

AI/ML - BASED ACADEMIC PERFORMANCE INTELLIGENCE

Dr.K.E. Kannammal¹, Balaji S², Balaji S S³, Bharath Sivanesh S⁴, Balaji S S⁵

¹Professor and Head of the Department of Computer Science & Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India, Email: hodcse@srishakthi.ac.in

²Student, Department of Computer Science & Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India, Email: balajis22cse@srishakthi.ac.in

³Student, Department of Computer Science & Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India, Email: balajiss22cse@srishakthi.ac.in

⁴Student, Department of Computer Science & Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India, Email: bharathsivaneshs22cse@srishakthi.ac.in

ABSTRACT : The AI Academic Performance Intelligence System is an intelligent web-based platform designed to analyze, monitor, and predict student academic performance using Artificial Intelligence and Machine Learning. The system provides role-based access for Admin, Staff, and Students, enabling efficient academic management through structured data analysis. A Random Forest Classifier is employed to predict student pass or fail outcomes based on historical academic data and Course Outcome (CO)-based performance metrics. The platform integrates a RAG-based chatbot powered by LangChain for academic query resolution and an LLM-based study planner using Grok for generating personalized study schedules. The frontend is developed using Next.js, while Django and FastAPI handle backend services with PostgreSQL as the database.

I. INTRODUCTION

The rapid advancement of digital technologies in higher education has created an urgent need for intelligent systems capable of transforming raw academic data into actionable insights. Traditional academic performance evaluation methods rely heavily on manual data entry, periodic report generation, and reactive interventions that often occur after academic failure has already taken place. This

reactive approach limits the ability of educational institutions to proactively identify and support students who are at risk of underperformance.

The AI Academic Performance Intelligence System is developed to address these limitations by providing a comprehensive, data-driven academic management platform. The system integrates machine learning-based performance prediction, CO-based analysis, and AI-powered student support tools within a unified, scalable architecture. By combining the predictive capabilities of the Random Forest Classifier with the conversational intelligence of a RAG-based chatbot and an LLM-driven study planner, the system enables educational institutions to shift from reactive to proactive academic management.

II. LITERATURE REVIEW

Amrieh, E. A., Hamtini, T., & Aljarah, I. "Mining Educational Data to Predict Student's academic Performance using Ensemble Methods", *International Journal of Database Theory and Application*, 2016.

The research conducted by Amrieh, Hamtini, and Aljarah presents a comprehensive investigation into the application of ensemble machine learning methods for predicting student academic performance using

educational data mining (EDM) techniques. The authors address a significant challenge in contemporary higher education: the inability to proactively identify students who are likely to underperform before academic failure occurs. By leveraging ensemble-based classifiers, the study aims to overcome the limitations of traditional single-classifier approaches, which are often prone to overfitting and variance in prediction accuracy. The study utilizes a real-world dataset collected from an e-learning environment, incorporating behavioral features such as student engagement metrics, login frequency, assignment submission patterns, and interaction with learning resources.

Yadav, S. K., Bharadwaj, B., & Pal, S. "Data Mining Applications: A Comparative Study for Predicting Student's Performance", *International Journal of Innovative Technology and Creative Engineering*, 2012.

The research compares three widely used classification algorithms: Decision Trees (ID3), Naive Bayes, and Artificial Neural Networks (ANN). These algorithms are evaluated on a dataset comprising student attributes such as family background, study habits, attendance records, internal assessment marks, and participation in extracurricular activities. The study measures predictive performance using accuracy and confusion matrix analysis. The findings reveal that while all three methods achieve reasonable predictive accuracy, the Decision Tree algorithm produces the most interpretable results, making it particularly suitable for educational applications where transparency of prediction reasoning is valued.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. "Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?", *Journal of Machine Learning Research*, 2014.

The study employs a rigorous experimental methodology, using stratified cross-validation and statistical significance testing to ensure the validity of comparative results. Algorithm families evaluated include Random Forests, Support Vector Machines, Neural Networks, Linear Discriminant Analysis, Naive Bayes, Decision Trees, Boosting methods, and

numerous others. Each algorithm is tested in its default parameter configuration as well as with tuned hyperparameters, providing insights into both out-of-the-box performance and optimized performance. The key finding of the study is that Random Forests — specifically the implementation based on Breiman's original algorithm — consistently rank among the top-performing classifiers across the widest range of datasets and problem types.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. "Systematic Review of Research on Artificial Intelligence Applications in Higher Education", *International Journal of Educational Technology in Higher Education*, 2019.

