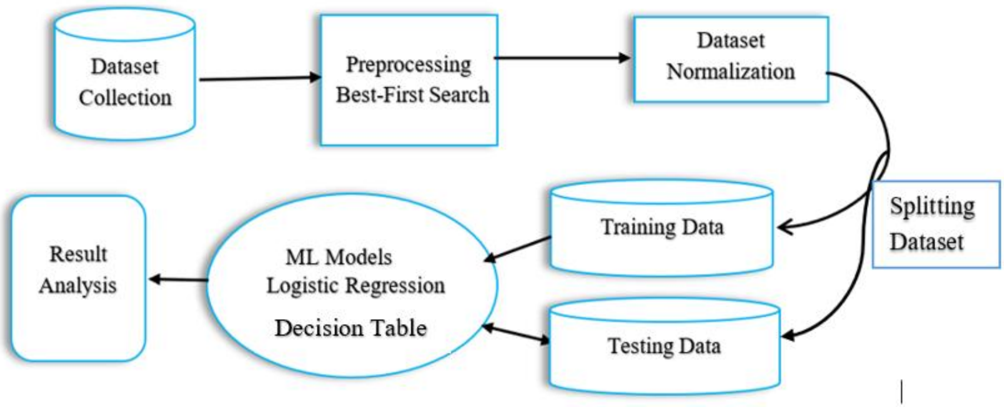

Fig. 1: **Research Methodology Flow Chart**

Dataset Collection

The Students' Academic Performance Prediction dataset comprises 33 attributes, with the detailed descriptions of these attribute values presented in Table 1.

Table 1: ***Dataset Attribute Description***

Attribute Description

S. No	Attribute Name	Value/Description	S. No	Attribute Name	Value/Description
1	'School'	'student's school'	18	'paid: extra classes'	'binary: yes or no'
2	'Gender'	'(binary: "F"- female or "M" – male)'	19	'Activities: extracurricular activities.'	'binary: yes or no'
3	'Age'	'Numeric'	20	'nursery: school'	'binary: yes or no'
4	'Address'	'(binary: "U"- urban or "R"- rural)'	21	'higher: interested in higher Edu'	'binary: yes or no'
5	'Famine'(family size)	'(binary: LE3<=3 or GT3>3)'	22	'internet - Internet access at home'	'binary: yes or no'
6	'status' (parental status)	'(binary: "T": living together or "A": apart)'	23	'romantic: relation'	'binary: yes or no'
7	'Med mother's education'	'(numeric: 0=none, 1=primary, 2=5th to 9th grade, 3=secondary, 4=higher Edu)	24	'Feud father education'	'(numeric: 0=none, 1=primary, 2=5th to 9th grade, 3=secondary, 4=higher Edu)
8	'free time - free time'	'(numeric: 1-5 (low to very high)'	25	'farmer: family relationships'	'numeric: 1-5 (bad to excellent'
9	'Mob mother's job'	'nominal'	26	'go out – going out with friends'	'numeric: 1-5 (low to very high'

10	'Fob father's job'	'nominal'	27	'Daly: use of alcohol'	'numeric: 1-5 (low to very high
11	'reason to choose this school'	'(Nominal: close to "home", "reputation", "course" preference or "other")	28	'Walk: weekend alcohol'	'numeric: 1-5 (low to very high'
12	'guardian'	'(nominal: "mother", "father" or "other")'	29	'health: current health status'	'numeric: 1-5 (bad to very good'
13	'Travel time'	(numeric: 1<15 min,2 15 to 30 min,3:1hrs, or 4>1 hrs.)'	30	'Studytime: weekly study time'	'(numeric: hrs 1<2, 2 : 2 to 5, 3: 5 to 10, or 4: >10)'
14	'absences'	'numeric: from 0 to 93'	31	'G1: 1st period grd'	'numeric: 0 to 20'
15	'Failures: past classes'	numeric: n if 1<=n<3 else 4'	32	'G2: 2nd period'	'numeric: 0 to 20'
16	'schoolsup: extra educ. support'	'binary: yes or no'	33	'G3: final grade'	'numeric: 0 to 20'
17	'famsup - family educ. support'	'binary: yes or no'			

Greedy Heuristic Search: Best-First Search

Feature selection is a method that reduces dimensionality, eliminates redundant attributes, and improves model prediction [27]. Best First Search is a heuristic search strategy where the most promising node to be expanded is determined by the evaluation function. The two versions of BFS are A (Best-First Search) and Greedy Best-First Search. Greedy BFS utilizes the Heuristic function search and enables us to leverage both the algorithms [28]. Greedy BFS utilizes the heuristic to rank nodes within the search space and estimate their potential. It assumes that the most promising node at each iteration is the one that will reach the goal efficiently and is found to work very well for optimization problems [29]. In our research work, the best first search algorithm selects the best 13 attributes out of 33. The selected attributes are gender, age, med, reason, failures, higher, romantic, farmer, G2, and G3 as shown in table 2.

Table 2: Best Attributes Selected by BFS

S.No	Attribute	S.No	Attribute
1	Address	8	romantic: relation
2	Family Size	9	Go out –going out with friends
3	Mother Education	10	Heath: current health status
4	Travel Time	11	G1:1st period
5	Failures; past classes	12	G2:2nd period
6	Paid: Extra Classes	13	G3: final grade
7	Internet access at home		

CLASSIFICATION MODELS

The construction of a model that can be used to classify a group of items which will later be used to assign class labels or the attributes of yet unknown objects in the future is known as classification [30]. In the organized intelligent integrated model, decision Table (DT) and logistic regression (LR) were used for student academic performance.

Decision Table (DT) Classifier

Decision Table (DT) has a hierarchical tree structure with vertices (root, internal and leaves) and edges [31]. It splits an internal vertex by splitting criteria and repeats the process until all or most rows are classified accordingly. Gina Impurity is a method to split attributes. It is the probability of incorrectly classifying a sample, found by using the formula below, where pi represents the probability. The attribute with the lowest Gina impurity is chosen as the splitting attribute [32].

$$\text{GiniImpurity} = 1 - \sum i\,(pi)^2 \quad \text{equ. 1}$$

The role of the Decision table algorithm in analyzing student academic performance is to create a simple, visual flowchart to predict if a student will succeed or fail, and to show why. The two main roles, it uses student data (like attendance, past grades, and study habits) to predict a future outcome (like "Pass" or "Fail"). Secondly it automatically ranks the factors that matter most, placing the strongest predictor (like "Previous GPA" or "Attendance Rate") at the top of the chart. Also helps educators instantly know what is driving a student's performance, allowing them to provide targeted help sooner (e.g., if attendance is the main problem, they focus on that).

Logistic Regression (LR) Classifier

Logistic Regression (LR) is a statistical model that estimates the probability of an event occurring based on independent variables [33]. It first calculates the beta parameters, or coefficients based on the maximum likelihood estimation (MLE) of all the independent variables, shown in the equation below, and after applying mathematical operations (logging and summing), a predictive probability is generated for each input [34].

$$\ln(\pi/(1-\pi)) = \beta 0 + \beta 1 * X1 + \ldots \beta k * Kk \quad \text{equ. 2}$$

The role of the Logistic Regression algorithm in predicting student academic performance is to calculate the probability (the chances) of a student achieving one of two outcomes (a pass/fail result) [35] and to clearly show which factors are most important in that prediction like,

- It calculates the chance of success: Instead of just guessing "Pass," the model gives you a percentage, like "There is an 80% chance this student will pass."
- Common Use: This is perfect for predicting things like:
- Pass vs. Fail in a course.
- Graduate vs. Dropout from a program.
- Factor Importance: The model assigns a "weight" to every piece of student data (attendance, homework grade, test score). A bigger weight means that factor is more important.
- Actionable Insight: Educators can look at the model and know exactly how much a factor matters: "For every class missed, the student's chance of failing goes down by 5%." This allows schools to target the specific factors that are the biggest risks [36].

RESULTS AND DISCUSSION

This section presents a comprehensive analysis of the experimental results obtained during this research, along with a detailed discussion of the findings about the research objectives. The empirical results demonstrate that individual classifiers offer distinct advantages for effectively detecting student performance evaluation and prediction. The performance of two classification algorithms is evaluated using the best and optimal feature selection strategies. The effectiveness of a classification algorithm can be determined by assessing its correlation coefficient, root mean squared errors, and mean absolute errors, particularly when using the best attribute assessment technique [37]. This evaluation is conducted on both the training and test data sets. To make it easier to understand for the performance of the presented model, the root mean square error (RMSE) is used as the measurement for the accuracy of the model in the experiment. Thus, if the value of RMSE is smaller, the value is nearer to true value and hence the model is more accurate.

Table 3: COMPARISON OF DECISION TABLE AND RANDOM FOREST CLASSIFIERS

Algorithm	Correlation Coefficient	Root Mean Squared Error	Mean Absolute Error
Decision Table	0.8966	1.8261	1.1969
Logistic Regression	0.916	1.6223	1.0354

Table 3. Shows the summarized results of the decision table and the Logistic Regression algorithm classifiers with attribute evaluation techniques. The classifiers are compared based on correlation coefficient, root mean squared error and mean absolute error. The comparative analysis indicated that the Logistic regression model performed better, demonstrating higher accuracy and a lower error (RMSE) as concluded by the study. This means the Logistic Regression provided more correct predictions that were closer to the actual student grades.

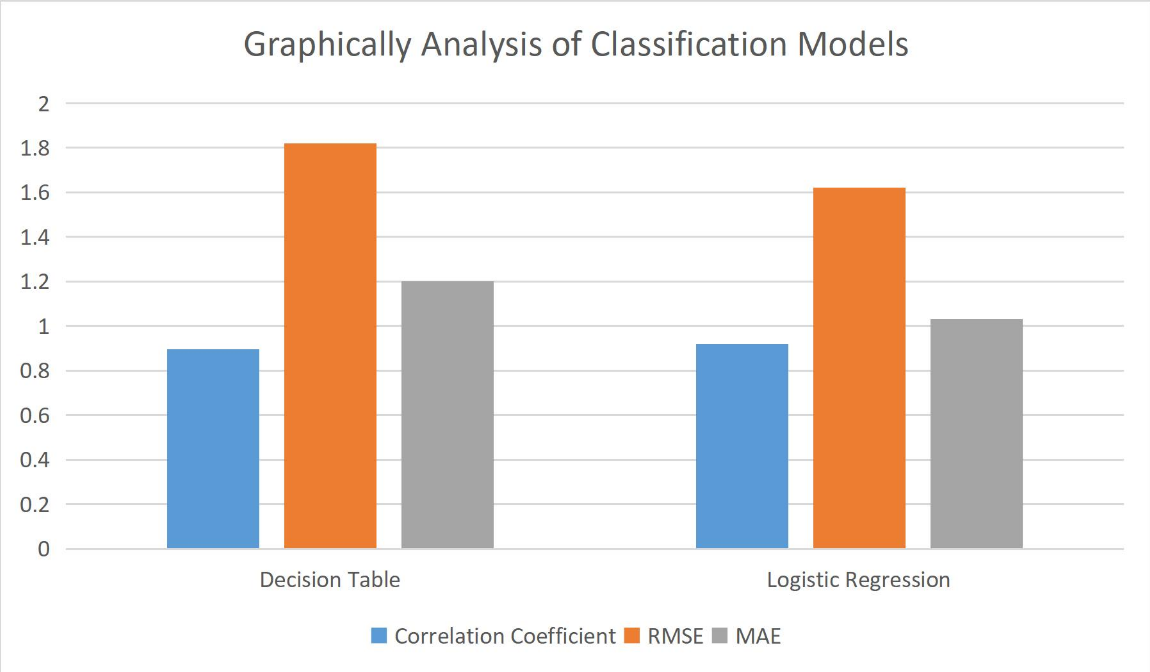

Fig 2: Graphically Comparison of Classification Models

Figure 2 presents a graphical comparison of the two models in terms of correlation coefficient, mean squared error, and mean absolute error, which shows that the LR outperforms in terms of RMS error.

Future Scope: Data mining is used for classifying large datasets, allowing various assumptions to be made. In the classification process, attributes are responsible for generating rules. Our proposed solution offers an effective malware detection system with enhanced accuracy and reduced execution time, leveraging its parallel processing capabilities, feature normalization, and selection methods. In further research, selection of attributes has been enhanced based on different artificial intelligence approaches that can be used as fitness factors for evaluation of attributes.

REFERENCES

[1] Rahman, M. A. U., Roman, M., Ahmad, S., Jan, M. S., & Nawab, H. U. (2021). Optimizing Collaboration: Insights into Inter-Team Coordination and Self-Management in Distributed Agile Software Development. *Webology (ISSN: 1735-188X)*, *18*(5).

[2] Roman, M., Ullah, A., Ullah, M. A., Hussain, F., Shams, S. S., Bint-e-Meraj, A., & Ali, S. (2025). Predicting Academic Success: A Machine Learning Approach Using Decision Tables and Random Forests Algorithms. *Spectrum of Engineering Sciences*, *3*(5), 205-213.

[3] Ullah, M. A., Ullah, A., Roman, M., Anwar, M., Siddiqi, M. U. M., Jaffar, M. H., & Ali, J. (2025). A Comparative Study of Machine Learning Algorithms for Cardiovascular Risk Prediction: Support Vector Machine, Gradient Boosting, And Rotation Forest. *Spectrum of Engineering Sciences*, 481-491.

[4] Abdi, A., & Zandipayam, A. (2019). Prediction of academic performance based on dimensions of academic identity and flourishing among students of the University of Medical Sciences. *Journal of Medical Education Development*, *12*(35), 19-26.

[5] Yaghoobi, A., & Bita, M. (2024). Meta-Analysis of studies on the relationship between identity styles and academic performance. *Quarterly of Educational Measurement*, *15*(55), 116-144.

[6] Bordbar, S., Ahmadinejad, P., Bahmaei, J., & Yusefi, A. R. (2024). The impact of mindfulness on academic achievement of students with the mediating role of adaptability: a structural equation modeling approach. *BMC Medical Education*, *24*(1), 1167.

[7] Tadese, M., Yeshaneh, A., & Mulu, G. B. (2022). Determinants of good academic performance among university students in Ethiopia: a cross-sectional study. *BMC medical education*, *22*(1), 395.

[8] Ghadampour, E., Farhadi, A., & Naghibeiranvand, F. (2016). The relationship among academic burnout, academic engagement and performance of students of Lorestan University of Medical Sciences. *Research in Medical Education*, *8*(2), 60-68.

[9] Ab Razak, W. M. W., Baharom, S. A. S., Abdullah, Z., Hamdan, H., Abd Aziz, N. U., & Anuar, A. I. M. (2019). Academic performance of university students: A case in a higher learning institution. *KnE Social Sciences*, 1294-1304.

[10] FATHI, A. M., Hosseinchari, M., Ahmadi, F., & Jowkar, B. (2021). Lived experiences of dropout university students: A qualitative content analysis.
[11] Yigermal, M. E. (2017). Determinant of Academic Performance of Under Graduate Students: In the Cause of Arba Minch University Chamo Campus. *Journal of Education and Practice*, *8*(10), 155-166.
[12] Moghadasi, J., & Keikavoosi-Arani, L. (2023). Investigating the factors influencing students' academic enthusiasm for a shift of paradigm among education managers shaping academic pedagogy. *BMC Medical Education*, *23*(1), 480.
[13] Anwar, M., Rahman, T., & Roman, M. (2025). VOICE-ACTIVATED SMART ENVIRONMENTS: DEEP LEARNING APPROACH FOR PASHTO SPEECH COMMAND PROCESSING. *Spectrum of Engineering Sciences*, *3*(5), 541-550.
[14] Singh, T. M., Reddy, C. K. K., Murthy, B. R., Nag, A., & Doss, S. (2025). Ai and education: Bridging the gap to personalized, efficient, and accessible learning. In *Internet of behavior-based computational intelligence for smart education systems* (pp. 131-160). IGI Global.
[15] Alghamdi, S., Soh, B., & Li, A. (2025). A Novel Optimised Feature Selection Method for In-Session Dropout Prediction Using Hybrid Meta-Heuristics and Multi-Level Stacked Ensemble Learning. *Electronics*, *14*(18), 3703.
[16] Yusoff, S. M., Othman, J., Mydin, A. M., Mohamad, W. W., Johan, E. J., & Yusoff, A. M. (2024). Students' Academic Performance: Prediction using Machine Learning Approaches. *Int. J. Acad. Res. Progress. Educ. Dev*, *13*(3), 1778-1792.
[17] Pan, J., Zhao, Z., & Han, D. (2025). Academic Performance Prediction Using Machine Learning Approaches: A Survey. *IEEE Transactions on Learning Technologies*.
[18] Ahmed, E. (2024). Student performance prediction using machine learning algorithms. *Applied computational intelligence and soft computing*, *2024*(1), 4067721.
[19] Sandeepa, A. G. R., & Mohottala, S. (2025, February). Evaluation of Machine Learning Models in Student Academic Performance Prediction. In *2025 5th International Conference on Advanced Research in Computing (ICARC)* (pp. 1-6). IEEE.
[20] Kalita, E., Alfarwan, A. M., El Aouifi, H., Kukkar, A., Hussain, S., Ali, T., & Gaftandzhieva, S. (2025, June). Predicting student academic performance using Bi-LSTM: a deep learning framework with SHAP-based interpretability and statistical validation. In *Frontiers in Education* (Vol. 10, p. 1581247). Frontiers Media SA.
[21] Agyemang, E. F., Mensah, J. A., Ampomah, O. A., Agyekum, L., Akuoko-Frimpong, J., Quansah, A., & Akinlosotu, O. M. (2024). Predicting Students' Academic Performance Via Machine Learning Algorithms: An Empirical Review and Practical Application. *Computer Engineering and Intelligent Systems*, *15*(1), 86.
[22] Jović, J., Kisić, E., Milić, M. R., Domazet, D., & Chandra, K. (2022). Prediction of student academic performance using machine learning algorithms. In *CEUR Workshop Proceedings* (Vol. 3454, pp. 31-39).
[23] Yadav, N. R., & Deshmukh, S. S. (2023, May). Prediction of student performance using machine learning techniques: a review. In *International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022)* (pp. 735-741). Atlantis Press.
[24] Yang, Y., Zhao, S., An, S., Li, X., & Zhang, Y. (2025). Student academic performance prediction via hypergraph and TabNet. *Journal of Big Data*, *12*(1), 119.
[25] Taher Mazandarani, M., Zand, Z., Khodabandelou, M. H., Mozaffari, F., & Sohrabi, B. (2025). Predicting Student Academic Performance: A Machine Learning Approach and Feature Analysis. *Interdisciplinary Journal of Management Studies*, *18*(3), 425-440.
[26] Hussain, F., Shams, S. S., Roman, M., Anwar, M., Shah, F. A., Meraj, A. B. E., ... & Uddin, M. A. (2025). MACHINE LEARNING IN HEALTHCARE: PREDICTING CHRONIC KIDNEY DISEASE THROUGH FEATURE-DRIVEN HEURISTIC MODELS. *Frontier in Medical and Health Research*, 3(7), 318-328.
[27] Shah, F. A., Meraj, A. B. E., Roman, M., Anwar, M., Zaib, A., Shams, S. S., & Hussain, F. (2025). Enhancing Weather Forecasting Accuracy: A Machine Learning Approach Using Genetic Algorithm and Random Forest. *Global Research Journal of Natural Science and Technology*.
[28] Heusner, M. (2019). *Search behavior of greedy best-first search* (Doctoral dissertation, Philosophisch-Naturwissenschaftliche Fakultät der Universität Basel).
[29] Wilt, C., & Ruml, W. (2015). Building a heuristic for greedy search. In *Proceedings of the International Symposium on Combinatorial Search*. Vol. 6, No. 1, pp. 131-140.

[30] Roman, M., Nawab, H. U., Ahmad, S., & Khan, I. A. (2022). K-Nearest Neighbor And Fuzzy K-Nearest Neighbor Algorithm Performance Analysis For Heart Disease Classification. *Webology(ISSN: 1735-188X)*, *19*(1).

[31] Moret, B. M. (1982). Decision trees and diagrams. *ACM Computing Surveys (CSUR)*, *14*(4), 593-623.

[32] Ahmed, A., Nipa, F. A., Bhuyian, W. U., Mushfique, K. M., Shahin, K. I., Nguyen, H. H., & Farid, D. M. (2024). Students' performance prediction employing Decision Tree. *CTU Journal of Innovation and Sustainable Development*, *16*(Special issue: ISDS), 42-51.

[33] Nick, T. G., & Campbell, K. M. (2007). Logistic regression. *Topics in biostatistics*, 273-301.

[34] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.

[35] Ochieng, Geoffrey O. (2023). "A Regression Model to Predict the Risk of Incomplete Grading of Student Assessments in Higher Education Institutions." PhD diss., KCA University.

[36] Mohamed Bahy Eldin, N., Elsemsar, A., & Kamal Amin, R. (2025). Predicting Student Performance: A Machine Learning Approach to Forecasting Pass/Fail Outcomes. *منارة الاسكندرية للعلوم التجارية*, *1*(1), 188-178.

[37] Roman, M., Naz, I., Luqman, M. A., Ali, J., Jan, M. S., & Nawab, H. U. (2024). Stroke Disease Prediction Using K-Nearest Neighbor and Decision Tree Algorithms with Machine Learning Pre-Processing Techniques. *Migration Letters*, *21*(S4), 2015-2027.