MACHINE LEARNING-BASED STUDENT PERFORMANCE PREDICTION IN HIGHER EDUCATION

Mohandass K[1], Aurangjeb Khan[2]
[1,2]SSCS, CMR University, Bangalore, India.
[1]E-Mail: mohan2004dass@gmail.com , [2]E-Mail: aurangjeb.khan83@gmail.com

Abstract : Forecasting academic achievement plays a vital role in enhancing learning results and facilitating proactive support mechanisms to minimize student attrition in higher education. Conventional approaches dependent on historical scores and instructor judgement frequently prove inadequate and are unable to detect academically vulnerable learners in time to deliver effective assistance. This paper aims to assess contemporary ML-driven approaches within the domain of learning analytics (EDM) to address these limitations. Through a comprehensive literature review of 12 recent studies (2019-2024), we examine the application of various ML algorithms — such as decision trees, Random Forest, support vector machines, and Artificial Intelligence (XAI) frameworks, based on identified research gaps, including the frequent lack of model interpretability and reliance on For single-institution datasets, we propose a robust methodology utilizing the UCI student performance dataset. The proposed framework compares decision trees, Random Forest, and naive Bayes classifiers to balance accuracy and transparency. The literature confirms that aggregated classifiers, especially the random forest algorithm, repeatedly demonstrate superior predictive performance (often exceeding 95%), making them highly effective for early identification of at-risk students and empowering evidence-based administrative decisions across academic institutions.

Keywords: Machine Learning; Educational Data Mining; Explainable AI; Classification Algorithms; At-Risk Student Detection.

2. INTRODUCTION

Scholastic achievement reflects a wide range of cognitive and behavioral engagement by learners, capturing their competency levels and accomplishments over a given study period. Prompt and reliable forecasting of academic outcomes holds considerable significance in higher education. It allows educational institutions to pinpoint students facing challenges, deploy customised academic support strategies, and continuously refine teaching approaches, thereby curbing withdrawal rates that burden institutional funding and compromise students' success future.

Traditionally, educators and administrators have attempted to assess the risk of academic failure. Failure was based primarily on manual records, previous homework, quizzes, and midterm exams. However, this approach reveals significant limitations of traditional education systems. Dependence on instructors' personal judgement and sparse historical data leads to a fundamentally passive approach, unreliable and operationally inadequate. Furthermore, legacy frameworks are generally ill-equipped to handle intricate, multifaceted learner information (such as demographics, socioeconomic status, and digital learning behaviours), consequently delaying the recognition of academically vulnerable individuals.

Machine learning (ML) and educational data mining (EDM) offer powerful solutions to these traditional shortcomings. Through the utilisation of extensive longitudinal and behavioural datasets, ML-based systems are capable of autonomously revealing latent trends, modeling complex nonlinear relationships, and anticipating a learner's terminal academic standing well ahead of course completion.

3. LITERATURE REVIEW

The following section offers a comprehensive synthesis of twelve contemporary investigations on machine learning-based student performance prediction in higher education. The reviewed works span the period 2020-2025 and encompass a diverse array of classification methods, data sources, and performance metrics spanning aggregated classifiers, neural architectures, and interpretability-focused methodologies.

Pan et al. (2025) investigated state-of-the-art approaches (decision trees, Random Forest, SVM, Naive Bayes) using datasets like OULAD and CUD. They found that combining multiple methods yield robust predictions. Gap: The absence of standardised performance benchmarks impedes unbiased cross-study comparisons [1].

Alwarthan et al. (2022) applied RF, ANN, and SVM to predict at-risk students during a university preparatory year, using a local dataset with SMOTE balancing. Their RF model achieved 99.66% accuracy and utilized LIME/SHAP for explainability. Gap: Confining data collection to one academic setting significantly narrows the scope of applicability [2].

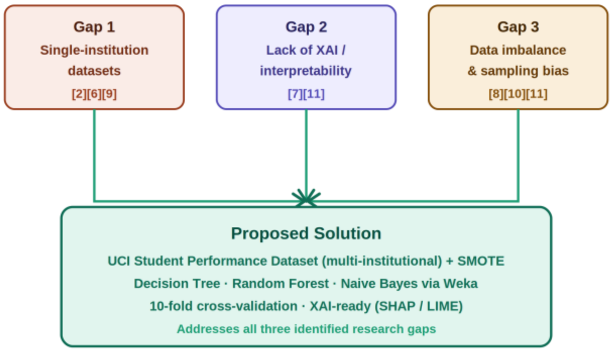

Fig. 1: Common Research Gaps Identified Across 12 Reviewed Studies and Proposed Solutions.

