International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Development of a Social Computing-Based Intelligent Class Record System for Early Detection of At-Risk Students Using Learning Analytics and the OULAD Dataset

John Ivan C. Maurat

Assistant Professor, College of Computing and Information Technologies, National University Philippines

Abstract

This study aimed to design, develop, and evaluate a social computing-based intelligent class record system for the early detection of at-risk students using learning analytics, social computing, and artificial intelligence-driven predictive modeling. Traditional class record systems primarily function as repositories of attendance and grades and often lack predictive and analytical capabilities necessary for timely academic intervention. Addressing this limitation, the proposed system integrates demographic, assessment, and engagement data derived from the Open University Learning Analytics Dataset (OULAD) to provide data-driven insights into student performance and engagement patterns.

The study employed a developmental research approach involving data preprocessing, feature engineering, machine learning model development, dashboard design, and usability evaluation. Several predictive algorithms, namely logistic regression, random forest, neural network, and extreme gradient boosting (XGBoost), were utilized and compared to classify students into pass, fail, withdraw, and distinction categories. In addition, Social Network Analysis (SNA) metrics such as degree centrality, clustering, and interaction frequency were incorporated to analyze peer engagement and collaborative behaviors.

The developed system features an interactive dashboard that visualizes student risk levels, engagement trends, assessment performance, and social interaction patterns. Furthermore, the system generates automated intervention recommendations based on predictive outputs and behavioral indicators to support proactive instructional decision-making. The usability and perceived effectiveness of the system were evaluated using the System Usability Scale (SUS) among selected educators.

Findings of the study revealed that ensemble and machine learning approaches demonstrated strong predictive capabilities in identifying at-risk students, while the integration of social computing enhanced the contextual understanding of learner engagement and persistence. The proposed intelligent class record system provides educators with a holistic and interpretable platform for monitoring academic performance and initiating timely interventions. The study contributes to the growing field of learning analytics by integrating predictive analytics, social computing, explainable artificial intelligence, and dashboard visualization into a unified educator-centered system that supports evidence-based and proactive educational practices.

Keywords: At-Risk Students, Dashboard, Early Warning System, Educational Data Mining, Educational Technology, Learning Analytics template

1. Introduction

Traditional class record systems in higher education primarily function as repositories for attendance and grades. While they support documentation, they lack analytical and predictive capabilities needed to identify struggling students early. As a result, interventions are often delayed, occurring only after consistent poor performance or disengagement. This limitation highlights the need for predictive analytics, which can transform student data into actionable insights and enable educators to anticipate academic risks before they escalate. By supporting proactive rather than retrospective responses, such systems can reduce dropout rates and improve learning outcomes, particularly in online and blended environments where real-time supervision is limited.

Within this context, the concept of at-risk students has become central to modern educational analytics. At-risk learners refer to students whose behavioral, academic, or engagement patterns indicate a higher likelihood of failure, withdrawal, or prolonged academic difficulty. Risk indications include inconsistent participation, delayed submissions, and decreased involvement, in addition to low grades. Since academic risk develops gradually over time, early identification and timely intervention are essential, emphasizing the need for systems capable of continuous monitoring and prediction.

Globally, dropout rates in tertiary education remain a persistent concern, with some institutions reporting rates as high as 30% (Tinto, 2012). In the Philippines, this problem is exacerbated by problems such as inadequate access to learning support, poor monitoring tools, and a quick transition toward flexible learning modes. These conditions highlight the critical need for sophisticated technologies that can help instructors identify at-risk students early on and intervene in a timely manner.

The subject of learning analytics (LA) provides a solid platform for addressing these difficulties by emphasizing the systematic gathering, analysis, and interpretation of learner data to enhance educational results. Rather than treating student records as static information, learning analytics interprets behavioral and performance data as dynamic indicators of future academic success. This shift enables educators to move from reactive evaluation toward proactive, evidence-based decision-making.

Empirical research utilizing datasets like the Open University Learning Analytics Dataset (OULAD) shows that predictive algorithms may accurately identify students depending on their likelihood of passing, failing, or withdrawing by examining early engagement indicators (Kuzilek et al., 2017). These findings reinforce the value of structured data analytics in identifying academic risk at an early stage.

