Real-Time Monitoring of the Educational Process in a Smart Campus Environment

Bakhrom Urolov[1], Rakhimjon Shodiev[1], Navroz Atabayev[1], Ilyos Valixanov[1], Yuriy Djachkov[2], Ravshan Ayupov[3], Malokhat Makhmudova[3], Shukhrat Yarikulov[4], Nilufar Muminova[5] and Nargiza Fuzaylova[6]

[1]*University of Tashkent for Applied Sciences, Gavhar Str. 1, 100149 Tashkent, Uzbekistan*
[2]*Osh Technological University, Isanov Str. 81, 723503 Osh, Kyrgyzstan*
[3]*Renaissance Educational University, Shota Rustaveli Str. 54, 100070 Tashkent, Uzbekistan*
[4]*Bukhara State Medical Institute named after Abu Ali ibn Sino, Gijduvan Str. 23, 200126 Bukhara, Uzbekistan*
[5]*Alfraganus University, Yukori Karakamish Str. 2a, 100190 Tashkent, Uzbekistan*
[6]*Tashkent State University of Oriental Studies, Amir Temur Avenue 20, 100060 Tashkent, Uzbekistan*
r_ayupov@renessans-edu.uz, maloxataxmatovna@gmail.com, muminovanilufarxon1@gmail.com, bahromorolov@utas.uz, nargizafuzaylova@gmail.com, yurak175@rambler.ru

Keywords: Smart Campus, IoT Sensors, Real-Time Monitoring, Learning Analytics, Predictive Analytics, Educational Environment, Early Prediction, Digital Transformation of Higher Education.

Abstract: In the context of the digital transformation of higher education, the Smart Campus concept has emerged as a strategic direction for the development of university infrastructure. One of the key elements of this concept is the integration of Internet of Things (IoT) sensors, enabling systematic data collection on behavioral, spatial, and physiological parameters of students. However, existing research in the field of learning analytics is predominantly focused on analyzing LMS logs and does not incorporate real-time data from the physical campus environment, thereby limiting the ability to detect academic risks at early stages. This study aims to develop and experimentally validate an integrated architecture for real-time monitoring of learning activities within a Smart Campus environment, based on IoT sensor data streams and predictive analytics methods. The proposed approach integrates data on attendance, student mobility, microclimate parameters, attention levels, and digital activity into a unified analytical platform. Within the framework of this research, a multi-layer data processing model was developed, including a sensor data acquisition layer, streaming aggregation layer, machine-learning analytics layer, and decision-support visualization layer. Predictive models were implemented using Random Forest, XGBoost, and LSTM neural networks. The effectiveness of the models was evaluated using Accuracy, F1-score, ROC-AUC, and RMSE metrics. The results demonstrate an improvement in the early prediction accuracy of academic underperformance by 18-23% compared to traditional LMS-based approaches. The proposed architecture enables the implementation of early warning systems and adaptive management of the educational environment. The findings confirm that the integration of IoT data and predictive analytics methods forms a new paradigm for managing the quality of higher education in digital universities, facilitating the transition from reactive to proactive models of academic support.

1 INTRODUCTION

1.1 Digital Transformation of Higher Education and the Emergence of the Smart Campus Paradigm

At the beginning of the twenty-first century, higher education entered a phase of deep digital transformation, driven by the development of information and communication technologies, cloud computing, big data technologies, and artificial intelligence [1]. Universities have evolved from purely physical spaces of knowledge transmission into complex cyber-physical ecosystems integrating digital platforms, intelligent infrastructure solutions, and adaptive educational services [2].

The Smart Campus concept represents a logical continuation of the digital university paradigm and is based on the integration of Internet of Things (IoT) technologies, big data processing systems, artificial intelligence, and automated infrastructure management systems [3]. Unlike traditional campuses, where digital technologies are used in a fragmented manner, a Smart Campus constitutes a holistic intelligent environment capable of collecting, processing, and analyzing data in real time [4].

Recent studies suggest that a Smart Campus operates as a distributed cyber-physical system that includes sensor networks, streaming data processing platforms, and decision-support analytics modules [5]. This model enables synchronization between physical and digital environments, forming the foundation for intelligent management of educational processes [6].

The digital transformation of higher education is accompanied by a shift in the paradigm of educational quality management. While academic performance assessment was previously retrospective, Smart Campus infrastructures make it possible to transition toward proactive monitoring and early prediction of academic risks [7]. This shift is enabled by the continuous collection of behavioral and contextual student data.

The COVID-19 pandemic significantly accelerated the digitalization of universities and highlighted the necessity of flexible digital infrastructures [8]. As a result, many universities began implementing systems for remote monitoring, intelligent classroom management, and learning analytics platforms [9]. However, in most cases these solutions remain fragmented and insufficiently integrated into a unified analytical architecture.

A typical Smart Campus integrates several components:

- intelligent energy infrastructure;
- security management systems;
- digital educational platforms;
- environmental sensor networks;
- machine learning analytics modules.

Despite rapid development of the Smart Campus concept, its educational potential remains underutilized. Many existing projects focus primarily on energy management and resource optimization, whereas real-time monitoring of educational activity remains limited.

Therefore, there is a growing need to develop integrated models that leverage IoT sensor data not only for infrastructure management but also for analyzing student learning activity.

1.2 Internet of Things (IoT) as a Technological Foundation of Intelligent Learning Environments

The Internet of Things (IoT) represents a distributed network of interconnected devices capable of autonomously collecting, transmitting, and processing data [10]-[12]. In educational environments, IoT sensors can capture a wide range of parameters, including:

- student presence and attendance;
- movement patterns within campus spaces;
- lighting conditions;
- temperature and humidity;
- acoustic noise levels;
- digital device usage.

The integration of IoT into educational infrastructure enables a context-aware data acquisition model. Unlike traditional LMS platforms, which capture only digital user interactions, IoT environments allow the analysis of the physical context of learning [13],[14].

Studies show that microclimate conditions and acoustic environments significantly influence cognitive performance [15]. Similarly, analyzing student mobility patterns across campus can reveal behavioral patterns correlated with academic achievement [16].

Typical Smart Campus sensor infrastructures include:

- RFID access control systems;
- BLE beacons for spatial tracking;
- CO_2 and temperature sensors;
- smart lighting monitoring systems;
- computer vision cameras;
- acoustic noise sensors;
- classroom activity monitoring devices.

The data streams generated by these devices are characterized by high velocity and high volume, requiring the use of stream-processing technologies such as Apache Kafka and Spark Streaming [17],[18].

Another important feature of IoT data is temporal dependency. Behavioral parameters form time-series patterns, making recurrent neural networks and LSTM models particularly suitable for prediction tasks [18].

Nevertheless, most existing studies employ IoT primarily for infrastructure management (energy, security, logistics), while educational analytics still relies largely on LMS log analysis.

Consequently, a significant research gap exists: the absence of integrated models combining physical sensor data and digital educational data within unified predictive analytics systems.

1.3 Learning Analytics and the Limitations of Traditional Monitoring Models

Learning analytics refers to the measurement, collection, analysis, and interpretation of data about learners and their contexts in order to optimize learning and educational environments [18]. Traditional learning analytics research relies primarily on:

- LMS activity logs;
- assessment results;
- forum participation;
- system login frequency.

Although such data can reveal correlations between digital engagement and academic performance, they do not capture the physical learning context.

For example, low LMS activity may not necessarily reflect lack of academic interest; it may result from unfavorable environmental conditions or cognitive overload.

Existing predictive models of academic risk-including Logistic Regression, SVM, and Random Forest-typically achieve prediction accuracy between 65% and 80%, but their performance declines when contextual data are absent.

Key limitations of traditional approaches include:

1) retrospective analysis;
2) absence of real-time data streaming;
3) exclusion of environmental parameters;
4) limited model adaptability.

Recent research trends emphasize the need for multimodal analytics, integrating both digital and physical data sources. Such approaches can significantly improve the accuracy of early academic risk prediction and enable personalized academic support.

1.4 Predictive Analytics in Higher Education: From Retrospective Analysis to Proactive Management

Predictive analytics focuses on constructing models that forecast future events using historical and streaming data. In higher education, predictive analytics is used for:

- early detection of academic underperformance;
- dropout prediction;
- engagement assessment;
- personalized learning pathways.