Zafari et al. (2022) conducted a systematic review exploring AI in K-12 education, focusing on ML, Deep Learning, and Intelligent Tutoring Systems. They found that RF, SVM, and DT provided the highest average accuracy for performance prediction. Gap: Many studies were excluded due to inaccessible full texts and insufficient attention to the creation of instructional support tools [3].

Gonzalez-Perez et al. (2025) reviewed data-driven learning analytics in higher education. Their analysis revealed that a substantial majority — approximately 80.6% — employed supervised learning techniques (e.g., predicting dropouts). Gap: There is a considerable disconnect between predictive models and the multifaceted nature of academic datasets, with data lacking integrated ethical and psychological frameworks [4].

Alhazmi & Sheneamer (2023) utilized Logistic Regression, RF, KNN, SVM, and XGBoost on a large institutional dataset of over 275,000 records. Using t-SNE for dimensionality reduction, their hybrid feature models showed significant accuracy improvements. Gap: The findings underscore the necessity of incorporating advanced neural network-based frameworks and blending non-academic features with academic ones [5].

Kala et al. (2024) proposed a hybrid deep learning model (PSO-DNN) alongside traditional models (KNN, LR, RF, DT) using a face-to-face programming course dataset and xAPI-Edu-Data. The PSO-DNN model outperformed classical models by 6% and utilized SHAP/LIME for interpretability. Gap: Concentrating exclusively on the institution constrains the wider applicability of the findings [6].

Alamri & Alharbi (2021) conducted a systematic review investigating explainable student performance models, evaluating Decision Trees, rule-based learning, and deep learning algorithms. Gap: Their review uncovered a critical shortage of research employing interpretability-oriented assessment criteria [7].

Alshanqiti & Namoun (2020) developed a hybrid regression model combining collaborative filtering, fuzzy rules, and Lasso linear regression, alongside a multi-Label Self-organizing Map (MLSOM). Tested on seven datasets, the model dynamically adjusted weights to achieve low RMSE. Gap: Publicly available data sources are frequently susceptible to selection bias and omit key indicators such as learner participation [8].

Pelima et al. (2024) conducted a meta-analysis of 70 academic papers that predict university graduation based on data mined from LMS and SIS. The most commonly used algorithms were SVM and RF. Gap: Systematic data skewness, limited sample populations, and poor cross-contextual transferability remain persistent challenges [9].

Sahlaoui et al. (2021) used ensemble methods (XGBoost, Extra Trees) together with SMOTE on the Jordanian xAPI-Edu-Data (480 records), which performed over 98% accuracy. To enhance transparency, SHAP values were integrated. Gap: Deemed inappropriate for deployment within active, real-time academic environments due to such a small offline dataset [10].

Asthana et al. — Modeling Educational Outcomes (2023), Learning Coefficients for Predictive Performance using Adaptive Assessments (October 2023) — Linear Regression, RF, DT, SVR applied to 91 CSE undergrad records. Linear Regression achieved 97% accuracy. Gap: The restricted data volume considerably hindered the evaluation of sophisticated neural architectures, introducing overfitting risks [11].

Mengash (2020) performs predictions for the university admissions performance of 2,039 students, reportedly applying ANN, DT, SVM, and Naive Bayes. With the use of pre-admission data, including SAAT and GAT scores, the ANN model delivered 79.22% accuracy. Gap: Contextual determinants, including personal attributes, familial circumstances, and financial standing, were excluded from analysis [12].

The studies consistently showcase three emerging research gaps: an over-reliance on single-institution datasets hindering generalizability [11], a lack of adoption of model interpretability frameworks [7], and neglect for non-academic predictors like socioeconomic and behavioral features [8]. The presence of these gaps underpins the motivation for the proposed method in Section 4, which proposes a comparative ML framework enhanced by XAI to tackle interpretability and dataset diversity.

4. PROPOSED METHOD

Based on the review of the literature, critical shortcomings identified encompass restricted transferability arising from reliance on data from isolated academic settings [2][6][9], data imbalance [5][10], and a persistent and alarming deficiency in algorithmic transparency [7][11]. The proposed method in the paper aims to mitigate these issues, and it integrates aggregated and transparent predictive classifiers, thus attaining an effective balance between classification accuracy and model transparency.