In addition, social computing extends learning analytics by examining how peer interactions and collaborative behaviors influence student performance. By incorporating social dimensions of learning, risk detection becomes more comprehensive, capturing not only individual performance but also engagement within learning communities.

Furthermore, AI and ML approaches including logistic regression, random forests, neural networks, and gradient boosting improve predictive accuracy by revealing complex patterns in educational data (Lakkaraju et al., 2017). When integrated into learning systems, these models can generate interpretable predictions that support instructional planning and intervention strategies.

In the Philippine educational context, the increasing adoption of technology-assisted learning has amplified the demand for data-driven monitoring tools. Although many institutions use learning management systems, most still lack advanced analytics capabilities for early risk detection. As a result, educators often rely on manual monitoring and subjective judgment, which limit timely intervention. Addressing this gap, this study proposes an integrated intelligent class record system that combines learning analytics, social computing, and AI-based predictive modeling using the OULAD dataset. Unlike previous studies that focus on isolated components such as prediction accuracy or visualization tools, this research unifies predictive analytics, social network analysis, and interactive dashboards into a single educator-centered platform. The system transforms traditional class records into an intelligent environment that monitors performance, predicts student risk, and visualizes engagement patterns, thereby enabling timely and targeted interventions.

This study is the first to integrate predictive modeling, social network analytics, and automated intervention recommendation into a unified class record system, thereby bridging the gap between advanced educational data mining techniques and their practical application in everyday teaching. By combining predictive accuracy with social learning insights and actionable recommendations, the proposed approach allows educators to conduct timely, targeted, and data-driven interventions, thereby enhancing student retention and academic performance.

2. Prepare Your Paper Before Styling

Foundations: Learning Analytics, EDM, and Theoretical Models

Learning Analytics (LA) and Educational Data Mining (EDM) are two complimentary but distinct research paradigms for extracting useful insights from raw digital learning data. EDM emphasizes computational rigor, focusing on algorithms, feature extraction, and prediction accuracy, whereas LA takes a broader approach, incorporating educational philosophy, institutional priorities, and actionable insights for educators and administrators (Siemens & Baker, 2012; Romero & Ventura, 2010). Scholars have advocated for greater interaction between the two communities to ensure that methodological breakthroughs are not isolated from the reality of teaching and learning.

Theoretical frameworks frequently impact the design and evaluation of LA tools. For example, the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) are frequently used to investigate instructors' adoption and views of dashboards, analytics systems, and intelligent class record technologies (Davis, 1989). In parallel, educational theories such as Tinto's model of student retention emphasize the significance of early detection of at-risk students, linking predictive analytics to current retention research (Tinto, 1993).

Datasets & Benchmarks: OULAD and Comparable Repositories

Open datasets, such as the Open University Learning Analytics Dataset (OULAD), have been helpful in enabling repeatable research. OULAD contains extensive student demographic, assessment, and virtual learning environment (VLE) activity data from over 30,000 students across multiple courses, making it a useful resource for creating predictive models (Kuzilek et al. 2017). Researchers employing OULAD found that behavioral factors including login frequency, assignment submission timeliness, and quiz attempts can accurately predict dropout and success.

Comparisons with different datasets, such as ASSISTments (for intelligent tutoring systems), MOOCs (for large-scale online courses), and institutional LMS logs, reveal that model transferability is significantly dependent on dataset properties (Joksimović et al., 2015). While clickstream elements are predictive in MOOCs, structured assessment data has a greater impact in OULAD. This highlights the challenge of generalizing across contexts and the risk of developing models that are overly specialized to specific situations (Sclater, 2015).

To overcome this, benchmark challenges and shared tasks are being developed to standardize evaluation and enable fair comparisons. These initiatives highlight sample biases, stimulate the use of different datasets, and promote open reporting of experimental techniques (Lakkaraju et al., 2017). Without such common benchmarks, it is impossible to determine whether prediction advances are due to better algorithms or merely to changes in data quality.

