

AI-Based Smart College ERP and Student Performance Prediction System

Prof. N. Ansari¹, Ana Khalid², bhavesh sahare³, Vaishnavi Pande⁴

^{1,2,3,4} *Department of Computer Science & Engineering, Nagpur Institute of Technology, Nagpur, India*

Abstract—College campuses produce a lot of administrative data every day. I see that this data covers admissions, student attendance, exam results, fee payments, class schedules and logs, from the learning management system. I also see that most colleges keep this data in files or software that do not talk to each other. This paper talks about a Smart College ERP system. Smart College ERP brings all functions together in one platform. Adds a Student Performance Prediction Engine that uses machine learning. Smart College ERP connects data from all areas. Uses it to help people make better choices. Smart College ERP has two parts. The first part is the ERP core. It handles admissions, attendance, exams, fees⁴ and timetables. The second part is the prediction layer. The prediction layer gathers data about student behavior cleans it and prepares it. It then creates features that help the models learn. Smart College ERP trains models such as Logistic Regression, Random Forest, Support Vector Machines, Gradient Boosting and Artificial Neural Networks. These models predict which students may be at risk who might drop out and what their future academic performance could be. We review research in educational data mining, learning analytics and AI-enhanced ERP systems. This review explains why we chose the design and technology for Smart College ERP. We also discuss challenges, like explainability, data privacy, fairness and accessibility. Any real-world use of Smart College ERP must consider these to keep it trustworthy and fair. We give examples of how Smart College ERP could help students, teachers and college staff. In the end we conclude that Smart College ERP does not replace educators or administrators. Instead Smart College ERP supports them. Smart College ERP acts as a tool to help institutions spot problems early and use resources wisely. Smart College ERP helps the right students get help when they need it.

Index Terms— I want you to pay attention to Artificial Intelligence, Enterprise Resource Planning (ERP) Student Performance Prediction, Educational Data Mining, Learning Analytics, Machine Learning, Early

Warning System, Higher Education Management and Explainable AI. I think Artificial Intelligence is exciting and I think Machine Learning is powerful. I find these topics, for research and practice.

I. INTRODUCTION

Higher-education institutions are using systems more and more to handle their daily tasks. ERP systems were created for colleges and universities to replace paper records with parts for admissions, attendance, exams, fees and schedules. These systems have made keeping records and cut down on work for staff. Most old ERPs are still not very active. They keep data. Do tasks automatically but they don't look at the data to see problems before they happen. Decisions about which students need help how to use resources or which groups might leave school are still made by people usually after a problem is already clear like failing classes, poor attendance or students leaving. Artificial Intelligence, machine learning can make this passive way of working more active. Educational Data Mining and Learning Analytics show that patterns in attendance test scores using learning tools and personal details can predict a student's future performance quite well. When this ability to predict is added to the ERP that a school already uses the system can not record what has happened but also show what might happen next. This lets teachers, advisors and school leaders act before it is too late.

This paper introduces an AI-Based Smart College ERP and Student Performance Prediction System. It brings these two areas together in one system. The system keeps the parts of a traditional college ERP but adds an AI Prediction Engine. This engine uses data already in the ERP like attendance, grades, when fees are paid and how well students follow the schedule along with data from learning platforms. It gives risk scores chances of dropping out and predictions about

performance. These are shown on dashboards for students, teachers, school staff and leaders.

The rest of this paper is set out like this. Section II looks at research on how ERPs have changed in education educational data mining, machine learning for predicting performance early-warning systems for dropping out and AI systems for campuses. It also shows the gap this work fills. Section III explains the system design and how it works, including the machine learning steps and the tools used. Section IV compares the AI methods that work best for the prediction part. Section V gives examples of how the system can be used. Section VI talks, about ethics, privacy and real-world issues. Section VII looks at the picture and limits of the system. Section VIII ends the paper.

II. LITERATURE REVIEW

2.1 Evolution of ERP Systems in Higher Education

I think Traditional education ERPs started as digitized versions of paper record-keeping. Traditional education ERPs were student information systems that stored admission details, attendance and grades in databases [1]. Over time Traditional education ERPs expanded into platforms that cover finance, human resources, library and hostel management. Traditional education ERPs are now often delivered as cloud-based subscription-priced services of on-premise installations [1]. Recent industry and enterprise-architecture literature argues that the next stage of this evolution involves connecting siloed systems. This includes the Student Information System (SIS) the Learning Management System (LMS) the ERP and even Internet-of-Things (IoT). sensors in classrooms. The connection is made through a shared data layer so that AI models can operate on a real-time view of the student journey instead of, on fragmented batch-updated exports [2].