The review categorizes AI applications in higher education into four primary domains: profiling and prediction of student performance and behavior, intelligent tutoring and personalized learning systems, assessment and evaluation automation, and institutional and administrative management including scheduling, resource allocation, and dropout prediction. Within each domain, the authors analyze the types of AI techniques employed, the nature of the educational data utilized, and the institutional contexts in which AI systems have been deployed. The review reveals that machine learning-based prediction and classification tasks constitute the largest and most rapidly growing area of AI application in higher education

Alfredo, R., Milesi, M., Echeverria, V., & Martinez-Maldonado, R. "Co-Designing AI-Powered Learning Analytics: Bringing Students and Teachers Together", *International Journal of Educational Technology in Higher Education*, 2025.

The research employs participatory design methodology, conducting a series of workshops with students and teachers from higher education institutions in which participants collaboratively define their analytics needs, evaluate prototype system designs, and provide iterative feedback on analytics visualizations, interaction patterns, and decision-support features. Through thematic analysis of workshop data, the authors identify a set of design

principles that characterize analytics systems perceived as genuinely useful and trustworthy by educational stakeholders, including interpretability of AI-generated insights, transparency of data sources and algorithms, relevance of analytics to specific instructional and learning contexts, and the preservation of human agency in acting on analytics recommendations.

III. EXISTING SYSTEM

Existing academic performance management systems in educational institutions largely depend on manual processes for data collection, performance evaluation, and student communication. Faculty members are required to manually enter marks into spreadsheets, generate individual performance reports through time-consuming calculations, and communicate results to students via email or physical notice boards. These processes are inherently error-prone, time-intensive, and fail to scale effectively with increasing student cohort sizes.

Current systems lack the analytical depth required to provide subject-wise, topic-wise, and Course Outcome-based performance insights. Performance evaluation is typically conducted at the end of a semester, making it impossible to identify at-risk students early enough for meaningful intervention. The absence of predictive capabilities means that faculty are only able to react to academic failure after it has occurred, rather than preventing it proactively. Existing platforms do not integrate AI-powered student support tools such as conversational chatbots or personalized study plan.

IV. PROPOSED SYSTEM

The AI Academic Performance Intelligence System proposes a comprehensive, intelligent platform that transforms traditional academic performance management through the integration of machine learning, large language model capabilities, and automated communication workflows. The system operates through three primary role-based modules — Admin, Staff, and Student — each designed to address the specific needs of its user group within a unified architectural framework.

The core predictive component of the system employs a Random Forest Classifier trained on historical student academic data including internal assessment marks, CO-wise performance scores, subject-level results, and attendance records. The model predicts student pass or fail outcomes with high confidence, enabling staff to identify at-risk students early and implement targeted academic interventions. The CO-based analysis module maps each examination question to a specific Course Outcome, enabling granular analysis of learning achievement at the topic and outcome level rather than relying solely on aggregate scores.

The AI Student Support module integrates two key intelligent components: a RAG-based chatbot powered by LangChain that provides contextually accurate responses to academic queries by retrieving relevant institutional data, and an LLM-based study planner using the Grok API that generates personalized daily and weekly study schedules based on student-specific weak areas, available study time, and examination timelines. The study planner provides learning resources including reference materials and video links alongside each scheduled task, enabling students to follow structured, self-directed improvement plans. Automated email notifications are managed through an n8n workflow that is triggered upon staff mark uploads, ensuring that students receive reports without manual communication.

V. IMPLEMENTATION

A) NEXT.JS (FRONTEND):

Next.js was used to develop the web-based frontend of the AI Academic Performance Intelligence System, providing role-differentiated dashboards for Administrators, Staff, and Students. Its server-side rendering and static site generation capabilities ensure fast page loads and optimal performance. The component-based architecture enables modular development of complex UI elements including interactive analytics charts, mark upload interfaces, and AI study plan visualization components.

B) DJANGO & FASTAPI (BACKEND):

Django serves as the primary backend framework managing user authentication, role-based access control, database interactions, and administrative operations. FastAPI handles high-performance asynchronous API endpoints for AI model inference, mark processing, and real-time data operations. The dual-backend architecture separates concerns effectively — Django manages structured data operations while FastAPI handles computationally intensive AI and ML service calls with minimal latency.

C) POSTGRESQL (DATABASE):

PostgreSQL serves as the primary relational database, storing all structured academic data including student records, mark sheets, CO mappings, examination schedules, and performance analytics. Its support for complex relational queries enables efficient subject-wise, topic-wise, and CO-based performance aggregations that power the analytics dashboards.