Predictive Modeling Approaches for At-Risk Detection

Supervised machine learning is the most utilized paradigm for predicting student performance and identifying at-risk kids. Traditional methods like logistic regression and decision trees provide interpretability, but they may struggle with nonlinear and multidimensional data (Chen & Guestrin, 2016). Ensemble techniques like Random Forests and XGBoost have emerged as viable alternatives, frequently beating simpler models due to their ability to capture complicated feature interactions (Brownlee, 2016).

Deep learning techniques including recurrent neural networks (RNNs) and temporal convolutional networks have been used on sequential VLE data to identify time-dependent behavioral patterns (Kovanović et al., 2017). These methods beat traditional models in situations with large amounts of sequential data, such as MOOCs, but they have trade-offs in terms of transparency and processing cost (Thompson and Baker, 2019).

Hybrid techniques that combine domain expertise (such as retention theories) with data-driven strategies are becoming increasingly prevalent. These models achieve a mix between prediction accuracy and interpretability, allowing researchers and educators to understand not just whether a student is at risk, but also why the algorithm came to that conclusion (Brown, 2020). Such strategies combine predictive modeling with real-world educational objectives, focusing on explainability and trust rather than accuracy.

Feature Engineering: VLE Interactions, Temporal Signals, and Assessments

Feature engineering remains a critical component of efficient predictive modeling in education. Early involvement, such as accessing the VLE within the first week or tackling the first assignment early, is consistently linked to better outcomes (Arnold & Pistilli, 2012). Timing-related factors, such as quiz submission delays or gaps in login time, serve as indicators of self-regulation and persistence (Lockyer et al., 2014).

Advanced techniques focus on temporal feature aggregation, distinguishing between short bursts of activity and persistent engagement that span weeks. Content-type characteristics, such as interactions with videos versus discussion forums, offer explanatory power, as different resource types may represent different learning strategies (Perez-Sanchez et al., 2019).

To increase interpretability, technologies like SHAP (SHapley Additive Explanations) rank feature importance, which allows educators to identify which behaviors most strongly predict risk (Lundberg & Lee, 2017). This not only improves model openness, but it also informs practical adjustments like encouraging early submission or increasing forum participation.

Social Computing & Social Learning Analytics.

Social computing takes analytics beyond individual behavior and captures the structure and dynamics of student interactions. Indicators of social network analysis (SNA), such as degree centrality, betweenness, and clustering coefficients, have been linked to perseverance and academic achievement. Degree centrality, betweenness centrality, clustering coefficient, and interaction frequency are examples of social network analysis (SNA) metrics that provide more information about student involvement and participation in learning environments. Degree centrality identifies highly connected learners, while betweenness centrality highlights students who serve as bridges between learning groups. Clustering metrics reveal collaborative learning communities and peer interaction structures. These metrics are educationally relevant because students with stronger social participation and collaborative engagement are often associated with improved academic persistence, motivation, and learning performance. Conversely, socially isolated learners may demonstrate lower engagement and increased academic risk. For example, students who are prominently positioned in discussion networks tend to perform better, but those on the periphery are more likely to drop out (Dawson, 2017).

Research on online forums and collaborative tools indicates that well-connected learners benefit from peer assistance, exposure to varied perspectives, and faster issue resolution (Brooks & Thompson, 2018). In contrast, solitary learners face hurdles to persistence, underscoring the need to develop collaborative activities that foster network integration.

Community detection methods (such as k-means and Louvain) identify clusters of learners with similar interaction profiles (Blondel et al., 2008). These clusters can inform varied approaches, such as motivating disengaged clusters and offering advanced resources to highly active groups (Fortunato & Hric, 2016). Thus, social learning analytics improves predictive modeling by embedding students in their social contexts.

Visual Analytics & Instructor Dashboards

Dashboards serve an important role in converting complex model outputs into actionable information for teachers. Effective dashboards prioritize clarity, actionable alerts, and drill-down capabilities for monitoring individual student trajectories (Verbert et al., 2014). According to research, dashboards with simple representations (traffic lights and risk flags) supplemented with contextual content (assessment history and activity breakdown) were more widely accepted (Conijn et al., 2017).