2.2 Educational Data Mining and Learning Analytics

Educational Data Mining or EDM uses data-mining methods on information gathered from learning environments. The main aim is to find patterns that can help make teaching and learning better [3]. Much of the research in EDM focuses on predicting how well students will do academically. This is done by looking at performance and behavior. The reason for this focus is simple: if we can spot students who are having trouble we can step in quickly with support [3].

Learning Analytics or LA is another field that works closely with EDM. It also looks at sets of data created by online learning systems. Its goal is to predict student results and adjust the content being taught based on what the data shows [4]. Both EDM and LA share a belief. That belief is central to this paper: institutions already collect amounts of data. If that data is properly linked and modeled it holds clues, about how students will perform in the future.

2.3 Machine Learning Approaches for Performance Prediction

The classic benchmark dataset in this space is the UCI Student Performance dataset, introduced by Cortez and Silva. The UCI Student Performance dataset records school-related attributes for secondary-school students in Mathematics and Portuguese-language courses. The UCI Student Performance dataset has been used extensively to test classification and regression approaches to grade prediction [5]. A separate used benchmark is the Open University Learning Analytics Dataset, known as OULAD compiled by Kuzilek et al. The Open University Learning Analytics Dataset captures virtual-learning-environment interactions, assessment submissions and final outcomes for, than 30,000 students across multiple course presentations [6].

Comparative studies consistently show that ensemble methods and tree-based methods beat linear models on these prediction tasks. A recent IEEE study that compared Support Vector Machines, Random Forest, k-Nearest Neighbours, Artificial Neural Networks and Logistic Regression for web-based performance prediction found that Random Forest and Artificial Neural Networks produced the accurate and robust predictions. A model that combined Decision Trees, Random Forest Support Vector Machines and Artificial Neural Networks when tested across the UCI, OULAD and NELS:88 datasets reached peak accuracies of 98.8 percent, 96.9 percent and 97.5 percent which shows that ensemble approaches generalize across different institutional data sources. Gradient-boosting methods have also performed strongly: an XGBoost-based model that was trained and validated on a fifty thousand-record dataset achieved an R-squared of 0.91. Cut mean-squared error by fifteen percent compared to traditional approaches. A comparative study of dropout and success prediction discovered that boosting algorithms such as

LightGBM and CatBoost tuned with the Optuna framework, outperformed classifiers. A comprehensive PRISMA-guided review of seventy-two studies drew a conclusion stating that tree-based models and neural networks dominate the field and that explainable-AI techniques are becoming a necessary complement, to raw predictive accuracy

2.4 Dropout Prediction and Early-Warning Systems

A specific and important use of performance prediction is the early-warning system or DEWS. This system is built to identify students who're likely to leave an institution before finishing their studies. The goal is to give retention staff time to step in and help these students stay on track. One of the most influential examples is the Wisconsin statewide DEWS, which Knowles describes as a flexible and modular platform. It can adjust to data sources and different outcome measures making it adaptable to changing needs.

Another example comes from Lee and Chung who created a machine-learning-based DEWS. Lee and Chung focused on boosting prediction accuracy for students who're, at risk of dropping out. At the time Bañeres and colleagues developed a system specifically for online higher-education environments, where finding at-risk students can be more difficult because of limited in-person interaction..

Recently researchers used a Random Forest model trained on administrative data from 1,385 bachelor's degree students. The model achieved an area under the curve or AUC of 0.961 and an accuracy of 88.6%. These results were better than those of models like logistic regression, gradient boosting and neural networks all tested on the same dataset

Martins and others conducted a study, at the Polytechnic Institute of Portalegre. They used machine-learning classification models to find students who were likely to fail to complete their degrees on time.. Their work supports the idea that predictive models can help institutions take steps.

Across all these studies a clear pattern emerges: models for dropout work best when they support human decision-making. They should not be used as standalone systems that make decisions. Instead they should be combined with assessments, from staff, who can interpret the results in context and act with care and judgment. This balance helps ensure that

interventions are both accurate and fair.