D) RANDOM FOREST CLASSIFIER (MACHINE LEARNING):

The Random Forest Classifier is implemented using the scikit-learn library in Python and trained on historical student performance data. The model takes CO-wise marks, subject-level scores, internal assessment results, and attendance data as input features and outputs a pass or fail prediction with an associated confidence score. The model's feature importance analysis enables the system to highlight the academic factors most strongly associated with student risk, providing staff with interpretable insights beyond the binary prediction outcome.

E) RAG CHATBOT (LANGCHAIN):

The AI Chatbot is built using LangChain's Retrieval-Augmented Generation framework, which retrieves relevant documents from a vector store containing institutional academic data before generating responses. This architecture ensures that chatbot responses are grounded in verified, institution-specific information rather than generic pre-trained knowledge. Students can query the chatbot about

subject concepts, performance insights, and academic guidelines, receiving accurate and contextually relevant responses in real time.

F) LLM STUDY PLANNER :

The AI Study Planner integrates the LLM through its API to generate structured, personalized study schedules. Students provide inputs including their subject, available study days, and daily time commitment, and the LLM generates a day-wise study plan with topic breakdown, prioritization based on weak areas, and recommended learning resources.

G) n8n (AUTOMATION WORKFLOW):

n8n is used to automate the email notification workflow that is triggered upon successful mark uploads by staff. The workflow processes the uploaded marks, retrieves student contact information from PostgreSQL, generates personalized performance summary emails, and dispatches them automatically without requiring manual intervention.

VI. FLOW DIAGRAM

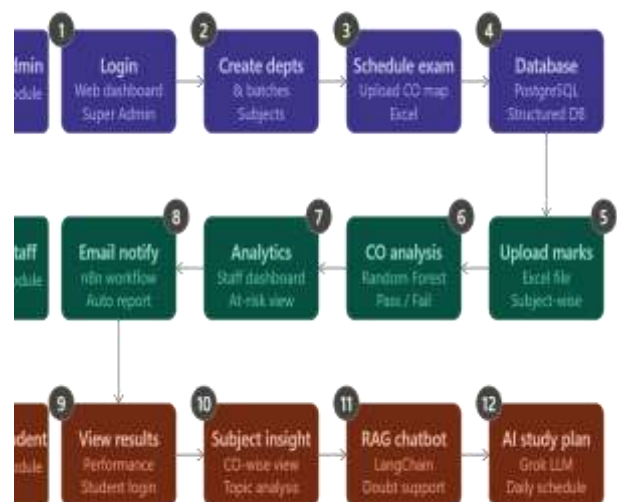


Fig. 1 Academic Performance Intelligence System

The system follows a three-tier workflow where the Admin manages institutional setup, Staff uploads marks and performs ML-based performance prediction, and

Students access results, AI chatbot support, and personalized study plans through an automated, data-driven academic intelligence platform.

VII. CONCLUSION AND FUTURE WORK

Future enhancements to the system include the integration of advanced deep learning models such as LSTM networks for sequential academic trend prediction, the development of a mobile application for student access on portable devices, the incorporation of multilingual support to serve linguistically diverse student populations, and the implementation of blockchain-based academic record verification for tamper-proof credential management.

Additionally, expanding the analytics module to include real-time performance monitoring during examinations and integrating adaptive quiz generation based on student-specific weak areas represent promising directions for continued system development.

VI. REFERENCES

- [1] Amrieh, E. A., Hamtini, T., & Aljarah, I. "Mining Educational Data to Predict Student's academic Performance using Ensemble Methods", International Journal of Database Theory and Application, 2016.
- [2] Yadav, S. K., Bharadwaj, B., & Pal, S. "Data Mining Applications: A Comparative Study for Predicting Student's Performance", International Journal of Innovative Technology and Creative Engineering, 2012.
- [3] Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. "Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?", Journal of Machine Learning Research, 2014.
- [4] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. "Systematic Review of Research on Artificial Intelligence Applications in Higher Education", International Journal of

Educational Technology in Higher Education, 2019.

[5] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. "Systematic Review of Research on Artificial Intelligence Applications in Higher Education", International Journal of Educational Technology in Higher Education, 2019.

[5] Alfredo, R., Milesi, M., Echeverria, V., & Martinez-Maldonado, R. "Co-Designing AI-Powered Learning Analytics: Bringing Students and Teachers Together", International Journal of Educational Technology in Higher Education, 2025.