- **Dataset:** The UCI Student Performance Dataset contains multi-institutional data standardized in terms of number of records and attributes comprising demographic, academic, and behavioral features across 33 features from a total of 649 student records to overcome the problem of local sampling bias.
- **Comparative ML Algorithms:** Decision Tree (DT), Random Forest (RF), and Naive Bayes (NB).

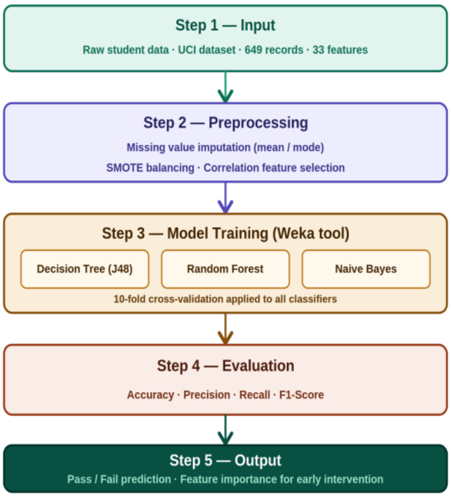

Fig. 2: Proposed Machine Learning Workflow for Student Performance Prediction using Weka.

Workflow:

a. As input, one gets raw student data: historical grades, demographic details, and behavioral metrics.

b. Preprocessing: Cleanses data by dealing with missing values. If numerical data is still necessary, it is discretized accordingly, and SMOTE (Synthetic Minority Over-sampling Technique) balances the class distributions (e.g., Pass vs. Fail).

c. Model — using Weka, you would input the pre-processed data. The DT, RF, and NB models are trained and tested using a 10-fold cross-validation scheme to avoid overfitting.

d. Output: The models predict the corresponding academic performance classes and also evaluate feature importance to provide a transparent basis for early interventions.

e. The systematic framework guaranteed consistency, openness, and cross-method comparability of results across all three classifiers.

5. EXPERIMENTAL SETUP

This section describes the dataset, preprocessing pipeline, algorithm configurations, and evaluation criteria used in the proposed framework.

- **Dataset Description:** The proposed methodology utilizes the UCI student performance Dataset, containing 649 student records with 33 attributes collected from two Portuguese secondary schools. Features include demographic attributes, social behaviours, and past academic grades. The target variable is the final grade (G3), categorized into binary classes: Pass (G3 $\geq$ 10) and Fail (G3 < 10). This dataset functions as a standardised reference point, guaranteeing experimental consistency.
- **Preprocessing steps:** First, data cleaning handles missing values via mean/mode imputation. Due to typical class imbalances in educational data (where minority classes often represent failing students), SMOTE is employed to rectify distributional imbalances within the dataset as proposed by Chawla et al. [13]. A subsequent attribute selection process isolates the most contributory input variables and reduces computational load.
- **Algorithms and parameters:**

1. Decision Tree (J48 in Weka): Standard pruning confidence factor (25), minimum instances per leaf (2) to maintain interpretability.
2. Random Forest: Number of trees set to 100, standard maximum depth, utilising the Gini Impurity criterion to handle complex, non-linear relationships.

   Recall is accorded the highest priority to minimise instances of undetected academically vulnerable learners.
3. Naive Bayes: Default parameters with kernel density estimators to handle numerical attributes probabilistically.

Evaluation Metrics: The models are evaluated using accuracy, precision, recall, and F1-Score. All models are assessed using a 10-fold cross-validation to ensure robust and unbiased performance estimation. Recall is accorded the highest priority to minimise instances of undetected academically vulnerable learners.

6. RESULT AND DISCUSSION

Table 1 presents a comparative summary of the expected performance of the three selected ML algorithms, based on findings synthesized from the reviewed literature.

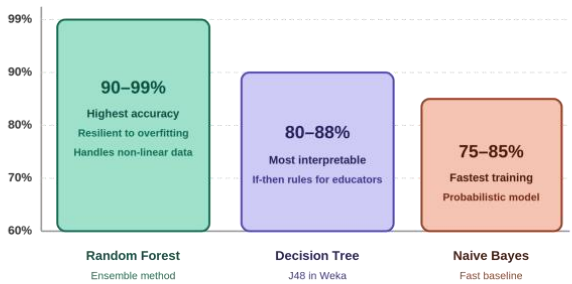

Fig. 3: Comparative Expected Accuracy Ranges of ML Algorithms [2][5][10][11][12].