2.5 AI-Enabled ERP and Smart Campus Platforms

As on the systems side, industry and applied literature describes how AI is beginning to be added to ERP and campus-management platforms. AI-enabled ERPs are reported to support analytics for dropout risk, automated exam evaluation, financial forecasting of fee revenue and conversational digital assistants that resolve routine student queries without staff involvement [17]. Enterprise-architecture analyses of AI-powered campus operations emphasize that the central challenge is not any AI model but the integration of SIS, LMS, ERP and facilities data behind a shared data layer with governed real-time pipelines since fragmented silos and legacy batch-reporting cycles otherwise prevent AI systems from acting on timely information [2]. Related work on campus assistants describes architectures composed of data-retrieval core-framework intelligent-services and communication layers that integrate learning-platform data, with chatbot-based interaction to support both students and instructors [18].

2.6 Research Gap

Two threads of research run largely in parallel. The literature on machine-learning for student performance and dropout prediction is extensive and methodologically mature. It is usually tested on separate research data sets instead of on the real-time data that flows through an institution's operational ERP. On the hand applied and industry literature on AI-enabled ERP and smart-campus platforms explains how the integration architecture works and what administrative benefits it offers but it rarely shows the details of the prediction pipeline, the model choice, the evaluation methodology or the explainability and governance safeguards that peer-reviewed EDM research stresses. We address that gap by proposing an architecture in which a validated machine-learning prediction pipeline is embedded directly inside the operational modules of a college ERP with explainability, privacy and fairness treated as first-class design requirements rather, than afterthoughts.

III. PROPOSED SYSTEM AND METHODOLOGY

3.1 System Overview

The proposed system is built with five layers. The first layer is the User Layer. This layer includes the four

main stakeholder groups: students, faculty, administrators and parents/management. The second layer is the Presentation Layer. It includes a web portal and a mobile application. The third layer is the Application Layer. It has the ERP core modules and the AI Student Performance Prediction Engine. The Figure 1 illustrates this architecture.

fourth layer is the Integration/Middleware Layer. It offers an API gateway and a notification service. The fifth layer is the Data Layer. It consists of a database, a data warehouse for behavioural and log data and cloud storage, for documents and backups.

