Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 15, Issue. 5, May 2026, pg.107 – 117

EduPredict: An AI-Driven Academic Performance Prediction System for Early Identification and Personalized Intervention of At-Risk College Students

Reijus M. Cruz[1]; Patrick B. Solis[2]; Rosa Mae Bascar[3]; Jomar A. Mondero[4]; Roden B. Pelegrino Jr.[5]; Jeffric S. Pisuena[6]; Kristine T. Soberano[7]

[1,2,3,4,5,6,7] Graduate School, State University of Northern Negros, Philippines

[1] reijus.orange@gmail.com; [2] patricksolis007@gmail.com; [3] rosamaebascar@gmail.com; [4] jomzmondero@gmail.com; [5] pelegrinorodenjr@gmail.com; [6] jpisuena@sunn.edu.ph; [7] ksoberano@sunn.edu.ph

DOI: https://doi.org/10.47760/ijcsmc.2026.v15i05.012

Abstract: Academic underperformance is a persistent challenge in higher education that conventional grade-based monitoring frameworks are structurally ill-suited to address in a timely manner. This paper presents EduPredict, an AI-driven student academic performance prediction and intervention management system built on a Random Forest classifier integrated within a four-tier analytics pipeline that encompasses descriptive, diagnostic, predictive, and prescriptive analytics. The system analyzes six class record indicators—attendance rate, class standing score, task average, examination average, grade trend, and year level—to classify students into high, medium, or low academic risk categories and generate individualized intervention recommendations. Evaluated on 856 class record entries from 172 college students at a Philippine state university, the model achieved an accuracy of 93.02%, precision of 0.9281, recall of 0.9302, and an F1-score of 0.9272. Five-fold cross-validation produced a mean accuracy of 91.96% (±1.59%), confirming stable generalization within the training domain. Feature importance analysis identified attendance rate as the dominant risk predictor, contributing 45.94% of total model weight. Beyond risk stratification, EduPredict incorporates a pre-examination minimum score projection module and a what-if simulation interface, extending the system's utility from passive risk identification to active faculty intervention planning. Results demonstrate the feasibility of deploying explainable, data-driven academic monitoring tools in resource-constrained state university settings.

Keywords: academic performance prediction; educational data mining; learning analytics; early warning systems; Random Forest; prescriptive analytics; at-risk students; higher education

I. INTRODUCTION

Academic underperformance is a documented and persistent problem in higher education institutions worldwide, affecting not only individual learner success but also institutional retention rates, accreditation standing, and the broader social returns on educational investment. Students who underperform often exhibit identifiable early warning signals—declining grades, sporadic attendance, and deteriorating formative assessment scores—well before their academic standing becomes critical. Yet in most Philippine state universities, academic monitoring continues to rely on periodic grade consolidations, end-of-term evaluations, and manual teacher observations: mechanisms that are fundamentally reactive and temporally misaligned with the pace at which academic decline typically unfolds. By the time a struggling student is formally identified under conventional practices, the window for low-cost, high-impact intervention has frequently already closed [1], [2].

The scale of this challenge in the Philippine higher education system is not trivial. State universities serve a disproportionately large share of students from economically constrained backgrounds, for whom academic failure carries outsized consequences—loss of scholarships, delayed graduation, and reduced employment prospects. Unlike private institutions that may have robust academic advising infrastructure, many state universities operate with limited faculty-to-student ratios and minimal dedicated support staff for academic monitoring. This structural constraint makes it particularly important that faculty be equipped with tools that multiply their capacity to identify and respond to at-risk learners without requiring disproportionate additional workload.

The rapid maturation of artificial intelligence and machine learning methodologies has opened substantive new avenues for addressing this gap. Machine learning models, particularly ensemble classifiers, are capable of detecting complex, non-linear patterns in academic data that are invisible to human reviewers working with raw grade records alone. Educational data mining (EDM)—the application of data mining techniques to educational datasets—has emerged as a productive sub-discipline for transforming institutional academic records into predictive and diagnostic insights. Romero and Ventura [6] established, through an updated survey of over 200 EDM studies, that classification, prediction, and clustering are the three most widely applied EDM techniques in higher education, with ensemble methods consistently outperforming single-classifier baselines across institutional and disciplinary contexts. The progression from descriptive reporting toward fully prescriptive, recommendation-capable analytics represents the current frontier of the field.