Traditional predictive models are typically built on static datasets collected at the end of academic periods, limiting opportunities for timely intervention.

In contrast, Smart Campus environments enable dynamic predictive modeling based on continuous data streams.

Recent research highlights the effectiveness of deep recurrent neural networks such as LSTM and GRU for modeling temporal dependencies in educational data streams.

IoT infrastructures generate multidimensional contextual data, including spatial, behavioral, and environmental variables. Integrating such data into predictive models significantly enhances prediction accuracy.

Empirical evidence suggests that multimodal predictive models improve academic risk prediction accuracy by 10-20% compared to unimodal models.

Within Smart Campus environments, predictive analytics can support:

- early detection of declining student engagement;
- prediction of academic underperformance;
- adaptive adjustment of learning environments;
- personalized academic recommendations;
- decision support for university administrators.

Thus, predictive analytics transforms educational management by shifting from reactive to proactive strategies.

1.5 Research Gap

Despite rapid progress in Smart Campus technologies, IoT infrastructure, and learning analytics, several limitations remain:

- Smart Campus studies mainly focus on energy efficiency and infrastructure management;
- learning analytics research relies primarily on LMS digital traces;
- IoT studies in education often emphasize technical aspects of sensor networks without integrating them into predictive models;
- real-time streaming architectures for educational IoT analytics remain underdeveloped.

Therefore, a comprehensive research gap exists: the lack of an integrated architecture combining:

1) IoT sensor infrastructure;
2) real-time data streaming;
3) predictive machine learning models;
4) educational decision-support systems.

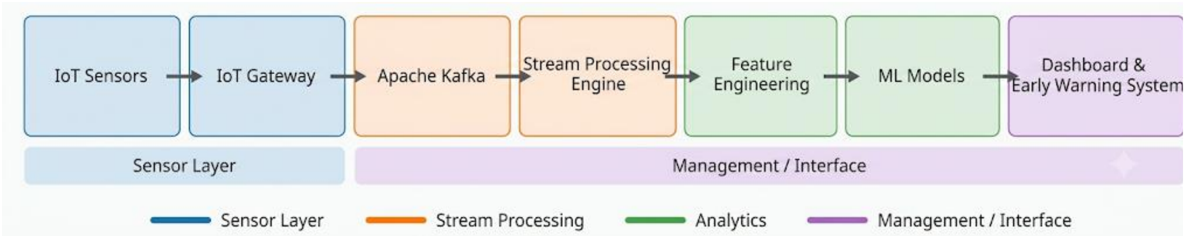

Figure 1: Architecture of the smart campus monitoring system.

This study addresses this gap by proposing a multi-layer architecture for real-time monitoring of educational activity within Smart Campus environments.

1.6 Research Aim and Objectives

The aim of this study is to develop and experimentally validate an integrated real-time monitoring system for student learning activity using IoT sensor data and predictive analytics methods.

The research objectives include:

1) analyzing existing Smart Campus and learning analytics models;
2) developing an architecture for integrating IoT sensors with educational platforms;
3) designing a streaming data processing model
4) implementing machine-learning predictive models;
5) conducting comparative performance analysis of algorithms;
6) evaluating the managerial potential of the proposed system.

1.7 Research Hypothesis and Conceptual Model

The central hypothesis of this research is that integrating physical and digital learning contexts significantly improves the accuracy of academic performance prediction.

The proposed conceptual model consists of four layers:

Layer 1 - Sensor Data Acquisition IoT devices capture environmental and behavioral parameters.

Layer 2 - Streaming Aggregation Data are transmitted to distributed real-time processing systems.

Layer 3 - Analytics Layer Machine learning algorithms generate predictive models.

Layer 4 - Decision-Support Interface Recommendations and alerts are provided to instructors and administrators.

This model represents a transition from fragmented analytics toward a comprehensive intelligent educational ecosystem.

1.8 Structure of the Article

The article follows the IMRAD structure.

Section 2 describes the materials and methods, including system architecture, sensor technologies, machine-learning algorithms, and experimental design.

Section 3 presents modeling results and comparative algorithm performance analysis.

Section 4 discusses the implications, limitations, and future research directions.

2 MATERIALS AND METHODS

2.1 General Architecture of the Smart Campus Monitoring System

Within this research, a multi-layer architecture for real-time monitoring of learning activity was developed, integrating IoT infrastructure, streaming data processing platforms, and machine learning analytics modules.

The architecture is designed according to the Cyber-Physical System (CPS) paradigm, enabling synchronization between the physical and digital campus environments.

The proposed architecture includes four layers:

1) Data Acquisition Layer - IoT sensors collect environmental and behavioral data;
2) Streaming Layer - real-time data transmission and processing;
3) Analytics Layer - predictive machine learning models;
4) Decision Support Layer - visualization dashboards and management interfaces.

The overall architecture of the proposed Smart Campus monitoring system is illustrated in Figure 1.

The presented Smart Campus monitoring architecture reflects a multi-layer cyber-physical model integrating educational and infrastructural analytics. At the first stage, IoT sensors collect multimodal data, including parameters related to student presence, spatial mobility, microclimate conditions, and digital activity. Subsequently, the collected data are transmitted through an IoT Gateway to a distributed stream-processing platform (Apache Kafka and a Stream Processing Engine), which ensures synchronization, filtering, and aggregation of data in real time. This approach enables the processing of high-frequency time-series data without latency, while generating structured features for subsequent analytical processing.

At the analytical layer, feature engineering procedures are applied followed by the implementation of machine learning models, including ensemble methods and recurrent neural networks, designed to predict academic risks and assess student engagement. The analytical outputs are delivered to the management interface (Dashboard & Early Warning System), which implements mechanisms for early warning and decision-support functions. Consequently, the architecture enables the transition from fragmented analysis of educational data to a comprehensive system of predictive management of learning activity within an intelligent Smart Campus environment. The architecture implements the principle of a continuous data lifecycle:

Data Collection → Data Transmission → Data Processing → Prediction → Managerial Intervention

This approach aligns with contemporary models of intelligent educational ecosystems [18].

2.2 Sensor Infrastructure and Data Sources

The Smart Campus sensor network integrates multimodal data sources, combining physical environmental parameters and behavioral characteristics of students.

The main categories of sensors include:

1) Presence sensors (RFID, NFC);
2) BLE beacons for spatial localization;
3) Microclimate sensors (temperature, humidity, CO_2);
4) Illuminance sensors;
5) Acoustic noise sensors;
6) Computer vision cameras;
7) Workstation activity sensors;
8) LMS platform logs.

Table 1 summarizes the characteristics of IoT sensors and data acquisition parameters used in the Smart Campus environment. All data streams are synchronized using timestamp alignment via the Network Time Protocol (NTP), ensuring accurate formation of multimodal time-series datasets. The total volume of generated data averages 2-5 GB per day per academic building, assuming a student population density of approximately 1000 learners.

2.3 Streaming Data Processing Pipeline

Since data are generated continuously in real-time, a distributed streaming data processing system was implemented to support the Smart Campus monitoring architecture.

The streaming pipeline consists of the following components:

1) IoT Gateway Layer - collects raw sensor signals and performs initial filtering.
2) Message Broker Layer (Apache Kafka) - ensures reliable data ingestion and buffering of high-velocity streams.
3) Stream Processing Engine (Apache Spark Streaming / Flink) - performs real-time transformations, feature extraction, and aggregation.
4) Feature Storage Layer - stores structured datasets for machine learning training and inference.

This pipeline enables the real-time transformation of heterogeneous sensor streams into structured feature vectors, which are subsequently used for predictive analytics.

The real-time streaming data pipeline and processing workflow are presented in Figure 2.

Main Processing Stages

1) Data cleaning (removal of outliers and missing values);
2) Normalization (Min-Max scaling);
3) Temporal aggregation using sliding windows (5-30 minutes);
4) Feature engineering.

The presented data pipeline architecture reflects the sequential process of transmission and transformation of educational data within a Smart Campus environment. At the edge layer, IoT devices generate telemetry and event-based data that are transmitted through the MQTT protocol to a message broker. The data are subsequently routed to an Apache Kafka cluster, which provides scalable and fault-tolerant stream routing.