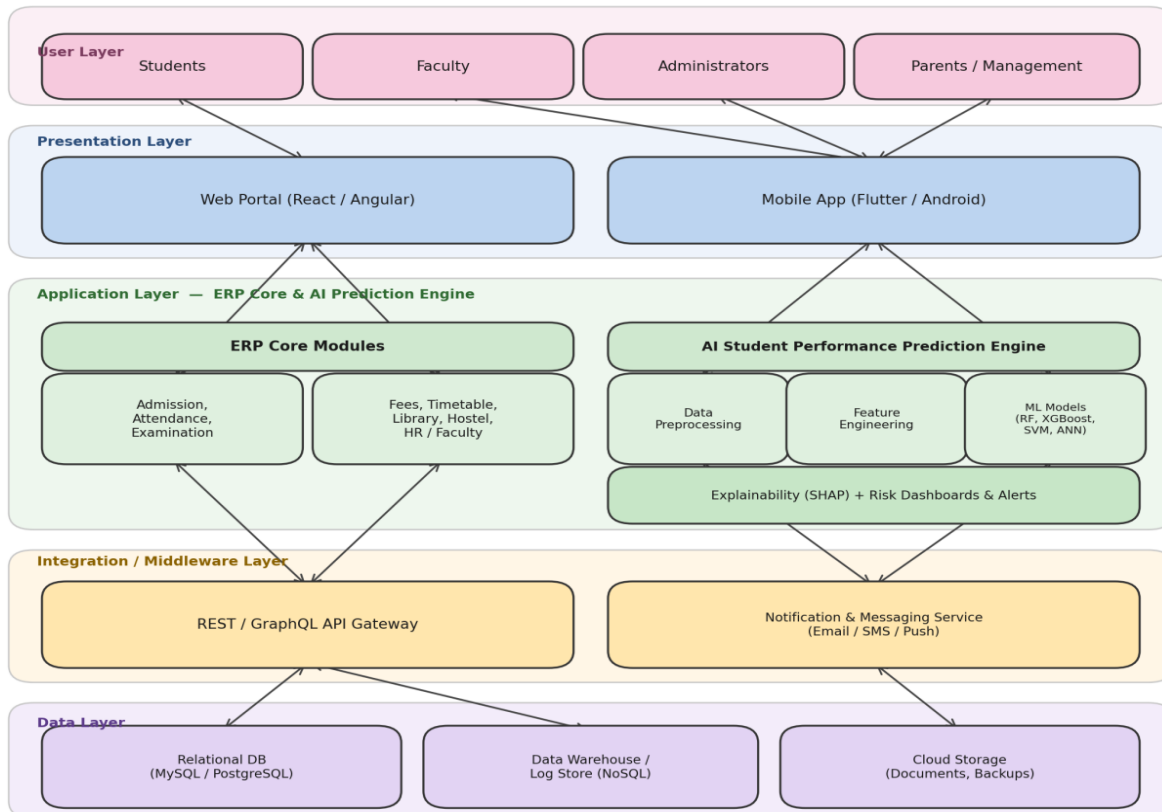


Figure 1. Layered architecture of the proposed AI-Based Smart College ERP and Student Performance Prediction System.

3.2 Core ERP Modules

The ERP core keeps the administrative functions you'd expect from a college management system. It provides the data that the prediction engine relies on.

Enrollment. This module captures records, entrance scores, past academic history and enrollment status.

Attendance Management. It records period-wise attendance using biometric, RFID or manual entry. Data is grouped by subject and by term.

Result Processing. This stores assessment marks, assignment scores and results from semester exams.

Fee & Finance. It tracks the fee structure, payment history, outstanding dues and scholarship allocations.

Resource Scheduling. This manages class timetables, faculty workloads and room assignments.

Library, Hostel and HR/Faculty modules. These handle the supporting resources, like library access, hostel accommodations and staff records.

Communication/Notification. This sends circulars, alerts and messages to students, faculty and parents.

3.3 AI Student Performance Prediction Engine

The ERP core keeps the functions you would expect from a college management system. It gives the data that the prediction engine depends on.

Enrollment. Enrollment captures records, entrance scores, past history and enrollment status. It helps us know who is in school.

Attendance Management. Attendance Management records period-wise attendance using biometric, RFID or manual entry. Attendance Management groups data by subject and by term. It helps us see who is in class.

Result Processing. Result Processing stores assessment marks, assignment scores and results from semester exams. Result Processing lets us see student progress.

Finance. Fee & Finance tracks the fee structure, payment history, outstanding dues and scholarship allocations. It keeps our records clear.

Resource Scheduling. Resource Scheduling manages class timetables, faculty workloads and room assignments. It helps teachers and students find the room at the right time.

Library, Hostel and HR/Faculty modules. Library, Hostel and HR/Faculty modules handle the supporting resources, like library access, hostel accommodations and staff records. They keep our campus life organized.

Communication/Notification.

Communication/Notification sends circulars, alerts and messages to students, faculty and parents. It keeps everyone in the loop.

3.4 Process / Data Flow

The end-to-end workflow goes through five stages: (1) Data Ingestion. ERP transactions. If connected LMS logs are collected all the time or in scheduled groups; (2) Preprocessing. The data that comes in is cleaned, turned into numbers and put into the feature store; (3) Prediction & Classification. The model that has been trained gives a score to each student who is still active and puts them into a risk group; (4) Dashboard Visualization & Alerts. Risk scores, charts that show trends and explanations from SHAP are shown on dashboards that are set up for different roles and alerts are sent to mentors and academic counsellors through the system that sends messages; and (5) Feedback & Intervention Tracking. The results from actions taken like mentoring meetings, remedial classes are written

back, into the system, which lets the model be trained again with the new data and lets people see how the model is working in real life.

3.5 Tools and Technologies

Layer | Representative Technologies

Frontend / Presentation React.js, Angular for web; Flutter, native Android/iOS for mobile

Backend / API | Django REST Framework, Node.js Express, Spring Boot; REST GraphQL API gateway

Machine Learning | Python, Scikit-learn, XGBoost LightGBM TensorFlow PyTorch, Pandas, NumPy, SHAP

Databases | MySQL, PostgreSQL for transactional ERP data; MongoDB, data-lake store for behavioural/log data

Visualization | Power BI, Plotly, Dash, Chart.js for dashboards

Cloud & DevOps | AWS, Azure, Firebase for hosting, storage CI/CD pipelines

Reference Datasets | UCI Student Performance dataset, OULAD

I think the system should be evaluated on prediction accuracy and F1-score of the risk classifier using held-out data. The system also needs to check usability of dashboards for faculty and mentors for example through a System Usability Scale survey. The system should measure response latency, for prediction and alert generation. The system must show scalability when many departments use it at the time. Finally the system must comply with privacy regulation when handling data as discussed in Section VI.

IV. AI TECHNIQUES AND COMPARATIVE ANALYSIS

The following table I summarize the machine-learning techniques commonly mentioned in the literature reviewed in Section II along with their usual role inside the proposed prediction engine, their known advantages and their degree of clarity. An important practical factor because predictions in this system help make choices, about real students..

TABLE I. A COMPARISON OF MACHINE LEARNING TECHNIQUES, FOR STUDENT PERFORMANCE PREDICTION

Algorithm	Typical Role in the Engine	Reported Strength (from literature)	Interpretability
Logistic Regression	Baseline pass/fail or at-risk classifier	Fast, stable baseline; easy to audit [7]	High
Decision Tree	Rule-based risk flagging	Human-readable decision rules [8]	High
Random Forest	Primary risk/dropout classifier	Consistently among the top performers; AUC 0.961 for dropout prediction [15]	Medium (via feature importance)
Support Vector Machine (SVM)	Grade-band classification	Strong on smaller, well-separated feature sets [7]	Low
XGBoost / LightGBM / CatBoost	High-accuracy performance/dropout prediction	R^2 of 0.91 on large datasets [9]; outperformed traditional classifiers [10]	Medium (with SHAP)
Artificial Neural Network (ANN)	Behavioural/engagement pattern modelling	Captures non-linear interactions; strong on OULAD-scale data [7], [8]	Low (requires XAI layer)
Hybrid / Ensemble (RF+SVM+ANN, etc.)	Production default model	Outperformed individual models across UCI, OULAD and NELS:88 (up to 98.8% accuracy) [8]	Low–Medium (with SHAP)

Two conclusions follow from this comparison. First no single algorithm is better in every setting. The correct algorithm depends on data volume on the balance between accuracy and interpretability that each institution needs and, on whether the main goal is a risk class or a continuous grade estimate. Second the literature shows that ensemble and hybrid arrangements usually give the raw accuracy.. This accuracy gain comes with a loss of interpretability. That loss is why the proposed engine combines its production ensemble model with a SHAP-based explanation layer, instead of offering only a plain probability score to end users.

V. ILLUSTRATIVE APPLICATION SCENARIOS

The following scenarios explain how the proposed system will be used in practice. They show example uses of the architecture described in Section III, based on patterns found in the literature reviewed in Section II and not on results from an institutional deployment.

5.1 Early Identification of At-Risk Students

When a students attendance drops below a set threshold and the students internal assessment scores go down during tests the prediction engine increases the students risk score and sends an alert to the assigned faculty mentor along with a brief SHAP-based explanation of the factors that contributed. This follows the intervention logic used in dropout early-warning systems reported in the literature [13] [14] where earlier detection is repeatedly identified as the factor that enables effective intervention.

5.2 Dropout Risk Forecasting to Support Counselling

At the beginning of each semester the engine creates a ranked list of students based on dropout probability using a Random Forest-based model that prior work has shown to deliver discriminative performance (AUC of 0.961), on administrative academic data [15]. Academic counsellors use this list with the explanations to prioritize outreach rather than relying solely on informal referrals.

5.3 Automated Fee and Resource Forecasting

Because the Fee & Finance module already tracks payment history the same data pipeline can support short-term forecasting of fee-revenue inflows. Highlight cohorts likely to require deferred-payment support consistent with the financial-forecasting capability described in applied AI-ERP literature [17].

5.4 Faculty Workload and Timetable Optimization

Attendance and scheduling data feed a secondary analytics view that helps administrators spot sections with high absenteeism clustered around specific time slots enabling more informed timetable adjustments.

5.5 Personalized Academic Advising Dashboard

Each mentors dashboard shows each mentors assigned students ranked by risk level with drill-down into attendance trend, assessment trend and the specific SHAP factors behind each score. Giving mentor a starting point for a conversation of a final verdict, on the student.

VI. ETHICAL, PRIVACY AND PRACTICAL CONSIDERATIONS

6.1 Data Privacy and Protection

The system manages personal and academic information, such as attendance, grades and financial records. Institutions that deploy such a system must align data handling with the laws that apply in their jurisdiction – for instance the General Data Protection Regulation (GDPR) in the European Union the Family Educational Rights and Privacy Act (FERPA) in the United States or Indias Digital Personal Data Protection (DPDP) Act. In line with the principle of data minimization highlighted in studies, on education governance and datafication [22] the proposed architecture is built to collect only the fields required for the specific predictive purpose to encrypt data when it is stored or transmitted and to use role-based access control so that for example a faculty mentor sees risk scores but does not see full financial records.

6.2 Algorithmic Bias and Fairness

I see that prediction models learn from institutional data.

Because of that prediction models can. Amplify existing biases.

For example prediction models may systematically under-predict the potential of students from

backgrounds when those backgrounds are linked to past outcomes.

In the literature people keep recommending bias auditing, across demographic subgroups.

The literature also stresses reporting of any disparities found.

This proposed engine follows those recommendations.

It includes a model-evaluation step.

That step happens every time not once at deployment..

6.3 Transparency and Explainability

A risk score that has no explanation can be seen as a verdict or as an opaque and untrusted output. The SHAP-based explanation layer that is described in Section 3.3 is added specifically to solve this problem. It follows the growing trend in EDM research to use explainable AI because explainable AI is needed before we can responsibly use predictive models, for academic decision-making [11] [21].

6.4 Guarding Against Over-Reliance

Predictive scores are designed to support, not replace the judgment of staff. In line with the caution raised by dropout prediction researchers that machine-learning outputs work best when combined with human evaluation the dashboards present risk scores as one input to a mentoring conversation. The system explicitly avoids automated actions such, as auto-generated warning letters that would remove a human decision-maker from consequential outcomes.

6.5 Accessibility

A system that is based on collecting data assumes that students have consistent access to devices and internet connections and that schools have reliable access to computer systems. Assumptions that are not always true for all schools or student groups. Because of this the system is designed so that the main functions of the ERP system can still work with internet and basic devices. It also makes sure that if some extra information from the learning management system is not available the engine can still make a prediction using the main ERP data..

6.6 Cybersecurity

Because the system brings school data together on one platform it also brings security risks together. Standard ways to protect data. Like encryption, access based on roles, tracking of activity and regular checks, for weaknesses. Are required parts of the part of the

system described in Section 3.1. They are not features added later.

VII. DISCUSSION

The architecture proposed in this paper addresses the gap identified in Section 2.6 by embedding a validated literature-grounded prediction pipeline inside the operational modules of a college ERP. The expected benefits fall into three categories: efficiency, through automation of routine record-keeping and reporting; earlier intervention, since risk scores are generated continuously from live ERP data rather than being reconstructed after the fact from disconnected exports; and more informed institutional planning since aggregated predictions can inform decisions about resource allocation, mentoring capacity and academic support programs.

The approach also has limitations. Prediction quality depends heavily on data completeness and consistency across departments and institutions with inconsistently recorded attendance and assessment data will see correspondingly weaker predictions—a cold-start problem that is most acute for newly enrolled students with little historical data. Models must also be retrained periodically to avoid concept drift as curricula, grading practices or student populations change over time. Finally integrating a prediction engine with legacy ERP systems that were not designed with an API-first architecture can require engineering effort, which is a practical adoption barrier for smaller institutions.

Future extensions of this work could incorporate data such as IoT-based smart-classroom sensors and biometric attendance as suggested in enterprise-architecture literature on AI-powered campus operations [2]; explore federated-learning approaches that allow models to be trained across institutions without centralizing raw student data addressing privacy concerns, at the source; and extend the notification layer with natural-language interfaces so that students and mentors can query their own risk profile conversationally rather than only through a dashboard.

VIII. CONCLUSION

This paper presented a Smart College ERP and Student Performance Prediction System. The system

combines the parts of a normal college ERP with a machine-learning prediction engine that can forecast students who are at risk the probability that a student will drop out and a student's expected academic performance. Drawing on a review of educational data mining, learning analytics, dropout early warning and AI-enabled campus-management literature the paper proposed a five-layer architecture. It described a five-stage prediction pipeline that is built around machine-learning models and uses SHAP-based explainability. The paper also listed privacy, fairness, transparency and accessibility safeguards that any real deployment of AI-Based Smart College ERP must include. The main conclusion is that AI-Based Smart College ERP is most valuable when it is used as a decision-support layer that brings explainable signals to human mentors, counsellors and administrators. It does not act as a system that makes decisions about students by itself. When used in this manner AI-Based Smart College ERP can help institutions shift from record-keeping to proactive data-informed academic support while keeping the final judgement about any individual student, in the hands of educators who know those students.

REFERENCES

- [1] Academia ERP, "Integration of AI in Education: Transforming with Academia ERP/SIS," 2024. [Online]. Available: <https://www.academiaerp.com/blog/integration-of-ai-in-education-transforming-with-academia-erp-sis/>
- [2] Ideausher, "Enterprise Architecture for AI-Powered Campus Operations," 2026. [Online]. Available: <https://ideausher.com/blog/ai-campus-operations-architecture/>
- [3] M. I. Uddin et al., "Role of convolutional features and machine learning for predicting student academic performance from MOODLE data," educational data mining study, 2023. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10631652/>
- [4] PMC/NCBI, "Enhancing e-learning through AI: advanced techniques for optimizing student performance," 2025. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11784796/>

- [5] P. Cortez and A. Silva, "Using data mining to predict secondary school student performance," in Proc. 5th Annual Future Business Technology Conf. (FUBUTEC), 2008, pp. 5–12.
- [6] J. Kuzilek, M. Hlosta, and Z. Zdrahal, "Open University Learning Analytics dataset," Scientific Data, vol. 4, art. 170171, 2017.
- [7] IEEE Xplore, "Student Performance Prediction System Using Machine Learning," IEEE Conference Publication, 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/11118334/>
- [8] Nature Scientific Reports, "Artificial intelligence in student management systems to enhance academic performance monitoring and intervention," 2025. [Online]. Available: <https://www.nature.com/articles/s41598-025-19159-4>
- [9] Frontiers in Education, "Machine learning models for academic performance prediction: interpretability and application in educational decision-making," 2025. [Online]. Available: <https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1632315/full>
- [10] Discover Artificial Intelligence, "Supervised machine learning algorithms for predicting student dropout and academic success: a comparative study," Springer Nature, 2024. [Online]. Available: <https://link.springer.com/article/10.1007/s44163-023-00079-z>
- [11] ScienceDirect, "Predicting student performance: A comprehensive review of machine learning, deep learning, and explainable AI approaches," 2026. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666920X26000093>
- [12] J. E. Knowles, "Of needles and haystacks: Building an accurate statewide dropout early warning system in Wisconsin," Journal of Educational Data Mining, vol. 7, no. 3, pp. 18–67, 2015.
- [13] S. Lee and J. Y. Chung, "The machine learning-based dropout early warning system for improving the performance of dropout prediction," Applied Sciences, vol. 9, no. 15, 2019.
- [14] D. Bañeres, M. E. Rodríguez, A. E. Guerrero-Roldán, and A. Karadeniz, "An early warning system to detect at-risk students in online higher education," Applied Sciences, vol. 10, no. 13, 2020.
- [15] Discover Education, "Predicting student dropout in higher education using Random Forest and survival analysis," Springer Nature, 2026. [Online]. Available: <https://link.springer.com/article/10.1007/s44217-026-01710-8>
- [16] M. V. Martins, D. Tolledo, J. Machado, and L. M. Baptista, "Early prediction of student's performance in higher education: a case study," Instituto Politécnico de Portalegre, 2021.
- [17] Vmedulife, "AI-Based ERP for Colleges | The Future of Smarter Higher Education," 2025. [Online]. Available: <https://vmedulife.com/blog/erp-lms/smarter-institutions-ahead-why-every-campus-needs-ai-based-erp-for-colleges/>
- [18] arXiv, "Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education," arXiv:2309.10892, 2023.
- [19] arXiv, "AI-Driven Early Warning Systems for Student Success: Discovering Static Feature Dominance in Temporal Prediction Models," arXiv:2512.12493, 2025.
- [20] arXiv, "Machine Learning-Driven Student Performance Prediction for Enhancing Tiered Instruction," arXiv:2502.03143, 2025.
- [21] arXiv, "A Human-Grounded Evaluation of SHAP for Alert Processing," arXiv:1907.03324, 2019.
- [22] B. Williamson and N. Piattoeva, "Education governance and datafication," Learning, Media and Technology, vol. 46, no. 1, pp. 1–14, 2021.