7. CONCLUSION

The presented study conducted a thorough examination of leading-edge ML approaches for student performance prediction and introduced a replicable framework designed to bridge identified shortcomings in prior literature. The literature demonstrates that ML-powered alert mechanisms equip teaching staff to administer prompt, individualised academic support, consequently lowering the incidence of student withdrawal. Among the algorithms reviewed, the random forest classifier repeatedly demonstrated the strongest predictive capability, achieving an accuracy rate of up to 99% [2] when paired with data balancing techniques like SMOTE. The proposed framework, utilising the UCI student performance dataset and the Weka tool, establishes a verifiable and interpretable reference framework for subsequent investigations. However, the study acknowledges several limitations in the current field, primarily an overdependence on geographically and institutionally constrained datasets that hinder progress, and the frequent failure to incorporate non-academic predictors. To make complex models understandable for non-technical educators, forthcoming investigations ought to emphasise the adoption of explainability-oriented methodologies, such as SHAP and LIME [2][6], within highly accurate ensemble models. Furthermore, future studies must aim to use large, cross-institutional datasets and real-time learning management system (LMS) data to dynamically and openly revise forecasts throughout the academic period.

8. REFERENCES

[1] J. Pan, Z. Zhao, and D. Han, "Academic Performance Prediction Using Machine Learning Approaches: A Survey," *arXiv preprint*, 2025.

[2] S. Alwarthan, N. Aslam, and I. U. Khan, "An Explainable Model for Identifying At-Risk Students in Higher Education," *IEEE Access*, vol. 10, pp. 107649–107668, 2022, doi: 10.1109/ACCESS.2022.3211070.

[3] M. Zafari, J. Safari Bazargani, A. Sadeghi-Niaraki, and S.-M. Choi, "Artificial Intelligence Applications in K-12 Education: A Systematic Literature Review," *IEEE Access*, vol. 10, pp. 61585–61612, 2022, doi: 10.1109/ACCESS.2022.3179356.

[4] L. I. González-Pérez, F. J. García-Peñalvo, and A. J. Argüelles-Cruz, "Data-Driven Learning Analytics and Artificial Intelligence in Higher Education: A Systematic Review," *IEEE Rev. Iberoam. Tecnol. Aprendiz*, Early Access, 2025, doi: 10.1109/RITA.2025.3615512.

[5] E. Alhazmi and A. Sheneamer, "Early Prediction of Students' Performance in Higher Education," *IEEE Access*, vol. 11, pp. 24687–24699, 2023, doi: 10.1109/ACCESS.2023.3250702.

[6] A. Kala, O. Torkul, T. Tunacan Yildiz, and I. H. Selvi, "Early Prediction of Student Performance in Face-to-Face Education Environments: A Hybrid Deep Learning Approach With XAI Techniques," *IEEE Access*, vol. 12, pp. 12456–12470, 2024, doi: 10.1109/ACCESS.2024.3516816.

[7] R. Alamri and B. Alharbi, "Explainable Student Performance Prediction Models: A Systematic Review," *IEEE Access*, vol. 9, pp. 33132–33143, 2021, doi: 10.1109/ACCESS.2021.3061368.

[8] A. Alshanqiti and A. Namoun, "Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification," *IEEE Access*, vol. 8, pp. 203827–203844, 2020, doi: 10.1109/ACCESS.2020.3036572.

[9] L. R. Pelima, Y. Sukmana, and Y. Rosmansyah, "Predicting University Student Graduation Using Academic Performance and Machine Learning: A Systematic Literature Review," *IEEE Access*, vol. 12, pp. 23567–23589, 2024, doi: 10.1109/ACCESS.2024.3361479.

[10] H. Sahlaoui, E. A. Abdellaoui Alaoui, A. Nayyar, S. Agoujil, and M. M. Jaber, "Predicting and Interpreting Student Performance Using Ensemble Models and Shapley Additive Explanations," *IEEE Access*, vol. 9, pp. 152688–152703, 2021, doi: 10.1109/ACCESS.2021.3124270.

[11] P. Asthana, S. Mishra, N. Gupta, M. Derawi, and A. Kumar, "Prediction of Students' Performance With Learning Coefficients Using Regression-Based Machine Learning Models," *IEEE Access*, vol. 11, pp. 74589–74602, 2023, doi: 10.1109/ACCESS.2023.3294700.

[12] H. A. Mengash, "Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems," *IEEE Access*, vol. 8, pp. 55462–55470, 2020, doi: 10.1109/ACCESS.2020.2981905.

[13] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.