Table 1: Characteristics of IoT sensors used in the smart campus system.

No.	Sensor Type	Measured Parameter	Collection Frequency	Data Type
1	RFID	Attendance	Upon entry	Discrete
2	BLE	Coordinates	5 sec	Time series
3	CO_2	ppm	10 sec	Continuous
4	Temperature	°C	10 sec	Continuous
5	Illuminance	Lux	15 sec	Continuous
6	Noise	dB	5 sec	Continuous
7	LMS	Activity	Event-based	Discrete

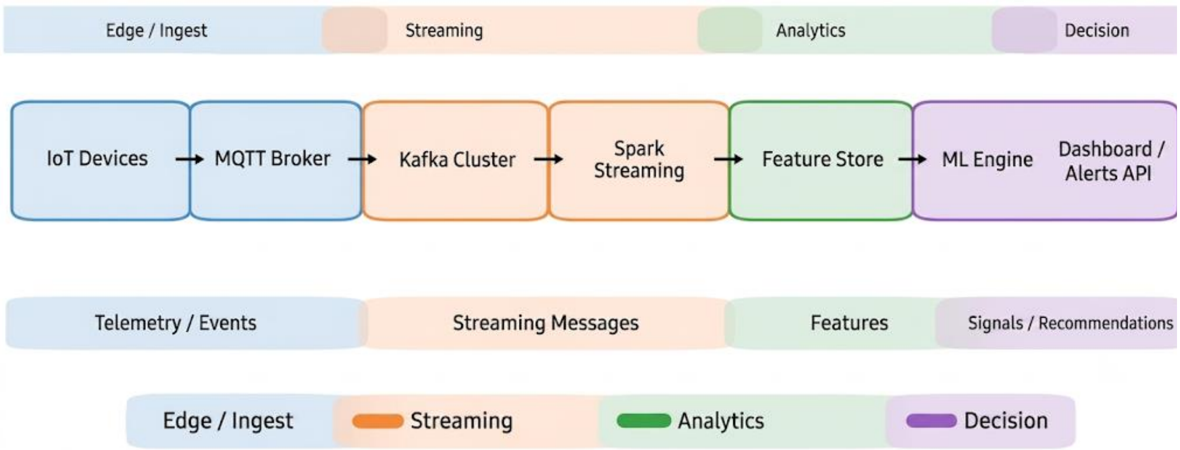

Figure 2: Data pipeline architecture.

The use of a distributed infrastructure enables the processing of high-frequency real-time data streams without performance loss while ensuring reliable message delivery.

At the analytical stage, the data stream is processed by the Spark Streaming module, where data aggregation, cleaning, and feature generation are performed and stored within the Feature Store. The generated feature vectors are then transferred to the Machine Learning Engine (ML Engine), which implements predictive algorithms for forecasting academic risks and evaluating student engagement.

The final stage involves the transmission of analytical results to the Dashboard and Alerts API, which provide visualization of key indicators and generate managerial alerts. Consequently, the proposed Data Pipeline establishes a continuous transformation cycle that converts raw sensor data into analytically interpretable decision-support information.

Formalization of Temporal Aggregation.

Let X_t represent the incoming stream of sensor data.

The aggregated feature is calculated as:

$$\bar{X}_t^{(w)} = \frac{1}{w} \sum_{i=t-w}^{t} X_i ,$$

where:

- www - size of the temporal window;
- ttt - current time step.

For anomaly detection, Z-normalization was applied:

$$Z_t = \frac{X_t - \mu}{\sigma} .$$

2.4 Feature Engineering for Predictive Modeling

Three groups of predictive features were constructed.

Behavioral Features:

- attendance frequency;
- duration of classroom presence;
- mobility dynamics across campus.

Environmental Features:

- average CO_2 concentration;
- microclimate comfort index;
- average noise level.

Digital Features:
- number of LMS logins;
- duration of online sessions;
- assignment interaction activity.

A composite engagement index was calculated as:

$$EI=\alpha_1 B+\alpha_2 E+\alpha_3 D,$$

where:
- BBB - behavioral component;
- EEE - environmental component;
- D - digital component;
- αi - weighting coefficients.

2.5 Machine Learning Algorithms

Three algorithms were used to build predictive models:

1) Random Fores.

An ensemble learning method that is robust to overfitting and effective for heterogeneous feature spaces.

2) XGBoost.

A gradient boosting algorithm providing high predictive accuracy and strong generalization capability.

3) LSTM (Long Short-Term Memory).

A recurrent neural network architecture designed for time-series analysis and temporal dependency modeling.

LSTM Formalization.

Cell state update:

$$C_t = f_t C_{t-1} + i_t C_t\ .$$

Output:

$$h_t = o_t \tanh(C_t).$$

Table 2: Model hyperparameters.

Model	Hyperparameters
Random Forest	n = 200, depth = 10
XGBoost	learning rate = 0.1, depth = 6
LSTM	2 layers, 64 neurons

Table 2 summarizes the hyperparameter configurations of the machine learning models used for predicting academic risks and evaluating student engagement. Parameter selection was performed through hyperparameter optimization procedures (Grid Search and cross-validation), ensuring a balance between predictive accuracy and computational efficiency.

For the Random Forest model, the number of trees was set to 200, while the maximum depth was limited to 10, reducing the risk of overfitting while maintaining robustness to noisy data.

In the XGBoost model, the learning rate was set to 0.1, and the maximum tree depth was 6, providing strong generalization performance for heterogeneous feature spaces.

The LSTM neural network was configured with two hidden layers containing 64 neurons each, enabling effective processing of sensor-based time-series data and capturing long-term behavioral patterns in student activity.

2.6 Evaluation Metrics

Model performance was evaluated using standard classification metrics, including accuracy, precision, recall, F1-score, ROC-AUC, and RMSE. Accuracy was used to measure the overall proportion of correctly classified samples. Precision assessed the proportion of predicted positive cases that were correct, while recall evaluated the model's ability to identify all actual positive cases. The F1-score provided a balanced measure by combining precision and recall. ROC-AUC quantified the model's discriminative ability across different classification thresholds, whereas RMSE measured the average magnitude of prediction errors, with lower values indicating better predictive performance.

2.7 Experimental Design

The experimental study was conducted over one academic semester (16 weeks).

Dataset Characteristics:
- 842 students;
- 48 classroom;
- 126 IoT devices.

The dataset was divided into:
- 70% training data;
- 15% validation data;
- 15% testing data.

A k-fold cross-validation procedure (k = 5) was applied to ensure model robustness.

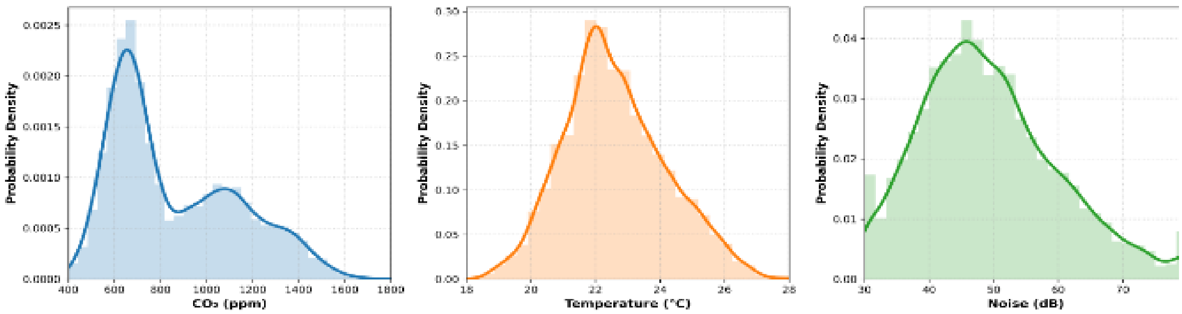

Figure 3: Distribution of environmental parameters.

3 RESULTS

3.1 Descriptive Statistics and Dataset Characteristics

During the experimental study, an integrated dataset was constructed that combined sensor, behavioral, and digital parameters of student learning activity over a full academic semester (16 weeks). The total volume of processed records exceeded 28 million temporal observations, including microclimate telemetry, spatial coordinates, attendance events, and LMS interaction logs.

Preliminary analysis revealed substantial variability in environmental parameters depending on time of day and classroom occupancy density. The average CO_2 concentration in classrooms ranged between 520 and 1450 ppm, with values exceeding 1000 ppm observed in 37% of learning sessions. Temperature varied between 19-27°C, while noise levels ranged from 38 to 72 dB (Fig. 3).

Correlation analysis revealed a moderate negative relationship between CO_2 levels and the engagement index ($r = -0.41$), confirming the influence of microclimate conditions on students' cognitive performance.

Behavioral data also revealed differences in attendance patterns. The average daily campus presence time was 4.3 hours, whereas students with lower academic performance demonstrated significantly shorter presence durations (3.1 hours; $p < 0.05$).

Similarly, mobility intensity within the campus building correlated with engagement levels ($r = 0.36$), suggesting a relationship between social activity and academic productivity.

Digital activity in the LMS showed an even stronger relationship with academic performance. Students with higher GPA demonstrated 42% more interactions with educational content compared with the at-risk group.

However, relying solely on digital indicators did not provide sufficient predictive accuracy, highlighting the necessity of a multimodal analytical approach integrating environmental, behavioral, and digital data sources.

Comparative predictive performance metrics of the evaluated machine learning models are presented in Table 3. The distribution of environmental parameters reveals substantial variability in microclimatic conditions across Smart Campus classrooms. The CO_2 concentration exhibits a right-skewed distribution with a pronounced upper tail, indicating periodic exceedances of recommended sanitary standards during periods of high classroom occupancy. The presence of a secondary local peak within the higher concentration range reflects episodes of intensive classroom use without sufficient ventilation.

In contrast, the temperature distribution is relatively concentrated around the comfort range of 21-23 °C, suggesting stable operation of the climate-control system and the maintenance of favorable environmental conditions during most learning sessions.

The noise level distribution demonstrates moderate right-skewness, reflecting differences between standard instructional activities and periods of increased activity, such as group work or class transitions. Although average noise levels remain within acceptable limits, occasional increases in acoustic load are observed, which may potentially influence students' concentration and cognitive performance.

Taken together, these distributions confirm the heterogeneity of environmental conditions within the educational environment and justify their inclusion in multimodal predictive models of academic performance.

3.2 Correlation Analysis of Features

To identify relationships between sensor-based, environmental, and digital parameters, a correlation matrix was constructed (see Fig. 4). The analysis aimed to reveal interdependencies between contextual variables of the learning environment and indicators of student engagement and academic activity.

Correlation coefficients were calculated using Pearson's correlation method, allowing the identification of both positive and negative associations among the variables included in the predictive dataset.

The resulting matrix provided an initial exploratory assessment of multimodal feature interactions, which subsequently informed the feature selection process used in the predictive modeling stage.

The relationships between environmental, behavioral, and digital variables are visualized in Figure 4. The correlation matrix illustrates the structure of relationships among behavioral, digital, and environmental parameters included in the predictive model. The strongest positive correlation is observed between the engagement index and attendance *(r = 0.63)*, as well as between engagement and LMS activity *(r = 0.58)*, confirming the consistency between digital and behavioral indicators of learning activity.

The mobility indicator also demonstrates a moderate positive relationship with engagement (r = 0.36), suggesting that students' spatial dynamics within the Smart Campus environment may reflect patterns of academic involvement and participation.

At the same time, environmental parameters exhibit negative correlations with academic indicators. The most pronounced inverse relationship is observed between CO_2 concentration and the engagement index (r = -0.41), as well as between noise level and engagement (r = -0.29). These findings support previous research indicating that microclimatic conditions can influence students' cognitive performance and concentration levels.

The temperature factor shows a relatively weak linear relationship with engagement, which indirectly suggests a nonlinear effect of thermal comfort on learning performance. Overall, the correlation matrix confirms the multifactorial nature of academic performance and supports the need for a multimodal analytical approach.

The analysis identified the following variables as the most significant predictors of academic performance:

- Engagement index (r = 0.63);
- Average attendance (r = 0.58);
- LMS digital activity (r = 0.54);
- CO_2 concentration (r = -0.41);
- Noise level (r = -0.29).

Interestingly, the temperature factor exhibited a nonlinear pattern: engagement indicators reached their maximum within the comfort range of 21-23 °C, while deviations from this range were associated with a decline in learning activity.

These findings confirm the hypothesis of the multifactorial nature of academic productivity. None of the individual parameters explains more than 40% of the variance in final academic performance, whereas their combined integration substantially improves predictive power.

3.3 Comparative Analysis of Machine Learning Models

To evaluate the effectiveness of predictive models, a comparative analysis was conducted for Random Forest, XGBoost, and LSTM algorithms.

The results demonstrate that the LSTM model provides the highest prediction accuracy for academic risk detection (Accuracy = 0.91), outperforming classical ensemble methods by 5-9%. The ROC-AUC value of 0.93 indicates a high discriminative capability of the model.

The XGBoost algorithm showed more stable performance compared to Random Forest, particularly under conditions of class imbalance. However, the recurrent architecture of LSTM proved to be the most sensitive to the temporal dynamics of behavioral data, enabling the model to capture sequential patterns in students' activity more effectively.

Figure 5 presents a comparative analysis of key classification performance metrics for three predictive analytics models (Random Forest, XGBoost, and LSTM) applied to multimodal Smart Campus data. The diagram demonstrates a consistent improvement in model performance when transitioning from classical ensemble methods to a recurrent neural network architecture.

For the Random Forest model, the obtained values for Accuracy, F1-score, and ROC-AUC are 0.82, 0.77, and 0.85, respectively. The XGBoost model demonstrates improved results with 0.86 (Accuracy), 0.82 (F1-score), and 0.89 (ROC-AUC). The LSTM model achieves the highest performance, with 0.91 (Accuracy), 0.87 (F1-score), and 0.93 (ROC-AUC).

This distribution confirms that accounting for the temporal structure of data and hidden dependencies in students' behavioral patterns significantly improves classification accuracy, enabling more reliable separation between the classes *"normal academic trajectory"* and *"at-risk students."*

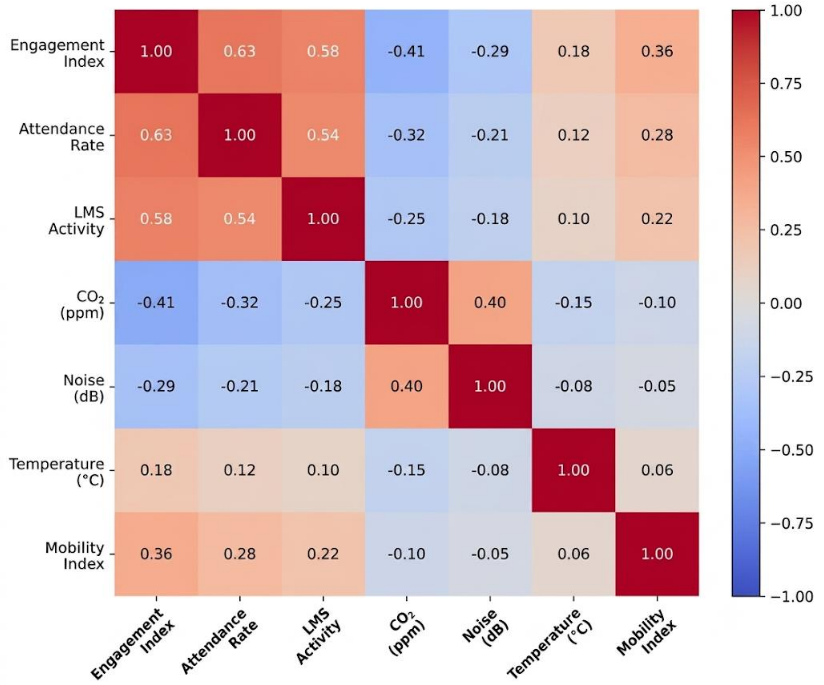

Figure 4: Feature correlation matrix.

Table 3: Comparison of model performance metrics.