Experimental studies have revealed that dashboards can enhance teachers' proactive outreach, particularly when combined with automated intervention suggestions (Ifenthaler & Yau, 2018). However, usability studies emphasize the importance of trust: educators are hesitant to act on dashboards they do not fully comprehend (Skiba & Deacon, 2020).

Transparency, adaptability, and consistency with teachers' existing processes are therefore key design elements. Dashboards should not overload users with data but rather highlight the most important indicators and explain their meaning, enabling informed and timely interventions.

Early-Warning Systems and Intervention Strategies

Early warning systems (EWS) use predictive analytics to provide actionable student support. The approaches range from simple rule-based alerts (e.g., low login frequency) to sophisticated triage systems that incorporate various signals (Macfadyen & Dawson, 2010). Evidence suggests that timely interventions, such as targeted emails, individualized comments, or academic coaching, can considerably increase retention and pass rates (Kramer et al., 2019).

However, effectiveness varies according to how interventions are designed and implemented. For example, generic "nudge" emails may have limited impact, whereas individualized coaching tailored to student needs produces greater results. Institutional capacity also matters; anticipatory alerts are only useful if personnel have the necessary resources and training to respond (Wise & Jung, 2020). This highlights the importance of incorporating predictive technologies into larger institutional support frameworks rather than deploying them in isolation.

Explainable AI (XAI) & Interpretability

As predictive models become increasingly complex, explainability becomes even more crucial. SHAP, LIME, and counterfactual explanations are techniques that help users understand which features contributed the most to a given forecast (Ribeiro et al., 2016; Ribeiro et al., 2018). These technologies can help teachers identify why a student is identified as at risk, thereby bridging the gap between data science and pedagogy.

However, issues remain. Explanations can place a cognitive strain on instructors, who may lack the technical skills to comprehend them successfully. Furthermore, the validity of post-hoc explanations is debatable; some claim that they oversimplify model behavior, thereby leading to false interpretations (Ross et al., 2021).

Despite these limitations, XAI is a promising avenue for increasing trust and adoption. Including explanations in dashboards—via plain language, simple visualizations, or scenario-based counterfactuals—can make risk forecasts more actionable for educators, thereby increasing the effect of predictive analytics.

Ethics, Privacy, and Responsible Learning Analytics

The use of learning analytics requires careful consideration of ethics and privacy. Students are typically unaware of the amount to which their data is recorded, raising questions regarding informed consent and autonomy (Selwyn and Facer, 2017). Best practices prioritize transparency in data use, opt-out options, and the idea of data minimization—collecting only what is necessary to encourage learning (Williamson & Piattoeva, 2019).

Beyond privacy, algorithmic fairness is a critical issue. Predictive methods may unintentionally reproduce prejudices, labeling minority or non-traditional students as "at risk," resulting in stigma and unequal treatment (Drachsler & Greller, 2011). Fairness-aware modeling, bias audits, and the use of diverse datasets are all strategies for mitigating such risks.

The DELICATE framework (Determination, Explain, Legitimate, Involve, Consent, Anonymize, Technical, and External) offers structured guidelines for responsible deployment (Drachsler & Greller, 2011). Finally, human-in-the-loop approaches remain essential, ensuring that analytics supplement but do not replace educator discretion. Integrating ethical precautions into design and implementation boosts trust and guarantees that analytics benefit students.

Evaluation Metrics, Validation, and Methodological Concerns

Evaluation procedures have a considerable impact on the results of predictive analytics. While accuracy remains a popular metric, it may be misleading in imbalanced datasets with a small proportion of at-risk kids. Precision, recall, F1-score, ROC-AUC, and PR-AUC are metrics that allow for more extensive evaluations (Powers 2011).

Methodological rigor is also important. Cross-validation procedures should take into consideration the temporal structure of the data; otherwise, models may unintentionally "peek into the future," thereby boosting performance estimates (Caballero & Penna, 2019). Selective-label bias, in which outcomes are only observed for a subset of students (e.g., those who pass tests), complicates validation and must be addressed directly.

Finally, the external validity of models is a recurring issue. Dataset shift—when models trained in one setting perform poorly in another—is a common observation (Johnson & Evans, 2020). The demand for cross-institutional validation and open benchmarking is increasing, ensuring that predictive tools are not only accurate but also portable and generalizable (Kohn & Hsieh, 2020).

Emerging Integrated Learning Analytics Systems

Recent studies increasingly emphasize integrating predictive analytics, social learning analytics, and dashboard visualization into unified educational systems. Arnold and Pistilli (2012) demonstrated how early alert analytics embedded in course dashboards significantly improved student persistence by enabling instructors to intervene based on behavioral signals rather than final grades. Their findings highlight that predictive analytics are most effective when tightly coupled with instructor-facing decision tools.

Similarly, Joksimović et al. (2015) explored multimodal learning analytics frameworks that combine clickstream analysis, assessment performance, and social interaction data. Their work revealed that hybrid models outperform single-source analytics by capturing both cognitive and social dimensions of learning. This supports the present study's integration of social computing with predictive modeling.

Ifenthaler and Yau (2020) examined intelligent analytics dashboards in blended learning environments and found that educators who received interpretable risk indicators demonstrated higher intervention frequency and improved instructional alignment. Their study reinforces the importance of explainable analytics and dashboard usability in promoting adoption.

More recent work by Viberg, Hatakka, and Mavroudi (2018) proposed a socio-technical learning analytics model that bridges algorithmic prediction with pedagogical context. They argue that analytics systems must incorporate human-centered interpretation layers to avoid purely technical decision-making. This aligns with the current study's goal of transforming predictive outputs into actionable instructional insights.

Collectively, these studies demonstrate a growing shift toward holistic analytics ecosystems that integrate predictive modeling, social engagement analysis, and visualization into a coherent framework supporting early intervention. The proposed intelligent class record system contributes to this evolving direction by embedding predictive and social analytics directly into the classroom record infrastructure.

3. Methodology

This chapter presents the research methodology employed in the design, development, and evaluation of the proposed Social Computing-Based Intelligent Class Record System for the early detection of at-risk students. It discusses the research design, data sources, system development process, data preprocessing procedures, predictive modeling techniques, social computing integration, dashboard development, evaluation procedures, and statistical treatment used in the study. The methodology was structured to ensure the systematic transformation of educational data into predictive and actionable insights that support proactive academic intervention. The methodology was derived from the objectives and framework presented in the dissertation.

3.1 Research Design

This study utilized a developmental research design integrated with quantitative analytical methods. Developmental research was appropriate because the study focused on designing, developing, and evaluating an intelligent class record system that integrates learning analytics, social computing, and artificial intelligence-based predictive modeling. The developmental approach enabled the systematic creation of a functional prototype while simultaneously evaluating its predictive capability and usability. Quantitative methods were applied in analyzing educational datasets, training predictive models, and evaluating system usability using numerical metrics. The study incorporated machine learning, social network analysis, and dashboard visualization techniques to transform educational data into actionable instructional insights. The research process involved data collection, preprocessing, predictive model development, dashboard implementation, system integration, and usability evaluation.

3.2 Data Source and Dataset Description

The study utilized the Open University Learning Analytics Dataset (OULAD) as the primary source of educational data. OULAD is a publicly available dataset containing demographic, assessment, registration, and Virtual Learning Environment (VLE) interaction data from more than 32,000 students enrolled in online courses. The dataset was selected because of its comprehensive representation of learner engagement, academic performance, and behavioral interaction patterns in digital learning environments.

The dataset includes the following components:

- Student demographic information such as age, gender, education level, and region;
- Assessment records including quizzes, assignments, and examination scores;
- Registration data reflecting enrollment and withdrawal information;
- VLE interaction logs containing clickstream and engagement behaviors;
- Student outcome classifications including Pass, Fail, Withdraw, and Distinction.

The integration of these multidimensional datasets enabled the system to generate predictive insights related to student performance and academic risk.

3.3 System Development Process

The development of the proposed intelligent class record system followed an iterative system development process composed of five major phases:

- