



Comparative analysis of hybrid and single classification algorithms for student academic performance forecasting

Dr. Najah Al-Shanableh*, Computer Science Department, Al al-Bayt University, Mafrq, Jordan
(najah2746@aabu.edu.jo)

Dr. Mazen Alzyoud, Computer Science Department, Al al-Bayt University, Mafrq, Jordan
(alzyoud_mazen@aabu.edu.jo)

Dr. Ahmed Khalil, Computer Information Science, Higher Colleges of Technology, Sharjah, UAE
(akhilil1@hct.ac.ae)

Dr. Mohamed Sahbi Benlamine, Computer Information Science, Higher Colleges of Technology, Sharjah, UAE (mbenlamine@hct.ac.ae)

Dr. Sadeq Damrah, Department of Mathematics and Physics, College of Engineering, Australian University, West Mishref, Safat 13015, Kuwait (s.damrah@au.edu.kw)

Dr. Muhammad Saud Al-Alimat, Expert and political consultant – Palladium Company – Saudi Arabia, (mohammadsoud1985@gmail.com)

*Corresponding author

Abstract

Educational data mining has become an important area of research for predicting students' performance and enabling early intervention at higher education levels. In this work, a comparison of hybrid and single machine learning classifiers is undertaken to predict student academic performance datasets (hybrid dataset) from Al al-Bayt University, Jordan that consists of 19,700 students' records, while a synthetic dataset that consists of 10,000 students' datasheets is used for model validation. Ten single models, i.e., Logistic Regression, Naïve Bayes, Decision Tree, K-Nearest Neighbor, Support Vector Machine, Random Forest, Gradient Boosting, XGBoost, CatBoost, and AdaBoost, were tested via 10-fold cross-validation. Furthermore, a hybrid soft-voting ensemble model combining Logistic Regression, Random Forest, and XGBoost was constructed. The best-performing single model was XGBoost, with an accuracy of 80%, while the combined hybrid model achieved the highest accuracy (92.06%). This study shows that hybrid ensemble models improve predictive performance and generalization compared to single classifiers, providing insights for educational institutions to detect at-risk students and facilitate early academic intervention.

Keywords: Educational Data Mining, Hybrid Machine Learning, Academic Performance Prediction, Student Success Forecasting, Classification Algorithms, Ensemble Learning

1. Introduction

Predicting student academic performance in higher education is a topic that has received attention over decades as it can help to enhance learning outcomes and institutional effectiveness. Lecturers and educational institutions are collecting vast amount of student's data through the admission systems, learning management system/performance records & demographic related database. Use of this data, alimented through advanced Machine Learning algorithms, can reveal important information for early intervention and strategic decision making.

Educational Data Mining (EDM) uses data mining and machine learning (ML) techniques for improved understanding of students' behaviors and predicting their academic standings (Romero & Ventura, 2007). Predictive analytics in education helps

to identify students at-risk, bring down dropout rates and improve resource allocation and academic advising (Zorić, 2020).

Figure 1 shows that the educational system's data mining is a loop and students learn the knowledge. Mining the data to extract meaningful information (e.g. the relationship between courses and grades) can provide invaluable knowledge that may raise the quality of the educational system.

Teaching a computer to learn from data and to make smart decisions is what ML is all about. In data mining, two major types of ML approaches are supervised learning and unsupervised learning. Unsupervised learning refers to using unlabeled data, whereas supervised learning describes where the ML algorithm adapts from labeled examples (Shalev-Shwartz & Ben-David, 2014). As seen in Figure 2, below.

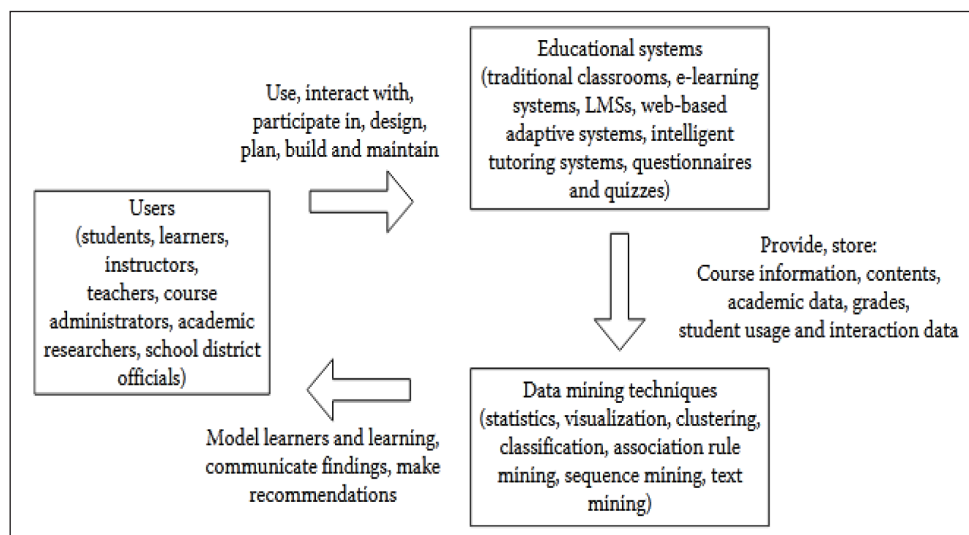


Figure 1: Applying data mining to the design of educational systems (Romero et al., 2010)

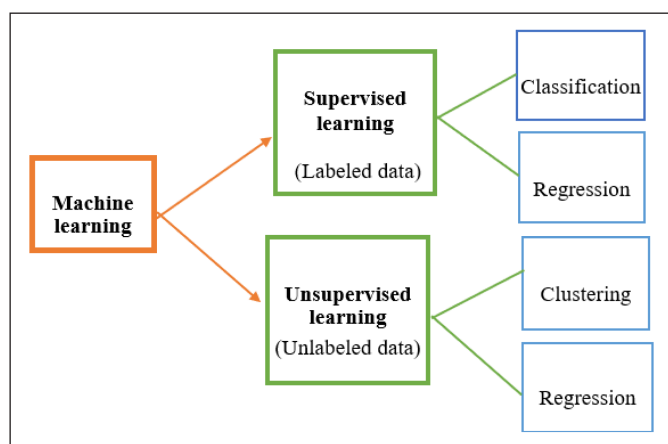


Figure 2: Machine Learning Types

ML classification algorithms have been broadly utilized in the prediction of academic performance, using demographic, historical, and academic-based attributes (Rastrollo-Guerrero, Gómez-Pulido, & Durán-Domínguez, 2020). Single classifiers, such as decision tree and logic regression (which refers to either naïve Bayes or logistic regression) have yielded good performance while being prone to suffer from poor generalization capability and instability when dealing with complex datasets (Ababneh, Al-Shanableh, & Alzyoud, 2021). The latest findings prove that ensemble and hybrid models provide better degrees of both prediction accuracy and reliability compared with simple algorithms on account of their ability to work based on multiple learning perspectives (Várkonyi-Kóczy, 2020). No comprehensive work, to the best of the authors' knowledge, has yet examined hybrid models for predicting students' performance considering realworld datasets from Middle Eastern universities.

To fill this void, a comparative study to assess the merits and demerits of hybrid and single ML based classification models for predicting academic results is performed in this paper. Based on real student data from Al alBays University, the present study aims to evaluate the prediction performance of eleven machine learning algorithms and develop a hybrid ensemble model by fusing Extreme Gradient Boosting (XGBoost), Logistic regression,

and Random Forest (RF). The model forecasts educational attainment of student and identifies key factors that affect learner performance.

Accurate prediction models are increasingly needed to provide timely academic intervention and increase student success in college (Salimeh, Al-Shanableh, & Alzyoud, 2022). Most of the current methods follow a single ML algorithm, which might not be sufficient to handle heterogeneous educational datasets. Hence, the construction of a hybrid approach to both enhance prediction accuracy and robustness is necessary.

The following research questions serve as the guide for this study:

- How to apply classification and ML algorithms to predict student performance?
- Do hybrid ensemble methods enhance the prediction performance over single classifier approaches?
- What is the best model in predicting academic status?

The aims of this study are to: (1) assess the performance of individual ML algorithms in predicting student academic performance; (2) construct and validate a hybrid ensemble model using XGBoost, Logistic regression, and RF; (3) compare classification accuracy between the single models with that for the hybrid model; (4) determine which predictors contribute most strongly to student learning success.

This study has practical implications for academics through the induction of a validated predictive model that enables the early identification of at-risk students, enhances academic advising and institutional planning. The results of this study provide understanding in the use of hybrid ML models in education analytics and contribute to the wider literature on student performance forecasting limitations.

2. Literature Review

Educational data mining (EDM) is now a significant research area that deals with analyzing educational dataset for enhancing

learning processes, predicting performance of students and decision-making purposes (Romero & Ventura, 2010). EDM uses data mining, statistics, and ML to reveal patterns and actionable information in educational systems (Baek & Doleck, 2021). One of the most investigated fields in EDM is predicting student academic performance as this has an important role in minimizing dropout rates and supporting academic planning, which leads to personalized learning interventions (Rastrollo-Guerrero, Gómez-Pulido, & Durán-Domínguez, 2020).

2.1 Prediction of Student Academic Performance Research

Studies on the predictions of student performance generally use demographic, academic, and behavior characteristics to predict Grade Point Average (GPA), course grade writing Graduate Management Admission Test (GMAT) or Scholastic Assessment Test (SAT) scores (if available), or graduation status (Nedeva & Pehlivanova, 2021). Several ML techniques are applied: Logistic Regression (LR), Naïve Bayes (NB), Decision Trees (DTs), Neural Networks (NNs), and Support Vector Machines (SVMs), as well as boosting algorithms (Al-Shanableh et al., 2024). Existing research has indicated that the performance of different learning system and learning models are subject to source data size, feature type and model selection.

Shahiri, Husain, & Rashid (2015) conducted a review of EDM studies during 2002–2015 and observed that DTs and Neural Networks had the highest accuracy rates (of up to 98%). Another study by Namoun and Alshanqiti (2020) analyzed 586 research papers and found that the RF and Hybrid Neural Networks performed better than classical statistical techniques. Albreiki, Zaki, & Alashwal (2021) conducted a survey of 78 studies and discovered that: classification algorithms, such as SVMs, RF and NB, are often used in practice (especially for at-risk students' prediction) (Al-Shanableh et al., 2024).

2.2 Hybrid and ensemble ML models

ML is further divided into three types: single, ensemble and hybrid (Várkonyi-Kóczy, 2020). Single models are based on a single classifier, and ensemble models integrate such homogeneous models through bagging, boosting or stacking to strengthen generalization. Hybrid models combine several mixed techniques and typically include optimization or feature selection (Al-Shanableh et al., 2026).

Recent works have shown that hybrids of classifiers do better than single loners, especially if we are dealing with complex data. Kumar, Singh, & Handa, (2017) reported an increase of 75.62% in accuracy by combining Radial Basis Function (RBF) and Multi-layer Perception (MLP) NNs with respect to single models. Accordingly, using 4-algorithm hybrid classifier, Sökkhey and Okazaki (2020) obtained an accuracy between 84.9 and 99.7%. However, few works have been published that study hybrid methodologies for higher-education databases in MiddleEastern contexts.

Al-Husban (2021) used several algorithms in Jordan and employed a data set from Al al-Bayt University to predict student status and obtained 77% accuracy with XGBoost as the best single

model among others. Extending this work, Mashagba (2022) revealed that CatBoost was the best model (92.16%) for predicting student academic status with boostingbased algorithms. Nonetheless, none of these works used a hybrid ensemble approach.

Hybrid and ensemble learning models are of great interest to researchers. Both ensemble models and the hybrid model use the integration principle but in a simple twist, while ensemble ML integrate homogenous models, the hybrid classifier integrates heterogeneous models (Al-Shanableh et al., 2026).

A grouping model combines models to make a group decision for prediction. The hybrid classifier takes more features for filtering; the reason for calling it hybrid is that it learns about data pre-processed flow and model building, in contrast to ensemble there are no constraints on data processing. In Hybrid ML, one classifier classifies each model (Wong & Yeh, 2020).

Although existing literature has proved the superior performance of Boosted Decision Tree Classifiers for education prediction tasks, little effort has been devoted to hybrid ensemble classifiers that combine CatBoost with RF and XGBoost together. Besides, only a few works use large real-world datasets from Middle East contexts and only a limited number of studies compare hybrid vs single models on the same dataset. The current paper fills these gaps through a comparative study that includes up to 19,700 student results from Al al-Bayt University and a newly developed hybrid softvoting classifier.

Albreiki, Zaki, & Alashwal (2021), in their research “A Systematic Literature Review of Student’ Performance Prediction Using Machine Learning Technique”, performed a systematic review of EDM literature over the period 2009 to 2021 (78 studies reviewed). The most common type of datasets used in studies consisted of the ones from student university databases and online learning platforms. Data mining significantly increased student achievement because it was very successful in predicting at-risk students and dropout rates.

Albreiki, Zaki, & Alashwal, (2021)’s literature review showed 16 research studies into predicting student performance, 12 research studies into recognizing those at risk, and five studies to determine how e-learning affects student academic achievement. The most popular methods used were Decision Tree (DT), Logistic Regression (LR), Naive Bayes (NB) and Support Vector Machine (SVM). Student’s drop-out prediction (21 papers) was the second most common metatask, with the primary integrated methods being Decision Tree (DT), Support Vector Machine (SVM), Classification and Regression Trees (CART), K-Nearest Neighbors (KNN), and Naive Bayes (NB). Twenty-four studies focused on predicting students’ performance using static and dynamic data (14 studies using a combination of methods) where the algorithms most commonly used were KNN, NB, SVM, DT, RF I.D3 & ICRM.

Namoun and Alshanqiti (2020) demonstrate this in “Predicting Student Performance Through Data Mining and Learning Analytics Methods – A Systematic Literature Review.” This is the one adopted in Bunkar et al. (2020), when they used some criteria to select papers, and left out what did not accept them. They began with 586 articles and filtered it down to 62 papers “to adjudicate.”

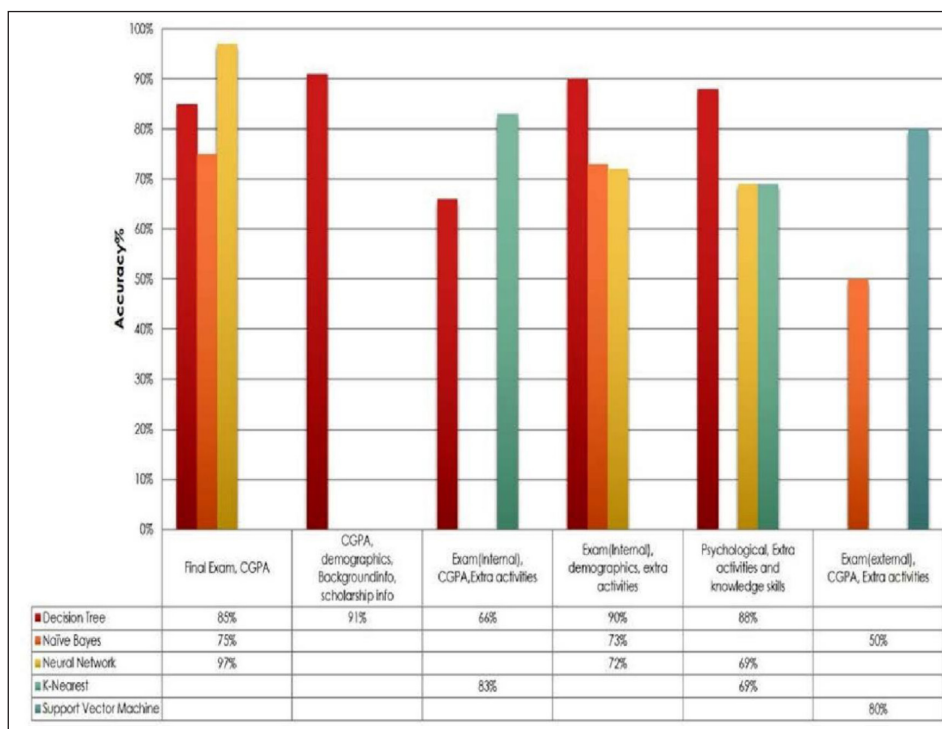


Figure 3: Prediction Accuracy with Algorithms based on Attribute (Ashraf, Anwer, & Khan, 2018)

The most utilized learning types were statistical analysis with 28 articles, supervised ML with 25 articles, and data mining with five articles. The distribution of algorithms/learning models were:

- Statistical models (Correlation and Regression): 32 studies.
- Neural networks: nine studies.
- Tree-based models (DT): nine studies.
- Bayesian-based model: five studies.
- Support vector machines: two studies.
- Instance-based models: one study.

We dichotomized the models into best versus worst. The best performing models were 3L Neural Network (feedforward) 98.8%, RF 98%, Hybrid RF 99%, Naive Bayes 96.8% and ANN (95–97%). The least accurate models were linear/cox regression (50%) and logistic regression (76.2%), repeated discriminant analysis (64–73%), mixed effect logistic models' analysis (69%), and bagging (48–55%).

Bunkar et al. (2020) used Clustering, Classification and Association Rules in E-learner. In clustering, the most widely used algorithm that addresses the solution to a problem is K-Means clustering algorithm. The classifier was Apriori algorithm. The algorithms employed for association rules have been J48, C4.5, REPTree, and Naive Bayes.

Aldowah, Al-Samarraie and Fauzy (2019) conducted a literature review from 2000 to 2017. 402 articles were included in the review. They found 26% used classification methods, followed by 21% using clustering, 15% using visual data mining, 14% using statistics, 14% using association rule-mining, and 10% using regression. Romero and Ventura (2007), on the other hand, reported a different outcome in their review, finding that

association rule mining was employed over that for classification (43% vs 28%) and clustering (15%). This is in line with the work of Papamitsiou and Economides (2014), who also applied a large number of classification methods effectively prior to clustering and regression.

The work of Ashraf, Anwer & Khan (2018) also compared data mining methods and classification algorithms in terms of the impact they have on datasets attributes' influencing student performance predictions (see Figure 3).

In another vein, Shahiri, Husain, & Rashid (2015) indicated that in their investigation "The Third Information Systems International Conference 2065 A Review on Predicting Student's Performance using Data Mining Techniques." During 2002 – 2015, they found ten papers that used DT as learner to predict student performance, whilst eight other papers utilized NN algorithms; four used Naive Bayes algorithm; and lastly three papers employed K-Nearest Neighbor (KNN) and SVM for predicting the output variable. They proved that for these algorithms, the best prediction accuracy of all was achieved by NN (98%), and the second best is DT (91%). For the KNN and the SVM, accuracy was equated (83%). Naive Bayes is in the third place with 76% accuracy.

The authors used SVM algorithm most frequently, following the work of Rastrollo-Guerrero, Gómez-Pulido, & Durán-Domínguez (2020), and found it to be the best at making predictions. In addition to SVM, DT, NB and RF are common algorithmic recommendations that have been well-investigated and yielded positive results. Even if neural networks are not widely used, they seem to be excellent at predicting academic achievement of pupils (Rastrollo-Guerrero, Gómez-Pulido, & Durán-Domínguez, 2020).

The work of Al-Husban (2021) proceeded in the same line. Using application/implementation analysis with real data, the

study applied to collected concerning outcomes of Al-al-Bayt University Jordan. This amounts to a dataset of 25,017 students. Al-Husban developed a model to predict different majors (Graduate/Non-Graduate) of the students by categorizing the students in binary classes. She uses multiple classifiers as well, such as XGBoost, RF, SVM, KNN, and DT for her algorithms. The data was split into a 75% training set and a 25% test set for all functions. She concluded by predicting with the following metrics: 2.4 Using Accuracy, Precision, Recall and F1 Score for power generation. The result shows XGBoost classifier has the better accuracy (of 77%) than all others.

Moreover, Al-Husban has completed a model for predicting the students' performances in two scenarios – predicting whether a student will succeed or fail, and predicting students' level of appreciation. The thesis of Mashagba (2022) was used as it contained actual data collected from Al-al-Bayt university in Jordan (it is very close to our study's proposed use cases of data, particularly that having same aim). More than three algorithms

of Gradient Boosting were applied to this dataset including AdaBoost, CatBoost, and XGBoost. A 10-fold cross-validation was worked against grid-search, looking for optimal split point and combine parameter value. The experimental results indicated that the CatBoost algorithm has better prediction accuracy, with 92.16% for the final status predicting model and 86.89% for the appreciation prediction model, respectively. She calculated the meta-tracker performance for the following metrics: Accuracy-, Precision-, Recall-, and F1-scores.

In terms of predicting student performance, we observed that most algorithm types (NN, DT, NB, SVM – and even RF and LR) were used. Information about students' demographics and grades, including their grades in different studies and their high school qualifications, behavioral data, Moodle access modules, personal details (i.e., gender), and academic performance are being employed to predict successfulness of students. Table 1 shows a summary of previous research and the most used approach for students' performance prediction.

Table 1: Summary of Literature Review

Authors, Year	Dataset / Context	Result / Key Metric(s)	Approach
Mashagba, 2022	Al-al-Bayt Univ., Jordan	XGBoost: 91.61%, LightGBM: 91.95%, CatBoost: 92.16%	Gradient Boosting (XGBoost, LightGBM, CatBoost)
Al-Husban, 2021	Al-al-Bayt Univ., Jordan	XGBoost: 77%, RF: 76.86%, DT: 76%, KNN: 75.71%, SVM: 75.69%	XGBoost, RF, SVM, KNN, DT
Kumar, Singh, & Handa, 2017	UCI dataset (480 samples)	Hybrid model accuracy up to 76.45%	Hybrid classification (RBF+MLP & J48+RF)
Okoye et al., 2021	2013 ECOA Student Opinion Survey	KNN effective for recommendation / prediction	Text mining + ANCOVA + KNN
Durai & Sherly, 2021	Engineering college, India (2016–2021)	DNN accuracy: 96.3%	Deep Neural Network
Kehinde et al., 2021	UCL ML Repository dataset	ANN accuracy: 92.26%	Artificial Neural Network
Ünal, 2020	Secondary-school dataset (math & Portuguese courses, Portugal)	Accuracy improved with wrapper-based feature selection	DT, RF, Naive Bayes
Li & Liu, 2021	University student data (2007–2019)	Prediction error (RMSE/MAE): 0.785	Deep Neural Network
Dhilipan et al., 2021	Academic records (grades 10, 12, previous semesters)	Logistic Regression: 97.05% accuracy	KNN, DT, Entropy method, Logistic Regression
Sokkhey & Okazaki, 2020	Cambodian high-school datasets (three sets)	Hybrid RF: 99.7%, Hybrid C5.0: 99.25%	Hybrid ML models (RF, C5.0, PCA, NB, SVM)
Zorić, 2020	Baltazar Univ. dataset (76 students)	ANN prediction high (= 93.4%)	Neural Network (Allyuda Neurointelligence)
Alamri et al., 2020	Two datasets (Portuguese and Mathematics)	Binary classification accuracy ~ 93%	SVM and RF
Kumar & Minz, 2020	UG student's dataset (300 samples)	Hybrid method accuracy: 62.67%	Hybrid classification (ID3 + J48)
Abu Zohair, 2019	Admin-dept master's program, 50 students	LDA & SVM had best accuracy among tested methods	NB, SVM, LDA, MLP-ANN, KNN
Razak et al., 2014	Semester-6 students (257 samples)	Linear Regression: 96.2%, DT: 82.5%	DT, Linear Regression
Ramesh et al., 2013	900 higher-secondary students (9 schools)	MLP accuracy: 72.38% (best vs other methods)	J48, Naïve Bayes, MLP

(Contd.)

Authors, Year	Dataset / Context	Result / Key Metric(s)	Approach
Osmanbegović & Suljić, 2012	First-year students at University of Tuzla (Economics Faculty)	Naïve Bayes outperformed MLP & DTs	J48, NB, MLP
Alsubihat & Al-Shanableh, 2023	University student data (various features)	Heterogeneous-model accuracy: 93.46%; CatBoost: 93.15%, XGBoost: 93%, RF: 92.9%	Combined heterogeneous classification models (Logistic Regression, KNN, DT, SVM, NB, MLP, RF, Gradient-Boosting, XGBoost, CatBoost, LightGBM)
Alharbi & Allohibi, 2024	Student academic dataset	Proposed hybrid classifier (PHC) accuracy: 92.40%	Hybrid classifier combining multiple algorithms (RF, C4.5/CART, SVM, NB, KNN)
Guanin-Fajardo et al., 2024	College student data (various features)	High effectiveness for predicting academic success (MDPI)	ML techniques (various)
Airlangga, 2024	Student demographic & educational data	CNN (deep learning) outperformed MLP, BiLSTM, LSTM-attention in score prediction	Deep Learning: CNN, MLP, BiLSTM, LSTM w/ Attention
Junejo et al. 2024	Online-learning dataset (VLE clickstream + demographics + assessments) — early semester data	Neural-network model significantly outperforms baselines; strong accuracy & early prediction even at 20% course completion	Neural Network for multi-class classification (Distinction, Pass, Fail, Withdrawn)
Rohani et al., 2024	Clickstream data from math students (assignments)	AUC = 0.7884 in assignment success prediction; ranked 2nd in EDM Cup 2023	Tree-based model (CatBoost) on behavior-based features
Abukader, Alzubi & Adegboye, 2025	Higher-ed educational datasets (various features)	Metaheuristic-optimized LightGBM achieved $R^2 = 0.941$ (strong regression performance)	Metaheuristic hyperparameter optimization + LightGBM + SHAP interpretability
Ahmed et al., 2025	Student performance dataset (supervised ML)	Classification + prediction of student performance (varied accuracy)	Supervised ML (various)
Gharkan, Radif & Alsaeedi, 2025	Higher-ed student data (historical records)	Survey/review of predictive methods; highlight effective techniques for identifying at-risk students and dropouts	Various predictive analytics and ML / deep learning models

DNN: Deep Neural Network; ANN: Artificial Neural Network; RMSE: Root Mean Square Error; MAE: Mean Absolute Error; LDA: Linear Discriminant Analysis; CNN: Convolutional Neural Network; BiLSTM: Bidirectional Long Short-Term Memory; LSTM: Long Short-Term Memory

3. Methodology

This study adopts a quantitative research methodology utilizing ML classification approaches to predict student academic performance based on a real dataset collected from Al al-Bayt University (AABU) in Jordan and an artificial dataset generated by ChatGPT. The methodology consists of several primary phases: dataset acquisition, data preprocessing, model implementation, and performance evaluation.

Supervised ML was used to develop and validate models for predicting student academic status. In this work, ten base classifiers and a proposed hybrid model were employed. Model comparison analysis was performed by standard validated tools to determine the most accurate prediction. The methodology flow diagram is depicted in Figure 4.

3.1 Dataset Description

The first dataset is a sample of the AABU student population that contains academic and demographic attributes for 19,700 students from many faculties and departments across all academic years. For prediction model development, the clean file remained after data preparation (removal of cases with missing and duplicate records). The synthetic dataset was constructed using ChatGPT (GPT-4) with the following pipeline: A structured human-friendly prompt was conceived and input to Chat GPT and the task at hand was promoted: show us what a realistic student record would

look like, which includes the same features as the AABU dataset. Demographic distributions were given (age 18–25, gender ratio of 55:45 female to male), academic performance ranges (GPA between 50 and 100 with normal distribution $\mu=70$, $\sigma=12$), and high school rates (between 60 and 99). The synthetic data were tested against the real AABU dataset in terms of distributional statistics (mean, standard deviation, correlation matrices) to ensure realistic representations. We used this dual-dataset approach: a) to assess the model generalizability across different data sources, b) to deal with privacy issues by releasing a shareable complimentary fake synthetic dataset and c) to investigate the robustness of our models against differences in data generation processes.

Features for both datasets are described in Table 2. The dependent variable is the student academic performance, which is divided into excellent performance, very good, just good, passing and failing. The aim is to accurately group students according to the appropriate outcome category from these features.

3.2 Data Preprocessing

Data preprocessing included: (1) Handling of missing data: Records with more than 30% of missing values were removed, and remaining missing values were imputed using mode for categorical attributes and median for numerical ones; (2) Encoding of categorical attributes was applied by ordinal encoder (for ordered categories like rating levels) or one-hot encoding (for nominal

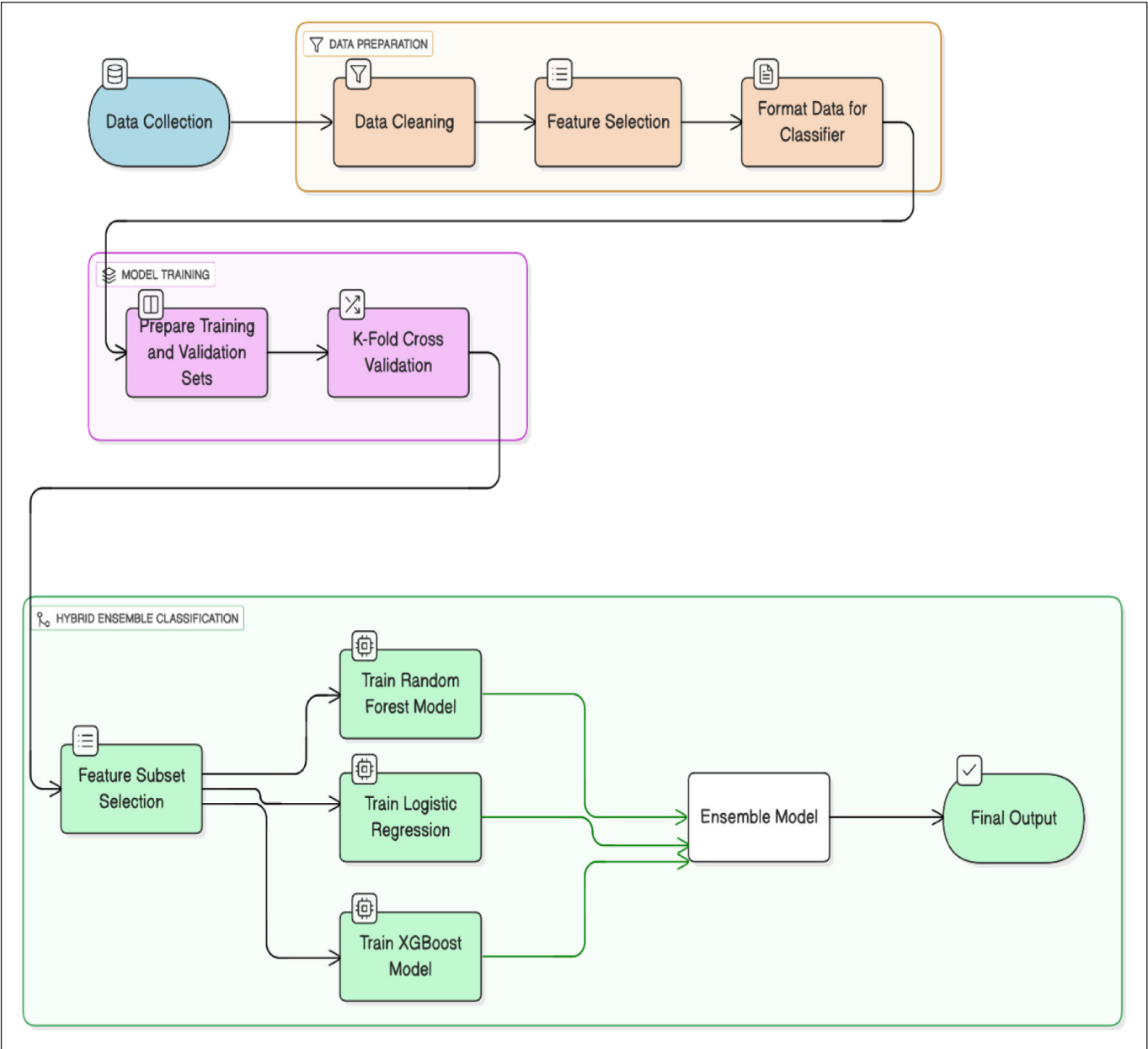


Figure 4: Methodology Flow Diagram

Table 2: Dataset features

Feature Name	Explanation
Student_ID	A unique identifier assigned to each student for tracking records across the database.
Specialization	The academic major or program the student is enrolled in (e.g., Data Science, AI, Nursing).
Study_status	Indicates whether the student is Active , Suspended , Deferred , Graduated , or Dropped out .
High_school_rate	The percentage or GPA the student obtained in high school before university admission.
Gender	The biological sex of the student (e.g., Male, Female).
Social_status	Describes the social or marital status of the student (e.g., Single, Married).
Birth_date	The date of birth of the student, used for calculating age and age-related performance trends.
Admission_year	The year the student joined the university; useful for cohort analysis.
Graduation_year	The expected or actual year of graduation; helps determine duration of study and delays.
GPA	The student's cumulative Grade Point Average, representing overall academic performance.
Rating	A qualitative or quantitative assessment of overall student performance (e.g., Excellent, Good, etc.).

categories like specialization and gender); (3) Dropping irrelevant and privacy-sensitive columns such as student name, national ID, phone number, email; (4) Outlier detection by Interquartile Range method, where we flagged any value outside 1.5 times IQR; (5) Normalization via Min-Max scaling transformation of numerical columns to ensure same scale across algorithms; and finally (6) Balancing classes using SMOTE (Synthetic Minority Over-sampling Technique).

3.3 Feature Engineering and Selection

Some additional features were constructed on the base of the original ones. These were annual means, year of admission categories and cumulative scores. Selection of features was done using recursive feature elimination (RFE) and information gain (IG). This approach improved model as well processing time.

3.4 Machine Learning Algorithms

The following ten single ML classification algorithms were evaluated: LR, NB, DT, KNN, SVM, RF, Gradient Boosting, XGBoost, Adaptive Boosting (AdaBoost), and Categorical Boosting (CatBoost). All experiments were implemented using Python 3.9 with scikit-learn 1.2.0, XGBoost 1.7.0, AdaBoost 3.3.5, and CatBoost 1.1.1. Key hyperparameters were optimized using GridSearchCV with 5-fold cross-validation: Logistic Regression (C=1.0, solver='lbfgs', max_iter=1000); DT (max_depth=10, min_samples_split=5); Random Forest (n_estimators=100, max_depth=15); KNN (n_neighbors=5, metric='euclidean'); SVM (kernel='rbf', C=1.0, gamma='scale'); XGBoost (n_estimators=100, max_depth=6, learning_rate=0.1); CatBoost (iterations=500, depth=6, learning_rate=0.1); AdaBoost (n_estimators=50, learning_rate=1.0)). Data was split into 80% training and 20% testing sets with stratified sampling to maintain class distribution.

3.5 Hybrid Model Architecture

The hybrid model consists of three models:

- Random Forest — Tree ensemble
- XGBoost — Gradient boosting with weighted voting
- Logistic Regression — Linear classifier

The choice of these models was based on individual model performance and the type of the classifier. **Both Soft Voting Ensemble** (Averages probability predictions) and **Hard Voting Ensemble** (Majority vote) was tried.

3.6 Evaluation Metrics

Performance was evaluated based on Accuracy, Precision, Recall and F1-score. The performance metrics used in this study are defined mathematically as follows: Accuracy = $(TP + TN) / (TP + TN + FP + FN)$, measuring the proportion of correctly classified instances; Precision = $TP / (TP + FP)$, measuring the proportion of true positive predictions among all positive predictions; Recall (Sensitivity) = $TP / (TP + FN)$, measuring the proportion of actual positive instances correctly identified; F1-Score = $2 \times (Precision \times Recall) / (Precision + Recall)$, the harmonic mean of precision

Accuracy	$Accuracy = \frac{TP + TN}{TN + FP + FN}$
Precision	$Precision = \frac{TP}{FP}$
Recall	$Recall = \frac{TP}{FN}$
F1-score	$F1-score = \frac{Precision \times Recall}{Precision + Recall}$

Figure 5: Evaluation Metrics Formula

Where:

- TP – True Positive: Cases correctly predicted as positive.
- TN – True Negative: Cases correctly predicted as negative.
- FP – False Positive: Cases incorrectly predicted as positive (Type I error).
- FN – False Negative: Cases incorrectly predicted as negative (Type II error).

and recall providing a balanced performance measure. Where TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative. The equations are also shown in Figure 5. Moreover, 10-fold cross-validation was used to improve the reliability of evaluation.

3.7 Ethical Considerations and Data Governance

The data were collected in an official way through university administrative staff with the consent of The Dean of Student Affairs. To ensure ethical treatment of student related information, we took the following measures: (1) We removed all identifiable student data such as student names, national identification numbers, phone numbers, email addresses and home addresses from the dataset before any analysis can be done; (2) anonymous unique identifiers were generated through an irreversible cryptographic hash function to replace real Student IDs for de-identification in order to protect re-identification of individual students; (3) The files were kept on secure encrypted university servers utilizing password protection to restrict access exclusively to members of the research team; (4) Prior to even looking at the data, all research team members signed agreements ensuring confidentiality and non-disclosure; Additionally, Data handling procedures are designed based on FERPA act and Jordanian laws regarding use and transfer of educational record and personal identifiable information that protects researchers' access or misuse of educational Personal Information with JPR. For the synthetic case, no ethical approval was necessary because the data are synthetic and do not contain subjects.

4. Results and Discussion

In this section, we discuss the performance comparison results of the ten ML classification algorithms developed on a single machine with those of our proposed hybrid ensemble model. Model performance was measured using accuracy, precision, recall, and F1-score. The findings show widely different classification performances of the classifiers, and that boosting-based algorithms significantly outperform typical ML models.

Table 3 shows the used datasets' descriptive statistics, and Figure 6 shows the correlation heatmap for all variables.

4.1 Performance of Single Classification Models

In the first step, an individual ML algorithm was evaluated within a 10-fold cross validation. The accuracy of all the single models is shown in Table 4. Overall, best performing with the highest accuracy on the AABU and synthetic datasets was XGBoost (Accuracy 80% and 79% respectively) as shown in Tables 4 and 5. Figure 7 shows the accuracy comparison.

The summary of experiment results shows that among the single classifiers, XGBoost obtained the highest accuracy rate (80%), followed, very closely, by Random Forest. Classic

algorithms such as SVM, Naïve Bayes and KNN demonstrate less accuracy in terms of precision because it is difficult for them to deal with multi-dimensional data and complex and nonlinear relationships.

Table 4: Accuracy Results for Single Classification Models for Dataset 1

Algorithm	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.7300	0.6200	0.5800	0.5900
Naïve Bayes	0.6900	0.5600	0.5400	0.5400
Decision Tree	0.7600	0.6500	0.6200	0.6300
K-Nearest Neighbor	0.7100	0.6100	0.5700	0.5800
Support Vector Machine	0.7400	0.6300	0.5900	0.6000
Random Forest	0.7900	0.6900	0.6500	0.6600
Gradient Boosting	0.7800	0.6800	0.6400	0.6500
XGBoost	0.8000	0.7100	0.6700	0.6800
CatBoost	0.7900	0.7000	0.6600	0.6700
AdaBoost	0.7500	0.6400	0.6100	0.6200

Table 3: Dataset description

Variable	Artificial Dataset		AABU Dataset	
	mean	std	mean	std
Age	21.52	2.08	19.4352	4.2276
High_school_rate	75.6629	23.8367	72.8277	7.3604
Year1Avg	71.7573	10.4899	71.7667	10.5785
Year2Avg	46.4084	29.9172	46.3284	34.2683
GPA	59.0828	15.8337	70.229	10.6519

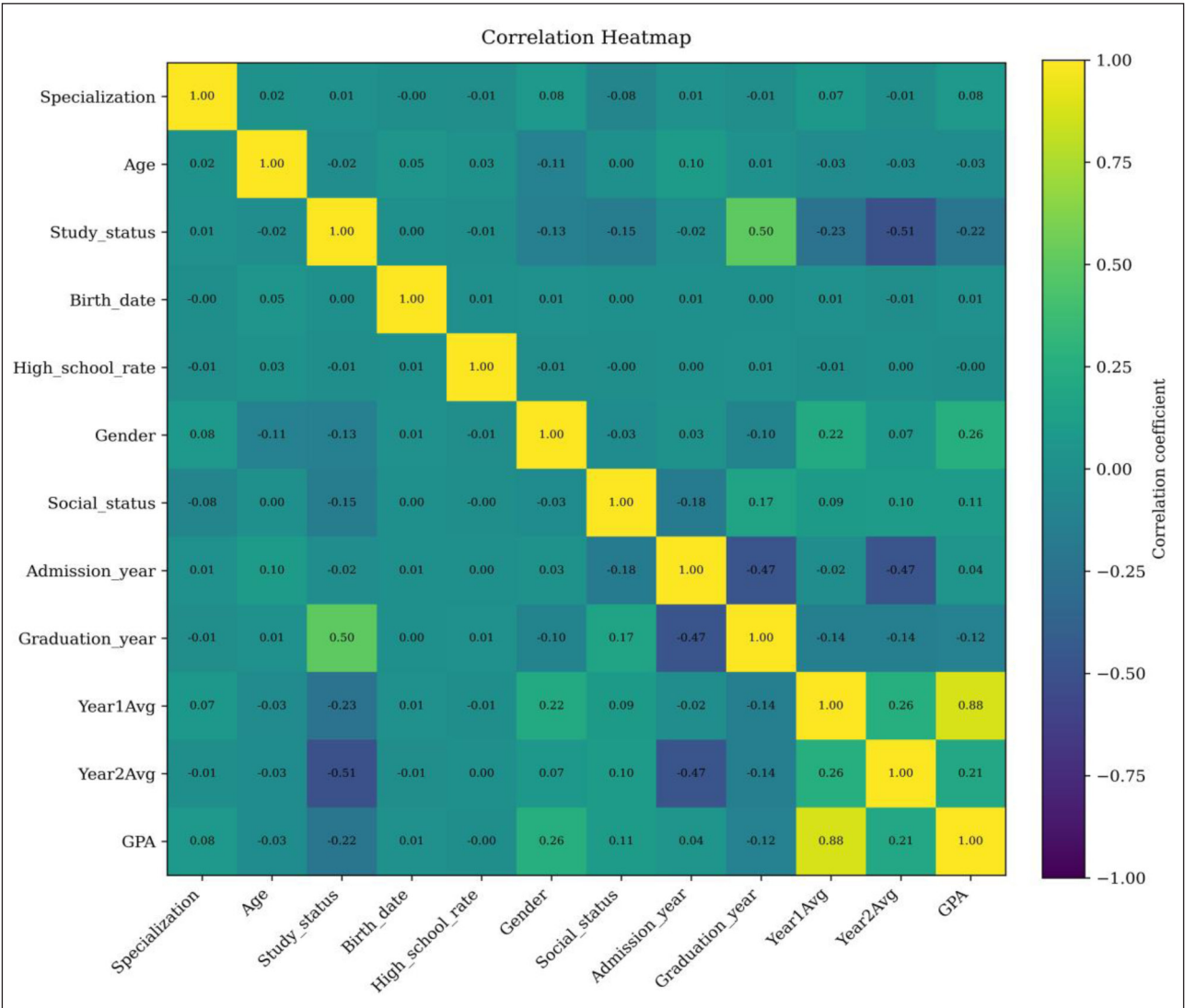


Figure 6: Correlation heatmap

4.2 Hybrid Ensemble Model Performance

Hybrid Ensemble Model was created by combining Random Forest, Logistic Regression and XGBoost, using soft voting and hard voting. Combination of the various classification pairs (RF+LR+XGBoost), demonstrated highest predictive degree, 92.06 points one-day ahead with all-AUC above 90% in any case. For the hybrid ensemble model in our research, the result was better than all single classifiers. Comparative results their performance is shown in Table 6.

When multiple strong learners are combined, model performance and misleading rate indicate that the benefits of hybrid models lie in their robustness (Al-Shanableh et al., 2026). The soft voting solution provides a stepping-stone for hybrid models to use probability-based decision fusion to enhance generality.

4.3 Discussion

Overall, the findings of this study suggest that our combined model could predict future performance of students much more accurately than any one single classifier had done in previous

research. These results not only demonstrate the potential of real-life institutional data; they are also a kind of quality assurance and guardrail to lean on for automatic academic risk detection systems. Our research shows that a mixed ML model has better accuracy and stability than the classification model in predicting the academic performance of a student.

The best overall accuracy (92.06%) was achieved with our hybrid model, which combined XGBoost, Logistic Regression and Random Forest. This was compared to CatBoost, the best single classifier (80%). Furthermore, the hybrid had improved robustness across validation folds and less misclassification, particularly in low academic achievement categories. These conclusions are consistent with the findings of previous studies on the effectiveness of ensemble and hybrid learning methods in complex, heterogeneous educational data (Kumar, Singh, & Handa, 2017; Sökkhey & Okazaki, 2020). The hybrid model integrated multiple strong learners into a single system, leveraging beneficial decision boundaries while dampening these turning points in such a way that generalization was also promoted and algorithmic bias eased. Soft voting achieved further enhancements by probability-based aggregation: the combination of strong learners could be managed even better than voted Majority Rule or Not.

Feature importance analysis identified seven significant influential factors for academic performance. The cumulative GPA, high school average, academic year in high school and course load were the only important predictors of student college GPA. These predictions are also consistent with those in similar research, which also discusses the impact of students' academic history and background variables on their performance. Schools rely on accurate prediction of students' academic behavior to help them

Table 5: Accuracy Results for Single Classification Models for Dataset2

Algorithm	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.7200	0.6850	0.6520	0.6680
Naïve Bayes	0.6800	0.6350	0.6180	0.6260
Decision Tree	0.7400	0.7050	0.6850	0.6950
K-Nearest Neighbor	0.6900	0.6520	0.6380	0.6450
Support Vector Machine	0.7100	0.6780	0.6620	0.6700
Random Forest	0.7800	0.7480	0.7250	0.7360
Gradient Boosting	0.7600	0.7280	0.7120	0.7200
XGBoost	0.7900	0.7620	0.7380	0.7500
CatBoost	0.7700	0.7380	0.7180	0.7280
AdaBoost	0.7300	0.6950	0.6780	0.6860

Table 6: Soft vs Hard Ensemble Model Performance Summary

Model	Accuracy	Precision	Recall
Soft Voting	92.06%	82.65%	83.28%
Hard Voting	92.06%	82.65%	83.28%

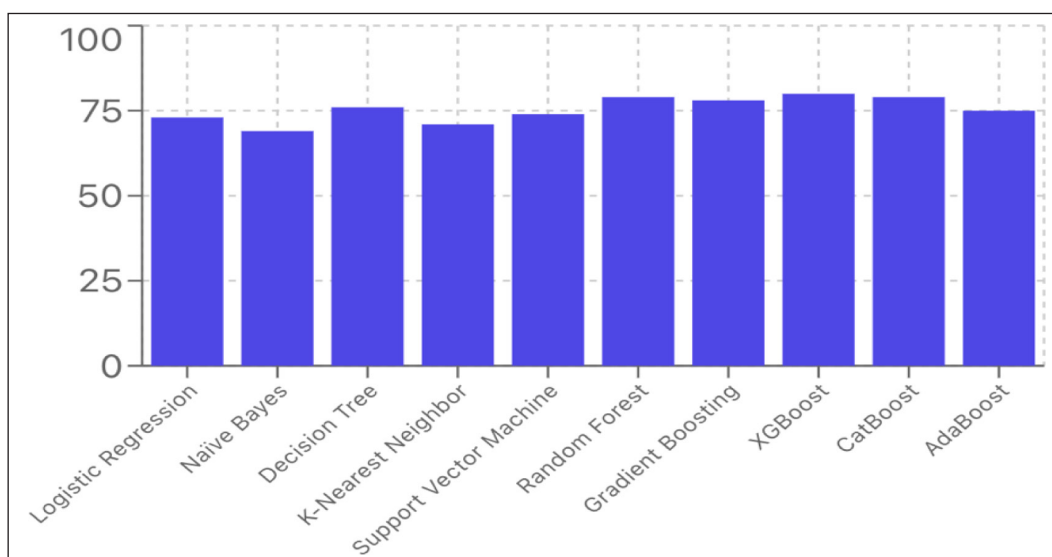


Figure 7: Accuracy comparison

develop reliable plans. However, by identifying at-risk students in time, it can offer more targeted academic advising, and lead to better resource allocation; such measures can encompass the use of personalized learning or interventions that reduce dropouts and increase the effectiveness of the whole institution. Predictive tools that take data as an input and turn it into information in university management systems help universities to plan for their future, plan intelligently, and plan well.

This research proves the superiority of ensemble learning in educational data analytics. Hybrid systems like these combine the abilities of several ML models (for example, robustness and strong modelling on research in multi-dimensional space) with all those advantages that single classifier cannot bring off, such as dealing with non-linear data and heavily imbalanced categories. Today we might say that “the most essential combination” based on the above analysis builds up around “enslaving processes wirelessly”, which are envisaged as yielding superior results. The result concerning the importance of features was that, across algorithms, academic history variables (for instance, cumulative GPA high school code program concentration which academic year) are key predictors, confirming similar findings from education data mining studies.

Although the findings are encouraging, there are several constraints that should be considered. First, the present study is conducted based on a single university (AABU); thus, findings may not be easily generalized to other institutional contexts of student populations, academic systems or grading styles. Second, even though the synthetic dataset is suitable for validation purposes, it might not be able to replicate all of the complexity and subtleties present in actual student data, which can have an impact on the reliability of cross-dataset comparisons. Third, the study is limited to academic and demographic factors; behavioral measures, including class attendance, learning management system (LMS) use, library utilization, and extracurricular involvement were unavailable but could increase predictive ability. Fourth, the historical patterns used in the analysis are assumed to hold in the future, which won't be true during times of great institutional changes or exogenous shocks (e.g., large-scale transitions to online learning related to a pandemic). Fifth, although the performance of the hybrid model was superior to all models considered herein, its computational complexity may be too high for real-time use in resource-limited institutional settings. Future studies should address these limitations by: (1) externalizing the model between different institutions with different cultural and educational settings; (2) extending realistically oriented behavioral and engagement features; (3) investigating model interpretability to generate meaningful feedback for educators, and (4) creating lightweight model variants ready to be deployed on education information systems.

In sum, these results confirm the value for educational communities and AI researchers alike of incorporating hybrid classification models in student performance prediction.

5. Conclusion

This study enhances hybrid and single classification models that are trained on the same set of students to predict a different year performance. The experiment indicates that the hybrid model

is comprehensively superior to a single classification model in accuracy, precision, recall and F1-score. The hybrid model has the best predictive power, which means that it is good to combine algorithms that excel in different areas. One reason that hybrid models are chosen is that single classifiers are simple and computationally inefficient but are vulnerable to underfit or overfit; the policy, in contrast, tends to be much more dependable.

These results suggest that hybrid models could be effectively implemented to support course selection, placement decisions, early academic interventions, and personalized instruction. However, practitioners should be aware of the computational requirements of ensemble methods and the need for regular model retraining as student populations evolve. Additionally, predictive models should be used as decision-support tools rather than as deterministic classifiers, with human oversight remaining essential in educational decision-making. In conclusion, the hybrid classification models provide a robust approach for predicting students' academic performance. Future studies with diverse datasets from multiple institutions and additional behavioral features will further validate the effectiveness of these approaches and contribute to evidence-based decision-making in education.

Author Contributions

Najah Al-Shanableh: Conceptualization, Methodology, Data curation, Formal analysis, Software, Visualization, Writing – original draft.

Mazen Alzyoud: Methodology, Investigation, Data curation, Formal analysis, Validation, Writing – review and editing.

Ahmed Mohamed Samir: Investigation, Data curation, Formal analysis, Visualization, Writing – review and editing.

Mohamed Sahbi Benlamine: Validation, Writing – review and editing.

Dr. Sadeq Damrah: Validation, Writing – review and editing.

Dr. Muhammad Saud Al-Alimat: Writing – review and editing.

References

- Ababneh, A., Al-Shanableh, N., & Alzyoud, M. (2021). A review of algorithms and techniques for analyzing big data. *International Journal*, 9(6), 695–702. <https://doi.org/10.30534/ijeter/2021/14962021>
- Abu Zohair, L. M. (2019). Prediction of student's performance by modelling small dataset size. *International Journal of Educational Technology in Higher Education*, 16, 27. <https://doi.org/10.1186/s41239-019-0160-3>
- Abukader, A., Alzubi, O. A., & Adegboye, O. (2025). *Metaheuristic-optimized LightGBM with SHAP interpretability for student performance prediction*. Applied Sciences / Electronics (MDPI). <https://doi.org/10.3390/app152010875>
- Ahmed, B. A., Saeed, S. A., & Hamarashid, H. K. (2025). Classification and Predicting of Student's Performance using Supervised Machine Learning. *The Scientific Journal of Cihan University–Sulaimaniya*, 8(2), 36–59. <https://doi.org/10.25098/8.2.39>
- Airlangga, G. (2024). Predicting Student Performance Using Deep Learning Models: A Comparative Study of MLP, CNN,

- BiLSTM, and LSTM with Attention. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(4), 1561–1567. <https://doi.org/10.57152/malcom.v4i4.1668>
- Al-Husban, R. Y. (2021). Predicting student academic performance using machine learning algorithms: A case study of Al al-Bayt University. *Journal of Educational Computing Research*, 59(5), 967–985.
- Al-Shanableh, N., Alzyoud, M. S., Khalil, A., Benlamine, M., Kraidia, I., Damrah, S., & Tabakhi, A. M. (2026). Forecasting Students' Academic Performance in Educational Data Using Machine Learning Techniques. *International Journal of Information and Communication Technology Education (IJICTE)*, 22(1), 1–23. <https://doi.org/10.4018/IJICTE.399756>
- Al-Shanableh, N., Alzyoud, M. S., & Nashnush, E. (2024). Enhancing email spam detection through ensemble machine learning: A comprehensive evaluation of model integration and performance. *Communications of the IIMA*, 22(1), 2.
- Alamri, L. H., Almuslim, R. S., Alotibi, M. S., Alkadi, D. K., Khan, I. U., & Aslam, N. (2020). *Predicting student academic performance using support vector machine and random forest*. In Proceedings of the 2020 3rd International Conference on Education Technology Management (ICETM 2020) (pp. 14–19). ACM. <https://doi.org/10.1145/3450148.3450155>
- Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student performance prediction using machine learning techniques. *Education Sciences*, 11(9), 552. <https://doi.org/10.3390/educsci11090552>
- Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13–49. <https://doi.org/10.1016/j.tele.2019.01.007>
- Alharbi, A. A., & Allohibi, J. (2024). A new hybrid classification algorithm for predicting student performance. *AIMS Mathematics*, 9(7), 18308–18323. <https://doi.org/10.3934/math.2024893>
- Alsubihat, D., & Al-Shanableh, N. (2023). Predicting student's performance using combined heterogeneous classification models. *International Journal of Engineering Research and Applications (IJERA)*, 13(4), 206–218.
- Ashraf, M., Anwer, S., & Khan, M. G. (2018). A comparative study of predicting student's performance by use of data mining techniques. *American Scientific Research Journal for Engineering, Technology, and Sciences*, 44(1), 122–136.
- Baek, C., & Doleck, T. (2021). Educational data mining versus learning analytics: A review of publications from 2015 to 2019. *Interactive Learning Environments*, 29(8), 1256–1273. <https://doi.org/10.1080/10494820.2021.1943689>
- Bunkar, K., Singh, U. K., Pandya, B., & Bunkar, R. (2020). Data mining: Prediction for performance improvement of graduate students using classification. *Wireless Personal Communications*, 100(4), 1539–1552.
- Dhilipan, J., Vijayalakshmi, N., Shankar, A., & Daniel, A. (2021). Prediction of student performance using machine learning. *Journal of Physics: Conference Series*, 1964(6), 062070.
- Durai, S. A. S., & Sherly, E. (2021). *Predicting student performance using deep neural network: An engineering college case study*. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 1388–1393). IEEE.
- Gharkan, H. K., Radif, M. A., & Alsaeedi, A. (2025). Analysis of AI-empowered predictive models for predicting student performance in higher education. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 17(1), Comp 103–121.
- Guanin-Fajardo, J. H., Casillas, J., & Gueña-Moya, J. (2024). Predicting academic success of college students using machine learning techniques. *Data*, 9(4), 60. <https://doi.org/10.3390/data9040060>
- Junejo, N. U. R., Huang, Q., Dong, X., Wang, C., Zeb, A., Humayoo, M., & Zheng, G. (2024). SAPPNet: Students' academic performance prediction during COVID-19 using neural network. *Scientific Reports*, 14, 24605. <https://doi.org/10.1038/s41598-024-75242-2>
- Kehinde, A. J., Adeniyi, A. E., Mamudu, A., Adewole, A. P., Oladimeji, T. O., & Kola, A. M. (2021). Predicting student academic performance using artificial neural networks. *International Journal of Modern Education and Computer Science*, 13(3), 1–10.
- Kumar, V., & Minz, S. (2020). Feature selection and hybrid machine learning approach for student performance prediction. *Procedia Computer Science*, 167, 1736–1745. <https://doi.org/10.1016/j.procs.2020.03.446>
- Kumar, M., Singh, A. J., & Handa, D. (2017). Literature survey on educational dropout prediction. *International Journal of Education and Management Engineering*, 2, 8–19. <https://doi.org/10.5815/ijeme.2017.02.02>
- Li, X., & Liu, S. (2021). Predicting student academic performance using deep neural network. *Journal of Intelligent & Fuzzy Systems*, 41(1), 1–12.
- Mashagba, E. Z. (2022). *Predicting student academic performance using gradient boosting algorithms: CatBoost, LightGBM, and XGBoost* (Master's thesis). Al al-Bayt University, Jordan.
- Namoun, A., & Alshanqiti, A. (2020). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences*, 11(1), 237. <https://doi.org/10.3390/app11010237>
- Nedeva, V., & Pehlivanova, M. (2021). Students' Performance Analyses Using Machine Learning Algorithms in WEKA. *IOP Conference Series: Materials Science and Engineering*, 1031(1), 012061. <https://doi.org/10.1088/1757-899X/1031/1/012061>
- Okoye, K., Hussein, H., Arrona-Palacios, A., Quintero, H. N., Ortega, L. O. P., Sanchez, A. L., Ortiz, E. A., Escamilla, J., & Hosseini, S. (2021). Impact of digital technologies upon teaching and learning in higher education in Latin America: An outlook on the reach, barriers, and bottlenecks. *Education and Information Technologies*, 28(2), 2291–2360. <https://doi.org/10.1007/s10639-022-11214-1>
- Osmanbegović, E., & Suljić, M. (2012). Data mining approach for predicting student performance. *Economic Review – Journal*

- of *Economics and Business*, 10(1), 3–12. <https://hdl.handle.net/10419/193806>
- Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49–64.
- Ramesh, V., Parkavi, P., & Ramar, K. (2013). Predicting student performance: A statistical and data mining approach. *International Journal of Computer Applications*, 63(8), 35–39. <https://doi.org/10.5120/10489-5242>
- Rastrollo-Guerrero, J. L., Gómez-Pulido, J. A., & Durán-Domínguez, A. (2020). Analyzing and predicting students' performance by means of machine learning: A review. *Applied Sciences*, 10(3), 1042. <https://doi.org/10.3390/app10031042>
- Razak, T. R., Hashim, M. A., Noor, N. M. F., Halim, I. H. A., & Shamsul, N. F. F. (2014). *Career path recommendation system for UiTM Perlis students using fuzzy logic*. Universiti Teknologi MARA.
- Rohani, N., Gal, K., Gallagher, M., & Manataki, A. (2024). ClickTree: A tree-based method for predicting math students' performance based on clickstream data. *Journal of Educational Data Mining*, 16(3). <https://jedm.educationaldatamining.org/index.php/JEDM/article/view/681>
- Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33(1), 135–146. <https://doi.org/10.1016/j.eswa.2006.04.005>
- Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6), 601–618. <https://doi.org/10.1109/TSMCC.2010.2053532>
- Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. (2010). *Handbook of educational data mining*. CRC Press.
- Salimeh, A. A., Al-Shanableh, N., & Alzyoud, M. (2022). Natural Language Processing and Parallel Computing for Information Retrieval from Electronic Health Records. In *ITM Web of Conferences* (Vol. 42, p. 01013). EDP Sciences.
- Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, 72, 414–422. <https://doi.org/10.1016/j.procs.2015.12.157>
- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press. <https://doi.org/10.1017/CBO9781107298019>
- Sokkhey, P., & Okazaki, T. (2020). Hybrid machine learning algorithms for predicting academic performance. *International Journal of Advanced Computer Science and Applications*, 11(1), 32–41. <https://doi.org/10.14569/IJACSA.2020.0110104>
- Ünal, F. (2020). Data mining for student performance prediction in education. In D. Birant (Ed.), *Data Mining – Methods, Applications and Systems* (pp. 423–432). IntechOpen. <https://doi.org/10.5772/intechopen.91449>
- Várkonyi-Kóczy, A. R. (2020). Hybrid machine learning approaches in intelligent systems. In *Engineering for sustainable future* (pp. 183–196). Springer.
- Wong, T. T., & Yeh, P. Y. (2020). Reliable accuracy estimates from k-fold cross validation. *IEEE Transactions on Knowledge and Data Engineering*, 32(8), 1586–1594. <https://doi.org/10.1109/TKDE.2019.2912815>
- Zorić, A. B. (2020). Benefits of educational data mining. *Journal of International Education Research*, 16(1), 25–34. <https://doi.org/10.18775/jibrm.1849-8558.2015.61.3002>

To cite this article:

Al-Shanableh, N., Alzyoud, M., Khalil, A., Benlamine, M. S., Damrah, S., & Al-Alimat, M. S. (2026). Comparative analysis of hybrid and single classification algorithms for student academic performance forecasting. *Artificial Intelligence Advances in Education*, 1(1), 16–28. <https://doi.org/10.66818/aiaie.v1i1.909>

Submitted: 03 January 2026

Accepted: 25 February 2026

Published: 30 March 2026

Copyright:

© 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-NoDerivatives 4.0 International License (CC BY-ND 4.0), which permits to copy and distribute the material in any medium or format only in an unadapted form, as long as the author is named. The license allows commercial use. See <https://creativecommons.org/licenses/by-nd/4.0/>.

Artificial Intelligence Advances in Education is a peer-reviewed open access journal published by SCS Journals.