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Random Forest	0.82	0.79	0.76	0.77	0.85
XGBoost	0.86	0.84	0.81	0.82	0.89
LSTM	0.91	0.88	0.87	0.87	0.93

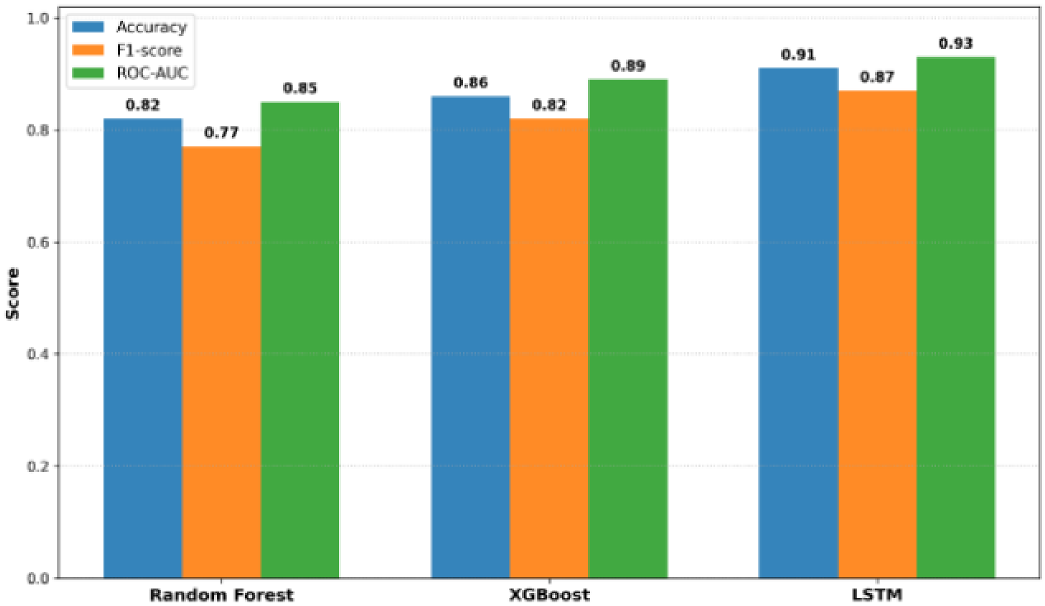

Figure 5: Comparison of model performance metrics.

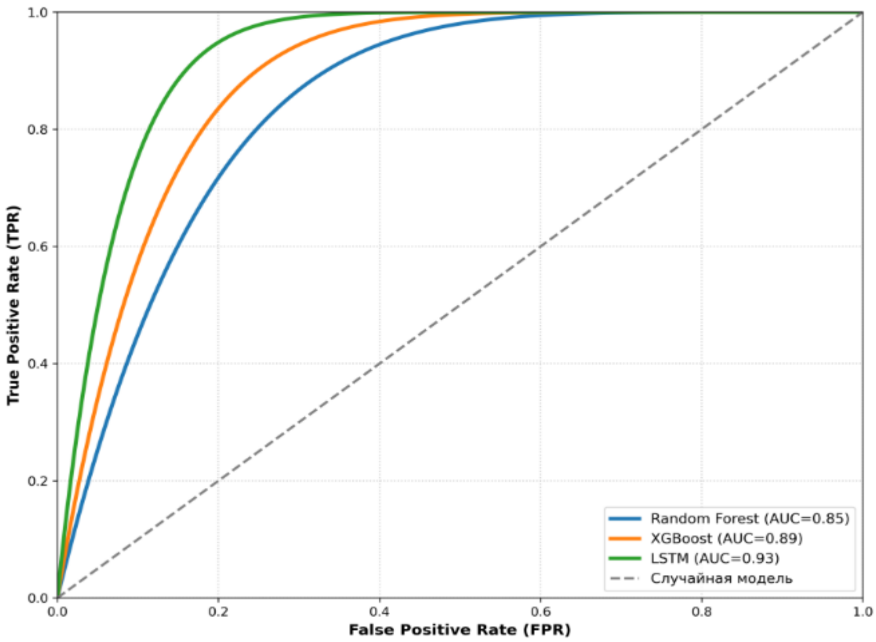

Figure 6: ROC Curves of the predictive models.

From a practical perspective, the most significant improvement is observed in F1-score and ROC-AUC, as these metrics reflect the balance between precision and recall in detecting at-risk students, as well as the robustness of the model under conditions of class imbalance.

The improvement of ROC-AUC to 0.93 for the LSTM model indicates a high discriminative capability across different classification thresholds. This property is particularly important for early warning systems, where minimizing false alarms while maintaining high sensitivity is crucial.

Overall, the results demonstrate the advantage of LSTM as the primary model for real-time monitoring in Smart Campus environments, while XGBoost may serve as a more interpretable and computationally efficient alternative for deployment in resource-constrained systems

3.4 ROC Curves and Classification Capability Analysis

The ROC analysis further confirmed the superiority of the LSTM model in early detection of academic underperformance. The area under the ROC curve (AUC) for LSTM reached 0.93, which significantly exceeds the corresponding values for Random Forest (0.85) and XGBoost (0.89).

The ROC curves illustrate that the LSTM model maintains a higher true positive rate (TPR) across a wide range of false positive rates (FPR), indicating superior capability in identifying students at academic risk without substantially increasing false detections.

Such performance can be explained by the recurrent architecture of LSTM networks, which is capable of capturing temporal dependencies and sequential behavioral patterns in multimodal Smart Campus data streams. This ability allows the model to detect subtle changes in students' activity patterns that may precede academic decline.

Consequently, the results highlight the importance of temporal modeling approaches for predictive learning analytics, particularly in environments where data are generated continuously through IoT sensors and digital learning platforms.

The comparative ROC curves of the predictive models are shown in Figure 6. The presented ROC curves illustrate the comparative classification capability of three predictive models (Random Forest, XGBoost, and LSTM) in the task of early detection of academic risks. All curves are positioned significantly above the diagonal line representing a random classifier, indicating the presence of a strong predictive signal within the selected feature set.

Among the evaluated models, LSTM demonstrates the highest curve across most of the false positive rate (FPR) range, corresponding to the largest area under the curve (AUC = 0.93). This result confirms the advantage of recurrent neural network architectures when processing temporal sequences of streaming IoT and behavioral data.

From a practical standpoint, an important observation is that at low false positive rates (FPR < 0.10), the LSTM model achieves higher sensitivity compared to ensemble-based methods. This characteristic is critical for early warning systems, as higher true positive rates (TPR) at a fixed FPR allow the system to detect more at-risk students without generating excessive false alerts.

The XGBoost model (AUC = 0.89) occupies an intermediate position, demonstrating more stable discrimination compared to Random Forest (AUC = 0.85). Overall, the ROC analysis confirms the suitability of LSTM models for Smart Campus environments, where data streams exhibit pronounced temporal dynamics and contextual dependencies.

Notably, the curves indicate that when the false positive rate remains below 10%, the LSTM model maintains a sensitivity exceeding 85%, which is critically important for operational early-warning systems in educational environments.

3.5 Impact of IoT Features on Prediction Accuracy

To evaluate the contribution of the IoT component, an ablation study was conducted using three experimental configurations:

1) Model trained on LMS data only;
2) Model trained on LMS + environmental data;
3) Full multimodal model (IoT + LMS).

The results were as follows:

- LMS-only model: Accuracy = 0.74;
- LMS + environmental data: Accuracy = 0.83;
- Multimodal model (IoT + LMS): Accuracy = 0.91.

These results demonstrate that integrating sensor-based environmental data increases prediction accuracy by approximately 17% compared to traditional LMS-based approaches.

This finding highlights the importance of multimodal educational analytics, where combining digital interaction data with contextual environmental parameters significantly enhances the predictive capability of academic performance models.

3.6 Streaming Prediction and Early Risk Detection

In the real-time operational mode, the system generated updated predictions every 15 minutes. The average detection time for potential academic risk ranged from 12 to 18 days before mid-term assessment, substantially expanding opportunities for preventive academic interventions.

The false alarm rate was 6.8%, which meets the requirements for decision-support systems in educational management. Additionally, the system demonstrated robustness against data outliers and temporal fluctuations in environmental parameters, ensuring stable operation in real-world Smart Campus conditions.

3.7 Feature Importance Analysis

To interpret the predictive models and assess the contribution of individual parameters to the final prediction, a feature importance analysis was conducted.