Machine learning applications in student performance prediction have accelerated considerably since 2020. Alyahyan and Düstəgor [4] conducted a comprehensive literature review and identified Random Forest and gradient-boosted decision trees as the most frequently adopted classifiers in student academic prediction, with reported accuracies frequently exceeding 85% across single-institution studies. Their review also highlighted that attendance, formative assessment scores, and prior academic performance consistently emerged as the three most influential predictors across diverse institutional contexts. Wang [1] demonstrated in a more recent study that AI-integrated student management systems deployed at the class record level—as opposed to systems relying on self-reported student data or learning management system logs—significantly improved the timeliness of academic monitoring and enabled earlier faculty response. Alkan et al. [2] extended this finding by showing that formative assessment data, when processed by machine learning models, could identify academically underperforming students substantially earlier than conventional final-grade cutoffs permit, providing a wider intervention window for faculty action.

Early warning systems (EWS) in education have been an active research area since the early 2010s, but their practical adoption in Philippine higher education institutions remains limited. Namoun and Alshanqiti [7] conducted a systematic review of 35 machine learning studies on student dropout prediction and found that attendance-related features ranked among the top three predictors in 94% of reviewed studies, regardless of institutional geography, academic discipline, or data granularity. This remarkable cross-contextual consistency underscores the behavioral significance of attendance as a leading indicator of academic risk—a signal available in every class record and therefore accessible even in institutions without sophisticated data infrastructure. Dutt et al. [8] similarly established that clustering-based behavioral profiling, when combined with classification models, produces more differentiated and instructionally actionable student groupings than classification alone, enabling faculty to tailor intervention types to distinct learner subgroups rather than applying uniform remediation strategies.

Despite these advances, a persistent operational gap remains between what contemporary prediction systems output and what educators actually need. Most existing systems produce risk scores or binary at-risk flags without providing faculty with concrete, student-specific guidance on what to do next, or the means to explore what would happen to a student's trajectory if a particular intervention were implemented. Cao and Mai [5] argued that the transition from predictive-only systems to prescriptive analytics platforms represents the critical next step in educational risk management: actionable guidance—not risk labels alone—is the currency that drives meaningful improvement in student outcomes. Mduma [9] further identified that data imbalance handling and the integration of behavioral context into prediction pipelines are among the most critical unresolved challenges in deploying student prediction systems in real institutional settings. These gaps define the design requirements that EduPredict was built to address.

EduPredict addresses these limitations by integrating risk prediction with prescriptive analytics in a unified, deployable platform designed for the operational realities of Philippine state universities. The system's primary contributions are fourfold: (1) a three-tier risk classification engine trained on class record data using a Random Forest classifier with cross-validated performance; (2) an automated, personalized intervention recommendation module that translates prediction outputs into specific faculty-assigned support actions; (3) a pre-examination minimum score projection module that converts abstract risk classifications into concrete, communicable academic targets for each student; and (4) a what-if simulation interface that allows faculty to evaluate the projected classification impact of targeted behavioral changes—such as attendance improvement or task completion—before committing institutional intervention resources. The system targets teachers, students, and parents within a Philippine state university context, where institutional support infrastructure is constrained and technology-enhanced early warning tools remain largely absent from daily academic operations.

The remainder of this paper is organized as follows: Section II enumerates the specific research objectives. Section III describes the methodology, encompassing the research design, conceptual framework, system architecture, dataset, preprocessing pipeline, analytics techniques, and evaluation protocol. Section IV presents and critically interprets the experimental results across five output modules. Section V concludes with a synthesis of findings, limitations, and directions for future research.

II. OBJECTIVES OF THE STUDY

The primary purpose of this research is to design, develop, and evaluate an AI-driven system that predicts the academic performance of college students, identifies those at risk of underperformance or failure, and generates timely, data-driven intervention recommendations to support academic recovery and learner success. The system is intended to empower teachers, academic advisers, and school administrators with actionable intelligence derived from class record data, enabling proactive monitoring and evidence-based decision-making throughout the academic term.

Specifically, this study sought to achieve the following objectives:

1. To develop a machine learning-based system that analyzes academic class record data to predict individual student performance levels and estimate the likelihood of academic failure using a Random Forest classification model.

2. To identify the most significant academic indicators influencing student performance outcomes, including attendance rate, class standing, task performance, examination scores, grade trends, and subject-specific patterns, through feature importance analysis.

3. To classify students into risk categories—high, medium, and low—based on their cumulative academic data, enabling faculty to prioritize monitoring and support resources toward the most vulnerable learners.

4. To generate pre-examination performance projections that compute the maximum achievable final grade and the minimum examination score required for each student to meet the institutional passing threshold, providing concrete academic targets prior to high-stakes assessments.

5. To deliver personalized, data-driven academic intervention recommendations—such as targeted remedial exercises, peer or faculty tutoring, academic counseling, or structured review sessions—tailored to each student's specific risk profile and performance gaps.

6. To incorporate behavioral and longitudinal performance analytics that identify distinct student performance patterns—including consistent performance, progressive decline, chronic underperformance, improvement trajectories, and disengagement signals—using clustering techniques.

7. To provide faculty, academic advisers, and administrators with prediction insights, confidence scores, and recommended actions aligned with each student's current academic status, supporting timely and differentiated instructional responses.
8. To evaluate the system's effectiveness in improving academic monitoring precision, accelerating at-risk student identification, and facilitating appropriate intervention planning through empirical assessment of model performance metrics and output quality.

III. METHODOLOGY

A. Research Design

This study employed a developmental research design, which is appropriate for investigations whose primary deliverable is a functional technological artifact evaluated for both technical performance and practical utility in educational settings [4]. The developmental paradigm guided the construction, testing, and iterative refinement of the EduPredict system across distinct implementation phases. A quantitative evaluation phase was embedded within the developmental framework to assess the machine learning model's performance using standard classification metrics and cross-validation procedures. The study further integrated four analytics modalities—descriptive, diagnostic, predictive, and prescriptive—each serving a distinct functional role in the academic monitoring pipeline.

B. System Architecture

EduPredict adopts a five-layer software architecture that integrates data management, artificial intelligence, and user interaction components within a cohesive and modular platform. The Data Sources layer constitutes the origin of all academic information, drawing from teacher-encoded class records, student assessment results, attendance logs, and historical performance data across academic terms. The Data Management Layer is responsible for preparing raw inputs for analysis through data collection, validation, cleaning, normalization, and feature engineering—including the derivation of grade trend coefficients and performance consistency indicators. The AI and Prediction Engine serves as the analytical core of the system, applying the Random Forest classifier and K-Means clustering to processed features to generate risk classifications, behavioral profiles, and performance forecasts. The Recommendation and Alert Module translates prediction outputs into actionable support by generating personalized intervention assignments, automated early warning notifications to students and parents, and structured faculty alerts. The User Interface Layer provides role-differentiated access to system outputs: teachers access risk dashboards, grade analytics, and intervention management tools; students view performance summaries, risk visualizations, and improvement suggestions; and parents receive notification alerts and student performance monitoring screens. The complete system was implemented using Python (Flask framework, SQLite database, scikit-learn ML layer) with an HTML/CSS/JavaScript frontend, enabling lightweight deployment without dependency on enterprise-grade infrastructure.

C. System Users

EduPredict was designed to serve three distinct user roles with role-appropriate feature access. Teachers function as the primary system users and are responsible for encoding class record data, monitoring student risk levels, and assigning interventions. Teacher-accessible features include class record encoding, student performance dashboards, risk classification monitors, grade trend analysis, behavioral cluster visualization, intervention assignment and tracking, pre-exam projection tools, what-if simulation, and analytics reports. Students access an individual performance interface that provides academic dashboards, risk level visualizations, grade trend graphs, pre-exam score projections, personalized recommendations, scenario simulations, and real-time notifications. Parents and guardians are included as system participants through a monitoring interface that delivers student performance summaries, risk-level notifications, and SMS-integrated alerts, reinforcing the home-school support connection and enabling early parental involvement when students are flagged as high risk.

D. Dataset and Participants

The study dataset comprised 856 academic records collected from 172 college students enrolled across selected courses at a Philippine state university from academic year 2023 onward. Multiple records per student across academic terms and subject enrollments provided longitudinal variability in the dataset. Students were drawn from diverse sections and performance levels to ensure class label distribution variability, which is essential for unbiased classifier training. Data were sourced exclusively from official institutional class records and encoded into the EduPredict system through teacher-managed Excel file uploads or direct manual entry. The study restricted data collection to class-record-available variables to ensure replicability in institutional contexts that lack sophisticated data infrastructure.

Six academic features were selected as model predictors based on their documented relevance to academic risk in the EDM literature [6], [7] and their consistent availability in standard Philippine higher education class records: (1) Attendance Rate—the proportion of scheduled class meetings attended; (2) Class Standing Score—a composite grade reflecting accumulated class participation, conduct, and standing components; (3) Task Average—the mean score across all assigned tasks, projects, and outputs; (4) Examination Average—the mean of mid-term and final examination scores; (5) Grade Trend—a scalar coefficient derived by fitting a linear regression model to each student's sequential grade entries, capturing the directional trajectory of performance over time; and (6) Year Level—the student's current academic year. The target variable was a three-class risk label—High, Medium, and Low—derived from cumulative grade thresholds aligned with the institution's established grading policy.

E. Data Gathering Procedure

The data gathering procedure followed a structured workflow. Teachers collected student academic records from class activities and organized the data in spreadsheet format aligned with the system's import schema. Completed records were uploaded into EduPredict via the Excel import feature, with supplementary records encoded manually as needed. The system performed automated validation on all uploaded records, flagging missing or inconsistent entries for teacher review. Validated records were stored in the relational database, after which the analytics engine processed the dataset to generate risk classifications, projections, and intervention recommendations. Notifications were then dispatched to students and parents based on generated risk outputs, completing the data-to-action cycle.

F. Data Preprocessing and Preparation

Raw academic records underwent a multi-step preprocessing pipeline before entry into model training. Missing value handling involved the removal of records with incomplete grade entries to avoid imputation-induced bias, given the dataset's modest size. All continuous features were scaled to the [0, 1] range using min-max normalization, ensuring that features with larger absolute ranges—such as raw examination scores—did not disproportionately dominate distance-based or threshold-based computations. Categorical fields, including year level, were ordinally encoded. Duplicate entries—operationally defined as multiple records for the same student-term-subject combination—were identified and removed. Grade Trend was derived programmatically by applying linear regression to each student's chronological grade sequence and extracting the slope as a scalar predictor, ensuring that directional performance dynamics were captured as a standalone feature rather than subsumed within grade averages. Standardized grade formats were applied across all records to ensure consistency across sections and academic terms.

G. Analytics Techniques and Models

Descriptive analytics was implemented to generate summary-level visualizations of student performance data, including class records tables, grade distribution charts, attendance statistics, risk level pie charts, section average grade summaries, and interactive dashboard elements. These outputs established the performance baseline from which diagnostic and predictive analyses proceeded.

Diagnostic analytics was applied to identify factors contributing to academic risk at the individual and group level. Feature importance scores derived from the Random Forest model were used to rank predictors by contribution. K-Means clustering (k=3) was applied post-prediction to assign behavioral performance profiles to student cohorts, operationalized as consistently performing, progressively declining, and chronically disengaged subgroups. Performance trend analysis and subject-difficulty profiling further enriched the diagnostic layer.

Predictive analytics constituted the system's primary analytical function. A Random Forest classifier [3] was employed to assign each student to one of three risk categories based on their academic feature vector. Linear regression was applied independently to model grade trends and compute pre-examination projections. Rule-based logic was used to calculate the maximum achievable final grade and the minimum examination score required for each student to cross the institutional passing threshold, given their current cumulative performance.

Prescriptive analytics generated actionable outputs from prediction results. Auto-generated recommendations—including directives for extra quizzes, peer or faculty tutoring, academic counseling, and attendance improvement—were mapped to specific risk classification outcomes through a rule-based assignment schema. A what-if simulation module allowed faculty to input hypothetical changes to student academic variables and observe the resulting predicted risk reclassification, enabling pre-intervention scenario planning. Intervention logs with status tracking, automated student and parent notifications, and targeted academic support suggestions completed the prescriptive layer.

H. Machine Learning Model Configuration

The Random Forest classifier [3] was selected as the primary prediction model for several well-documented reasons: its ensemble architecture provides natural resistance to overfitting relative to individual decision trees; its feature importance scores offer built-in model interpretability without post-hoc explainability methods; and its robustness to feature correlation and class imbalance makes it appropriate for the multi-collinear academic indicator dataset used in this study [4]. The model was implemented using the scikit-learn library [10]. Hyperparameter optimization was conducted through exhaustive grid search with stratified 5-fold cross-validation, optimizing for macro-averaged F1-score to account for unequal class distribution. Final hyperparameters comprised 100 decision tree estimators, a maximum tree depth of 10, and a minimum sample split threshold of 5. Class weights were set to 'balanced' to mitigate the effect of skewed risk label distribution on training, ensuring that the minority class (High Risk) received proportionally weighted learning attention.

I. System Development Methodology

EduPredict was constructed using the Incremental Development Model, which organizes system construction into a sequence of independently testable modules that are progressively integrated into the complete system. This approach was selected because it accommodates iterative refinement based on formative testing outcomes without requiring full system restart, which is particularly appropriate for complex multi-module academic systems. The ten development stages were executed sequentially: user authentication module; student management module; class record encoding module; analytics engine integration; prediction module implementation; recommendation engine development; intervention tracking module; notification module; what-if simulation module; and reports and analytics dashboard. Each module was subjected to functional testing and user acceptance evaluation prior to integration with adjacent modules, ensuring that defects were identified and resolved at the component level before compounding in the integrated system.

J. Evaluation Protocol

Model performance was evaluated using four standard binary-class and multi-class classification metrics: accuracy (proportion of correctly classified instances), precision (ratio of true positives to all predicted positives), recall (ratio of true positives to all actual positives), and F1-score (harmonic mean of precision and recall). Stratified 5-fold cross-validation was employed as the primary validation strategy, preserving class proportions across all training-evaluation splits. The narrow standard deviation of cross-validated accuracy was used as an empirical indicator of model stability. Given the dataset's size of 856 records, a separate held-out test partition was not reported independently of cross-validation, a limitation that is discussed further in Section IV.

K. Ethical Considerations

Ethical standards were observed throughout the data collection, processing, and system deployment phases of the study. Student academic records were handled exclusively for the purposes of this research and system evaluation, with confidentiality maintained through secure login authentication and role-based access control. No sensitive personal information beyond academic records was collected or disclosed. Students and faculty were informed of the system's purpose and operational scope. Data were stored within the institution's local system infrastructure and were not shared with external parties.

IV. RESULTS AND DISCUSSION

A. Model Performance

Table I presents the classification performance metrics of the EduPredict Random Forest model. The system achieved an overall accuracy of 93.02% on the class record dataset, with precision, recall, and F1-score values of 0.9281, 0.9302, and 0.9272, respectively. The close alignment between precision and recall indicates that the model avoids the asymmetric error profile common in imbalanced classification tasks, where high accuracy is achieved through systematic over-prediction of the majority class at the expense of minority class detection. In the context of academic monitoring, this balance is operationally important: a model with high precision but low recall would miss a large proportion of genuinely at-risk students, while a model with high recall but low precision would overwhelm faculty with false alerts, eroding trust in the system over time.

The 5-fold cross-validated mean accuracy of 91.96% (±1.59%) is particularly informative. The narrow standard deviation across folds indicates that the model's performance is stable across different partitions of the training data and is not a product of a fortuitously favorable single train-test split. This internal consistency is reassuring given the dataset's modest scale of 856 records and suggests that the model captures generalizable academic patterns within the institutional context rather than memorizing idiosyncratic record-level features.

These performance figures are comparable to those reported in analogous single-institution studies employing Random Forest on class record data [1], [2]. However, it is important to note that these metrics reflect performance within the training institution's specific grading structure, course composition, and student population. Cross-institutional replication is necessary before broader generalizability claims can be empirically substantiated. The absence of a separate held-out test set, necessitated by the dataset's size, represents a methodological constraint that future work with larger datasets should address.

TABLE I. Model Performance Metrics

Metric	Value	Interpretation
Accuracy	93.02%	Strong overall classification performance
Precision	0.9281	Low false-positive identification rate
Recall	0.9302	High true-positive detection rate
F1-Score	0.9272	Balanced precision–recall trade-off
Cross-Val. Accuracy (5-fold)	91.96% ± 1.59%	Stable generalization across data partitions

B. Feature Importance Analysis

Table II presents the ranked feature importance scores extracted from the trained Random Forest classifier using Gini impurity reduction. Attendance Rate emerged as the overwhelmingly dominant predictor, contributing 45.94% of the model's total decision-making weight—more than double the contribution of the second-ranked feature and collectively greater than all remaining features combined. This finding aligns with the cross-contextual evidence documented by Namoun and Alshanqiti [7], who reported attendance as a top-three predictor in 94% of reviewed studies spanning diverse institutional and disciplinary settings. The physiological and behavioral mechanisms underlying this relationship are well-established: regular class attendance provides students with structured exposure to instructional content, formative feedback cycles, and social academic reinforcement, all of which compound into measurable grade outcomes over time.

Within the institutional context of this study, attendance's dominant importance weight may be further amplified by the grading policy's formal inclusion of attendance as a component of the class standing score, creating a structural feedback loop in which attendance influences both the Attendance Rate feature directly and the Class Standing Score feature indirectly. This collinearity does not invalidate the finding but does suggest that disentangling the direct behavioral effect of attendance from its policy-mediated grade effect would require more granular data than class records alone can provide.

Class Standing Score ranked second at 21.62%, reflecting the cumulative influence of ongoing academic participation and conduct components embedded in the institutional grading scheme. Task Average (11.91%) and Examination Average (11.12%) together accounted for approximately 23% of total model weight, suggesting that formative and summative assessment data carry meaningful collective predictive power even when neither individually approaches the influence of attendance or class standing. This pattern is consistent with the observation that formative assessment data—when aggregated—can serve as a more reliable predictor than any single summative examination score [2]. Grade Trend contributed 8.31% of model weight, serving as an early directional signal: a declining slope in a student's grade trajectory may predict future risk independently of the student's current absolute grade level, enabling intervention before grades cross critical thresholds. Year Level was the least influential predictor at just 1.09%, confirming that academic risk is more accurately characterized by behavioral and performance dynamics than by the student's position in the academic program.

These findings carry direct instructional implications for faculty intervention design. Given that attendance accounts for nearly half of the model's predictive weight, institutional interventions focused on improving student attendance—through engagement strategies, proactive outreach, or policy adjustment—are likely to produce greater risk reclassification outcomes than those focused solely on examination preparation or task submission. The What-If Simulator results in Section IV-D empirically validate this inference.

TABLE II. Feature Importance Rankings

Rank	Feature	Importance Score	Contribution (%)
1	Attendance Rate	0.4594	45.94
2	Class Standing Score	0.2162	21.62
3	Task Average	0.1191	11.91
4	Examination Average	0.1112	11.12
5	Grade Trend	0.0831	8.31
6	Year Level	0.0109	1.09

C. Pre-Examination Score Projections

Table III presents sample outputs from EduPredict’s pre-examination score projection module. The module applies rule-based grade computation logic to each student’s current cumulative academic record to determine two values: the maximum achievable final grade given the student’s current standing, and the minimum examination score required to cross the institutional passing threshold. These projections are generated and made available to both teachers and students before high-stakes examinations, providing a concrete, individualized academic target rather than a generalized risk label.

Among the three sampled students, Student A and Student C face achievable minimum score requirements of 76% and 70% respectively, placing them within a realistic performance range for a prepared student. Student B, however, requires 97% (48 of 50 points) to pass—a threshold that represents an extremely narrow margin and warrants immediate, prioritized academic intervention. For Student B, the pre-examination projection does not merely confirm High Risk classification; it communicates the quantitative severity of that risk in terms that are directly actionable for both the student and the supervising faculty member. This translation of abstract risk into concrete numerical targets is one of EduPredict’s key operational differentiators from prediction-only systems, and aligns with the prescriptive analytics paradigm advocated by Cao and Mai [5] as the critical frontier of educational risk management.

TABLE III. Sample Pre-Examination Score Projections

Student	Min. Score Required	Percentage (%)	Feasibility Status
Student A	38 / 50	76	Achievable
Student B	48 / 50	97	Critical — Immediate Counseling Needed
Student C	35 / 50	70	Achievable

D. What-If Simulation Results

Table IV presents the results of EduPredict’s what-if simulation module, which allows faculty to input hypothetical changes in individual student academic variables and observe the resulting predicted risk reclassification. Three scenarios were tested using a representative High Risk student profile. In Scenario 1, improving attendance rate to 75% alone was sufficient to reclassify the student from High to Medium risk, a result consistent with the feature importance analysis showing attendance as the primary classification driver. In Scenario 2, raising task average to 80% without changing attendance produced no reclassification, confirming that task performance operates as a supporting rather than primary determinant in the classifier’s decision boundary at this risk level. In Scenario 3, combining both improvements (attendance at 75% and task average at 80%) produced the same reclassification as attendance improvement alone, confirming that the marginal contribution of task improvement is conditional on the prior satisfaction of the attendance threshold.

These results demonstrate the simulation module’s practical value as a pre-intervention planning tool. Faculty can use the simulator to evaluate the predicted return on different intervention strategies before allocating counseling, tutoring, or monitoring resources. For the student profile tested, the evidence unambiguously directs attention toward attendance-focused intervention as the highest-yield action. This kind of prescriptive specificity is precisely what Cao and Mai [5] identified as the missing operational bridge between prediction-capable systems and systems that actively drive improved student outcomes.

TABLE IV. What-If Simulator Results

Scenario	Original Class	Predicted Class	Outcome
Attendance Rate → 75%	High	Medium	Improved
Task Average → 80%	High	High	No Change
Attendance 75% + Task Average 80%	High	Medium	Improved

E. Automated Intervention Report

Table V presents a sample automated intervention report generated by EduPredict for a five-student cohort. Four of the five students were classified as High Risk, with three receiving directives to improve attendance immediately—a recommendation consistent with the feature importance findings that identify attendance as the dominant risk driver. Student S4 was additionally directed toward scheduled review sessions, indicating that while attendance was a contributing factor, academic preparation gaps also required targeted remediation. Student S5, the sole Low Risk student in the sample, was assigned a maintenance recommendation, reflecting a proactive orientation even for students currently on track.

The variation in required minimum examination scores across High Risk students—ranging from 31 to 48 out of 50—illustrates the system's capacity to communicate urgency gradients rather than uniform risk labels. Student S2, who requires 48 of 50 points to pass, faces a materially more precarious situation than Student S3, who requires only 31. Without this quantitative differentiation, both students would receive identical High Risk labels, and faculty might allocate intervention resources uniformly rather than directing proportionally greater attention toward students with the narrowest academic margins. EduPredict's individualized reporting therefore serves not only as an alert mechanism but as a resource allocation guide that enables faculty to deploy limited academic support capacity where it is most urgently needed.

TABLE V. Sample Automated Intervention Report

Student	Risk	Max. Grade (%)	Min. Exam Score	Status	Recommendation
S1	High	81.2	42 / 50	At Risk	Improve attendance immediately
S2	High	76.9	48 / 50	At Risk	Improve attendance immediately
S3	High	90.0	31 / 50	At Risk	Improve attendance immediately
S4	High	84.4	38 / 50	At Risk	Attend scheduled review sessions
S5	Low	92.0	29 / 50	On Track	Maintain current performance

F. Overall Discussion

Across the five output modules evaluated, EduPredict demonstrates a consistent and coherent analytical progression from raw class record data to differentiated, student-specific academic guidance. The system's integration of descriptive, diagnostic, predictive, and prescriptive analytics within a single platform addresses a well-documented operational gap in educational technology: the disconnect between systems that flag risk and systems that equip educators to act on that risk in practical, low-overhead ways [5], [8]. The feature importance analysis, what-if simulation, and pre-examination projections collectively provide faculty with three distinct lenses for understanding and responding to academic risk: the structural drivers of risk (feature importance), the most efficient intervention targets (simulation), and the student-specific stakes of inaction (pre-exam projections).

EduPredict should be understood as an intelligent decision-support tool rather than an autonomous academic management system. Its classifications and recommendations are designed to inform, not supplant, the professional judgment of educators. The system's value lies in its capacity to process and synthesize class record data at a scale and speed that no individual teacher can match manually, surface patterns that would otherwise remain invisible until students have already fallen critically behind, and translate those patterns into actionable, concrete guidance that faculty can implement within the normal rhythms of academic monitoring.

G. Limitations

The findings of this study are subject to several limitations that must be acknowledged. First, the dataset of 856 records from 172 students at a single institution is relatively modest in scale; the reported classification metrics reflect performance within one specific institutional context and cannot be assumed to transfer directly to institutions with different grading systems, course structures, or student demographics without retraining. Second, the feature set is restricted to indicators available in standard class records, excluding potentially high-value predictors such as learning management system engagement data, submission timestamps, socioeconomic status, scholarship classification, and extracurricular commitments—variables that prior research has identified as meaningful contributors to academic risk [7], [9]. Third, the study does not yet include longitudinal outcome validation: whether students identified by EduPredict as High Risk who received system-generated interventions demonstrated measurable grade improvement in subsequent assessments has not been formally tracked or reported. This is the most significant gap between the current system’s demonstrated performance and its claimed real-world efficacy, and addressing it is a primary priority for the study’s next phase.

V. CONCLUSION

This paper presented EduPredict, an AI-driven academic performance prediction and intervention management system that integrates a Random Forest classifier with a four-tier analytics pipeline to support proactive, data-driven student monitoring in Philippine state university settings. Evaluated on 856 class record entries from 172 college students, the system achieved 93.02% classification accuracy with stable cross-validated generalization (91.96% ± 1.59%), demonstrating that explainable, machine learning-based academic monitoring is feasible and reliable within the class record data paradigm available to most Philippine higher education institutions.

The feature importance analysis confirmed that the attendance rate emerged as the dominant predictor of academic risk, consistent with broader EDM literature, and the what-if simulation confirmed that improving attendance reliably shifts student risk classifications. Together with pre-examination projections and automated intervention reports, the system bridges the gap between prediction and actionable, student-specific academic support.

EduPredict does not replace the professional expertise of educators; it expands the scope and precision of what educators can monitor and act upon within the constraints of normal academic workloads. By making patterns in class record data visible at the student and section levels, and by generating intervention guidance at the speed of automated computation rather than manual analysis, the system enables a shift from reactive grade-based monitoring to proactive, evidence-informed academic support—a transition that the literature consistently identifies as the key lever for reducing academic failure rates in resource-constrained educational institutions [1], [5].

Future research should address the study’s current limitations through three primary directions. First, longitudinal outcome tracking should be conducted to validate whether EduPredict-generated interventions produce measurable improvement in student grade trajectories and retention across subsequent academic terms, establishing the system’s empirical efficacy beyond classification performance metrics. Second, multi-institutional deployment and cross-validation should be pursued, incorporating institutions with diverse grading systems, disciplinary compositions, and student demographics to rigorously assess the model’s transferability and identify conditions requiring retraining. Third, feature set expansion should incorporate engagement proxies such as learning management system activity, assignment submission patterns, and socioeconomic indicators, with SHAP (SHapley Additive exPlanations) analysis applied to maintain interpretability as model complexity increases. These extensions would advance EduPredict from a single-institution prototype toward a scalable, institutionally validated platform for AI-driven academic support in Philippine higher education.

References

[1]. Y. Wang, "Artificial intelligence in student management systems to enhance academic performance monitoring and intervention," Scientific Reports, vol. 15, 2025. doi: 10.1038/s41598-025-19159-4

[2]. B. B. Alkan, S. Kuzucuk, N. Alkan, and A. Sinan, "Using machine learning to predict student outcomes for early intervention and formative assessment," Scientific Reports, vol. 15, 2025. doi: 10.1038/s41598-025-23409-w

[3]. L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001. doi: 10.1023/A:1010933404324

[4]. E. Alyahyan and D. Düstəgor, "Predicting academic success in higher education: Literature review and best practices," International Journal of Educational Technology in Higher Education, vol. 17, no. 1, pp. 1–21, 2020. doi: 10.1186/s41239-020-00177-7

[5]. W. Cao and N. T. Mai, "Predictive analytics for student success: AI-driven early warning systems and intervention strategies for educational risk management," Educational Research and Human Development, 2025. doi: 10.28991/erhd-2025-01-01-743

[6]. C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," WIREs Data Mining and Knowledge Discovery, vol. 10, no. 3, e1355, 2020. doi: 10.1002/widm.1355

[7]. A. Namoun and A. Alshanqiti, "Predicting student performance and dropout in higher education using machine learning approaches: A systematic literature review," IEEE Access, vol. 9, pp. 36321–36338, 2021. doi: 10.1109/ACCESS.2020.3047736

[8]. A. Dutt, M. A. Ismail, and T. Herawan, "A systematic review on educational data mining," IEEE Access, vol. 5, pp. 15991–16005, 2017. doi: 10.1109/ACCESS.2017.2654247

[9]. N. Mduma, "Data balancing techniques for predicting student dropout using machine learning," Data, vol. 8, no. 3, p. 49, 2023. doi: 10.3390/data8030049

[10]. F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