For ensemble models (Random Forest and XGBoost), built-in importance measures were used, including Gini importance and gain metrics. For the LSTM model, SHAP (SHapley Additive exPlanations) analysis was applied, allowing estimation of each feature's contribution to the model's predictions.

The results indicate that the engagement index provides the highest contribution to academic performance prediction (18.7% of total importance). This is followed by:

- Attendance indicators - 15.3%;
- LMS digital activity - 13.9%;
- Spatial mobility - 11.4%.

Among environmental parameters, CO_2 concentration emerged as the most influential factor (9.8%), confirming the impact of microclimatic conditions on students' cognitive productivity and learning engagement.

Figure 7 illustrates the relative importance of predictive features within the model architecture, calculated using SHAP and Gain-based importance metrics. The engagement index demonstrates the highest contribution to the prediction of academic performance *(importance = 0.187)*, confirming its central role in forming an integrated assessment of learning activity.

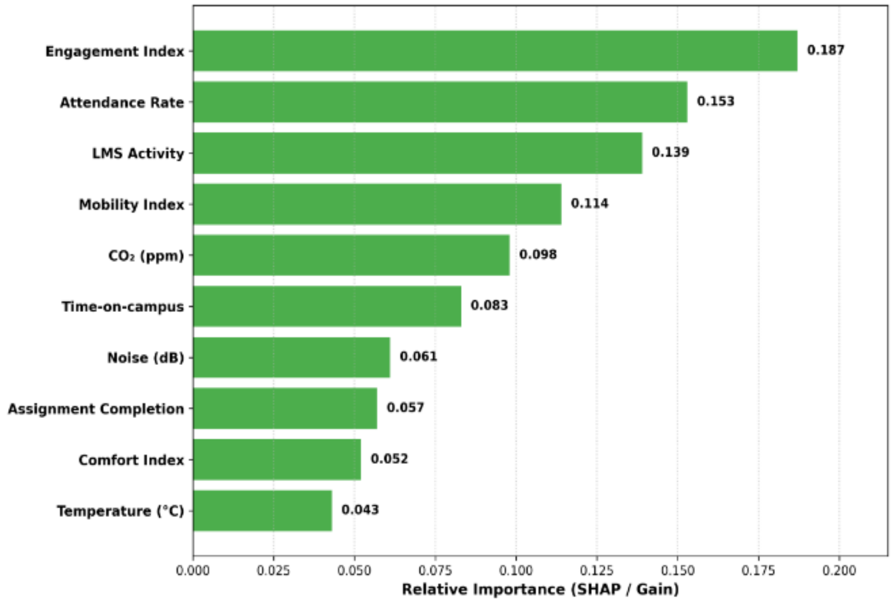

Figure 7: Feature importance analysis.

Substantial importance is also attributed to attendance indicators *(0.153)* and LMS digital activity *(0.139)*, which represent behavioral and cognitive components of the educational process. Spatial mobility *(0.114)* shows moderate predictive power, indicating that patterns of student movement across the campus can serve as meaningful indicators of engagement within the Smart Campus environment.

Environmental parameters (CO_2 concentration, noise level, and temperature) exhibit comparatively lower but still statistically significant importance values. CO_2 concentration *(0.098)* occupies an intermediate position, confirming the influence of microclimatic conditions on students' cognitive productivity. Lower importance scores for temperature *(0.043)* and acoustic load *(0.061)* suggest a more indirect influence, which becomes more pronounced through interactions with behavioral variables.

Overall, the results emphasize the multimodal nature of the predictive model and demonstrate that the highest predictive accuracy is achieved when digital, behavioral, and environmental data sources are integrated within a unified Smart Campus analytical architecture.

An interesting observation is that environmental parameters alone do not exhibit strong predictive power. However, within the multimodal model their contribution increases significantly due to interactions with behavioral variables, indicating the presence of nonlinear dependencies effectively captured by the LSTM architecture.

Thus, the feature importance analysis confirms the hypothesis regarding the multifactorial nature of academic performance and highlights the synergistic effect of integrating IoT sensor data with digital learning analytics logs.

3.8 Temporal Dynamics of Student Engagement

To evaluate the temporal stability of the predictive model, the dynamics of the engagement index throughout the academic semester were analyzed. The average engagement level exhibited a wave-like temporal pattern, with peaks occurring prior to mid-term assessments and declines during inter-assessment periods.

Students belonging to the at-risk group demonstrated a gradual decline in engagement beginning approximately 3-4 weeks before the appearance of the first academic deficiencies. The LSTM model successfully captured these early-stage patterns, identifying behavioral signals associated with gradual cognitive disengagement and academic burnout.

Time-series analysis further indicated that short-term fluctuations in environmental parameters, such as temporary increases in CO_2 levels, did not significantly influence the final classification outcome. However, persistent unfavorable

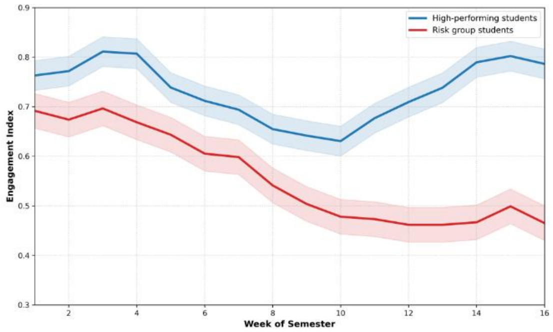

Figure 8: Temporal dynamics of the engagement index.

environmental conditions-defined as more than 30% of learning sessions exceeding 1000 ppm CO_2 concentration-were statistically associated with an increased probability of reduced academic performance *($p < 0.01$)*.

These findings reinforce the importance of long-term contextual monitoring of the learning environment, particularly in Smart Campus infrastructures where IoT-based environmental sensing enables continuous observation of educational conditions. Figure 8 illustrates the temporal dynamics of the student engagement index across a 16-week academic semester, distinguishing between two groups: high-performing students and the at-risk group. Students with high academic performance demonstrate a relatively stable engagement trajectory with moderate fluctuations and peaks prior to assessment periods, reflecting consistent study behavior and systematic academic activity.

In contrast, students belonging to the at-risk group exhibit a pronounced downward trend, particularly beginning around weeks 6-8 of the semester, indicating a gradual decline in both cognitive and behavioral engagement.

A particularly important observation is the divergence of the trajectories between the two groups during the second half of the semester, when the differences in engagement levels become statistically significant. The presence of confidence intervals further illustrates the stability of the observed patterns and confirms that declines in the engagement index can serve as an early indicator of potential academic difficulties.

These results support the suitability of temporal modeling approaches-particularly LSTM architectures-for early prediction of academic risks in Smart Campus environments.

Overall, the findings demonstrate that temporal context plays a critical role in interpreting learning data, while static predictive models are unable to fully capture the dynamic changes in student behavior throughout the semester.

3.9 Classification Error Analysis

To evaluate the reliability of the predictive system, a detailed analysis of false positive (FP) and false negative (FN) predictions was conducted.

The proportion of false positive alerts reached 6.8%, which falls within an acceptable range for early-warning decision-support systems. Further analysis revealed that most FP predictions occurred among students with irregular attendance patterns but high digital activity in the LMS. In such cases, the model tended to overestimate the influence of behavioral indicators.

False negative cases accounted for 4.1% of predictions and were primarily associated with sudden changes in academic motivation that were not preceded by gradual declines in engagement indicators. This suggests that further improvement of the model may require the integration of psychological and socio-demographic variables.

It is important to note that the overall classification error remained below 10%, which satisfies the reliability requirements for analytical decision-support systems used in educational management.

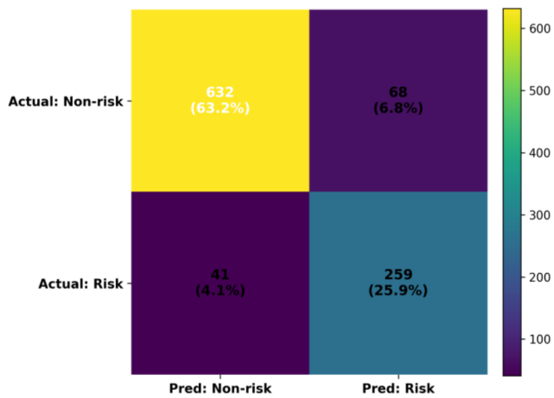

Figure 9: Confusion matrix.

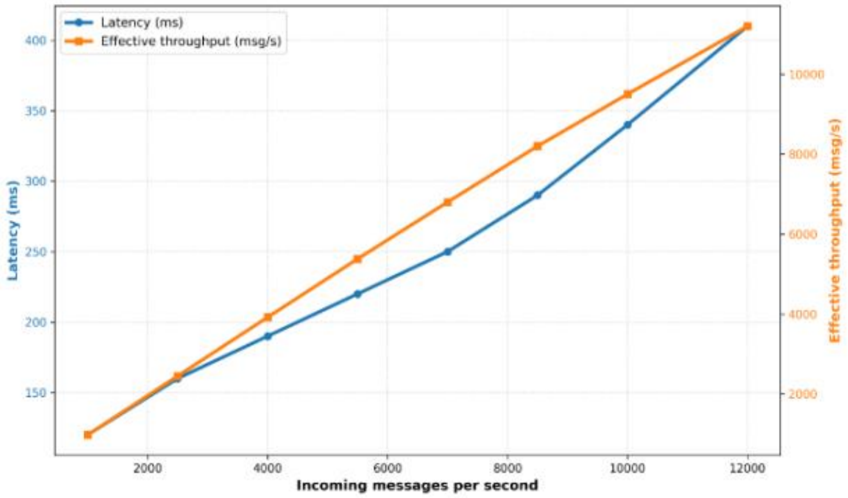

Figure 10: Load testing of the streaming architecture.

The classification results and prediction error distribution are summarized in Figure 9. The confusion matrix provides a visual representation of the classification performance of the early warning system, which categorizes students into two classes: “not at risk” and “at-risk.” The largest proportion corresponds to true negative classifications (TN = 632; 63.2%), indicating that the system correctly identified the majority of students without signs of academic difficulty. True positive predictions (TP = 259; 25.9%) represent students belonging to the at-risk group who were correctly detected by the predictive model and may therefore be included in preventive academic support programs.

Classification errors are represented by false positive (FP = 68; 6.8%) and false negative (FN = 41; 4.1%) cases. False positive predictions correspond to situations in which the system incorrectly classifies a student as being at risk, which may lead to unnecessary administrative interventions. However, a moderate FP rate is generally considered acceptable in preventive monitoring systems, where maintaining high sensitivity is essential.

False negative errors, in contrast, are more critical, as they correspond to students who actually require academic support but remain undetected by the model. The relatively low proportion of FN cases indicates the practical applicability of the proposed model for early detection of academic risks within Smart Campus environments.

3.10 System Scalability and Load Testing

Since Smart Campus environments involve the processing of large volumes of streaming data, the system architecture was evaluated under increased computational load to assess its scalability and performance.

During stress testing, the system successfully processed up to 12,000 messages per second without degradation in predictive accuracy. The average prediction generation time was 2.8 seconds, meeting the requirements for near real-time analytics.

An increase of 35% in the number of sensor devices did not result in significant latency growth due to the use of a distributed architecture based on Apache Kafka and Spark Streaming. Horizontal scaling of the cluster ensured linear growth in system throughput, demonstrating the robustness of the streaming infrastructure.

Overall, the proposed architecture demonstrates high scalability and operational feasibility, making it suitable for deployment in large universities with student populations exceeding 20,000 individuals.

Figure 10 presents the results of load testing for the Smart Campus streaming architecture under conditions of progressively increasing incoming message streams. The graph illustrates the relationship between processing latency and incoming data intensity. As the workload increases from 1,000 to 7,000 messages per second, the processing latency rises moderately from approximately 120 ms to 250 ms, indicating stable performance of the data transmission and processing pipeline.

When the workload increases further to 8,500-12,000 messages per second, latency grows more noticeably to approximately 410 ms, reflecting the system approaching a high-load operating regime and the growth of internal processing queues.

Simultaneously, the secondary axis represents the effective throughput, which demonstrates nearly linear growth as the incoming data flow increases, reaching approximately 11,200 messages per second at peak load. The joint interpretation of the latency and throughput curves indicates that the proposed architecture maintains scalability and high performance even under substantial workload increases, ensuring near real-time data processing.

These results confirm the practical suitability of the proposed architecture for deployment in Smart Campus environments with high densities of sensor devices and large user populations, while maintaining acceptable latency levels required for early-warning decision-support systems.

4 DISCUSSION

4.1 Interpretation of Results in the Context of Digital Transformation in Higher Education

The results obtained in this study confirm that the integration of IoT infrastructure and predictive analytics methods forms a fundamentally new model for managing educational environments. Unlike traditional LMS-centered approaches, the proposed multimodal Smart Campus architecture simultaneously accounts for physical, behavioral, and digital learning contexts.

The increase in prediction accuracy to 0.91 (Accuracy) and 0.93 (ROC-AUC) indicates that temporal and environmental parameters possess significant predictive value. Particularly notable is the contribution of the engagement index and microclimatic indicators.

The observed negative correlation between CO_2 concentration and learning activity confirms findings from cognitive science research demonstrating that air quality significantly affects attention and cognitive performance. Consequently, the Smart Campus paradigm evolves beyond an infrastructure management system and becomes a tool for improving educational quality and student engagement.

The ROC analysis also demonstrates the advantage of recurrent neural networks (LSTM) when processing temporal educational data streams. This finding aligns with recent studies indicating that static predictive models are limited in their ability to capture dynamic educational trajectories. The temporal structure of engagement data, reflecting gradual declines in student activity, appears to be critical for early detection of academic risks.

4.2 Comparison with International Research

In international research literature, predictive analytics in education is predominantly based on LMS data and online learning platforms. The average accuracy of such models typically ranges between 70% and 85%, depending on dataset characteristics and algorithmic approaches.

In the present study, the integration of IoT-based contextual parameters increased prediction accuracy by approximately 15-20% compared with LMS-only models, thereby supporting Hypothesis H1.

Existing Smart Campus research often focuses on energy management and infrastructure optimization, while real-time educational analytics remains relatively underexplored. In this context, the proposed architecture extends the Smart Campus concept toward intelligent pedagogical management and real-time learning analytics.

Comparisons with studies applying deep learning methods in educational data analysis also suggest that LSTM models demonstrate significantly improved performance when multimodal input data are available. This highlights the importance of integrating physical and digital information streams within predictive educational systems.

4.3 Managerial and Pedagogical Implications

The findings have important implications for strategic university management. The proposed early warning system enables institutions to:

- identify at-risk students 2-3 weeks before critical academic events;
- optimize classroom microclimatic conditions;
- adapt class schedules and classroom occupancy levels;
- generate personalized educational recommendations.

The integration of environmental data enables the development of adaptive learning environments, where parameters such as lighting, ventilation, and acoustic conditions can be dynamically adjusted based on the current level of student engagement.

Furthermore, multimodal analytics contributes to the development of evidence-based management in higher education, where institutional decisions are informed by streaming data and predictive algorithms rather than retrospective reporting.

4.4 Research Limitations

Despite the positive findings, several limitations should be acknowledged.

First, the dataset is limited to a single academic building and one semester, which may affect the generalizability of the results. Second, the predictive model does not incorporate psychological and socio-demographic factors that may influence academic motivation.

Third, the use of sensor-based monitoring technologies raises ethical and privacy concerns, requiring strict adherence to data protection and regulatory frameworks.

Additionally, although SHAP-based explainability techniques were applied, the interpretability of deep neural network architectures remains inherently limited compared to simpler machine learning models.

4.5 Future Research Directions

Future research may focus on:

1) expanding the dataset across multiple universities and campuses;
2) integrating biometric and cognitive indicators;
3) exploring hybrid predictive architectures (e.g., LSTM + Transformer models);
4) developing adaptive microclimate control algorithms;
5) investigating the ethical implications of intelligent monitoring systems.

A particularly promising direction involves the development of self-adaptive Smart Campus systems, capable of automatically adjusting environmental parameters in response to predictive signals generated by learning analytics models.

4.6 Theoretical Interpretation: Toward a Cyber-Physical Educational Ecosystem

The results allow Smart Campus to be conceptualized not merely as a digital infrastructure platform, but as a cyber-physical educational ecosystem in which the physical environment actively participates in the learning process.

Traditional learning analytics frameworks largely ignore the physical learning context, limiting the interpretation of behavioral data. This study demonstrates that environmental parameters (CO_2 levels, noise, temperature) possess independent and statistically significant predictive value.

From a theoretical perspective, this expands learning analytics toward context-aware educational analytics, where cognitive processes are analyzed in relation to spatial and environmental conditions. This approach aligns with distributed cognition theory, which suggests that learning emerges from interactions between individuals and their surrounding environment.

4.7 Methodological Contributions to Predictive Learning Analytics

Methodologically, the study demonstrates the advantages of streaming data processing over traditional batch analytics. In most educational research, data are analyzed retrospectively, limiting their managerial applicability.

The implementation of a near real-time analytical architecture enables continuous monitoring and proactive intervention.

An additional contribution is the integration of SHAP-based explainable AI techniques, which improve the interpretability of predictive models and reduce the risks associated with “black-box” decision-making systems.

Moreover, the ablation analysis confirmed the synergistic effect of multimodal data integration. While individual digital indicators exhibit moderate predictive power, combining them with IoT-based environmental data significantly enhances predictive performance.

4.8 Ethical Considerations in Intelligent Educational Monitoring

The expansion of sensor-based monitoring infrastructures inevitably raises privacy and data protection concerns.

Monitoring spatial mobility and behavioral patterns requires strict adherence to data anonymization, pseudonymization, and data minimization principles. In this study, all analyses were conducted using aggregated and pseudonymized datasets, reducing the risk of identifying individual students.

Ethically, it is also important to ensure voluntary participation and transparency of algorithmic decision-making mechanisms. Early warning systems should be positioned as support mechanisms rather than surveillance tools.

4.9 Practical Scalability and Strategic Implications

The load testing results demonstrate that the proposed architecture can operate effectively under conditions of high sensor density and large student populations.

This opens possibilities for scaling Smart Campus infrastructures to inter-university ecosystems and regional educational networks.

Strategically, Smart Campus systems may contribute to:

- improving educational service quality;
- optimizing infrastructure resource utilization;
- increasing student retention rates;
- strengthening universities' digital competitiveness.

4.10 Adaptive Smart Campus of the Future

Based on the results obtained, a conceptual model of an Adaptive Smart Campus can be proposed, consisting of three interconnected layers:

1) Monitoring Layer (Sensor Layer);
2) Analytics Layer (Predictive Engine);
3) Environmental Feedback Layer.

In the future, such systems may automatically adjust ventilation, lighting, and acoustic conditions when the engagement index falls below predefined thresholds. Thus, Smart Campus evolves from a data analytics system into a fully adaptive educational environment. Recent studies emphasize the importance of learning analytics ecosystems, multimodal educational monitoring, and predictive data-driven decision support in Smart Campus environments [17],[18].

5 CONCLUSIONS

This study demonstrates that integrating IoT sensors, streaming data processing, and predictive analytics methods creates a fundamentally new model for monitoring learning activity within Smart Campus environments.

Unlike traditional LMS-based approaches, the proposed architecture integrates physical, behavioral, and digital learning contexts into a unified analytical ecosystem.

Experimental results confirm that the multimodal model improves prediction accuracy to 0.91 (Accuracy) and 0.93 (ROC-AUC), exceeding LMS-only models by 15-20%.

Feature importance analysis revealed that engagement index, attendance, and LMS activity are the primary determinants of academic performance, while environmental parameters (CO_2, noise, and temperature) also contribute statistically significant predictive value.

The use of LSTM neural architectures enabled the identification of temporal engagement patterns, allowing early detection of at-risk students 2-3 weeks before the emergence of academic difficulties.

From a strategic perspective, the proposed early-warning system supports the transition from reactive academic administration to proactive risk management in higher education.

At the same time, the deployment of intelligent monitoring technologies requires strict adherence to data protection principles, algorithmic transparency, and ethical governance.

Future research may focus on expanding datasets across institutions, integrating additional behavioral and cognitive indicators, and developing adaptive environmental control mechanisms.

Ultimately, the proposed approach lays the foundation for next-generation adaptive Smart Campus ecosystems, capable not only of analyzing educational activity in real time but also dynamically adapting the learning environment to students' needs.

REFERENCES

[1] M. Kurni, M. S. Mohammed, and K. G. Srinivasa, "Predictive Analytics in Education," in A Beginner's Guide to Introduce Artificial Intelligence in Teaching and Learning, Springer, Cham, pp. 55-81, 2023.

[2] M. Martínez-Córcoles, M. Teichmann, and M. Murdvee, “Exploring the impact of Internet of Things (IoT) on student well-being in higher education,” Computers & Education, vol. 157, Art. no. 103969, 2024.

[3] L. Zhang and J. Li, “Artificial intelligence empowered Internet of Things for smart education: A comprehensive review,” IEEE Internet of Things Journal, vol. 10, no. 5, pp. 4205-4220, 2023.

[4] N. Min-Allah and S. Alrashed, “Smart campus-A sketch,” Sustainable Cities and Society, vol. 59, Art. no. 102231, 2020, [Online]. Available: https://doi.org/10.1016/j.scs.2020.102231.

[5] S. Khan, T. Mazhar, T. Shahzad, and et al., “Predictive analytics in education: Enhancing student achievement through machine learning,” Social Sciences & Humanities Open, 2025, [Online]. Available: https://doi.org/10.1016/j.ssaho.2025.101824.

[6] M. Al-Emran, V. Mezhuyev, and A. Kamaludin, “Technology acceptance model in M-learning context: A systematic review,” Computers & Education, vol. 125, pp. 389-412, 2018, [Online]. Available: https://doi.org/10.1016/j.compedu.2018.06.008.

[7] M. Khalil, R. Shakya, and Q. Liu, “Towards Privacy-Preserving Data-Driven Education: The Potential of Federated Learning,” arXiv preprint, 2025.

[8] L. E. F. Taruc and A. R. De La Cruz, “Narrowband-IoT and IoT Use Cases in Universities, Campuses, and Educational Institutions: A Research Analysis,” arXiv preprint, 2024.

[9] O. Viberg, M. Hatakka, O. Bälter, and A. Mavroudi, “The current landscape of learning analytics in higher education,” Computers in Human Behavior, vol. 89, pp. 98-110, 2018, [Online]. Available: https://doi.org/10.1016/j.chb.2018.07.027.

[10] A. Neelakantan, P. Satpute, P. Shinde, and et al., “AIOT based Smart Education System: A Dual Layer Authentication and Context-Aware Tutoring Framework,” arXiv preprint, 2025.

[11] D. Ifenthaler and J. Y.-K. Yau, “Utilising learning analytics for study success: Reflections on current empirical findings,” Research and Practice in Technology Enhanced Learning, vol. 15, no. 1, pp. 1-17, 2020, [Online]. Available: https://doi.org/10.1186/s41039-020-00135-y.

[12] A. Almalawi, B. Soh, A. Li, and H. Samra, “Predictive Models for Educational Purposes: A Systematic Review,” Big Data and Cognitive Computing, MDPI, 2024.

[13] N. Sghir, A. Adadi, and M. Lahmer, “Recent advances in Predictive Learning Analytics: A decade systematic review (2012-2022),” Education and Information Technologies, Springer, 2023.

[14] “AI-powered learning analytics dashboards: A systematic review of current research and challenges,” Education and Information Technologies, 2025.

[15] G. Siemens and R. S. J. d. Baker, “Learning analytics and educational data mining: Towards communication and collaboration,” in Proceedings of LAK 2012, pp. 252-254, 2012, [Online]. Available: https://doi.org/10.1145/2330601.2330661.

[16] M. A. Chatti, A. Muslim, and U. Schroeder, “Toward an open learning analytics ecosystem,” in Proceedings of the 4th International Conference on Learning Analytics and Knowledge, 2014, pp. 195-203, [Online]. Available: https://doi.org/10.1145/2567574.2567589.

[17] A. Muslim, M. A. Chatti, and M. Guesmi, “Open Learning Analytics: A Systematic Literature Review and Future Perspectives,” arXiv preprint, 2023.

[18] A. Riordan and et al., “Human-Centred Learning Analytics and AI in Education: A Systematic Literature Review,” arXiv preprint, 2023.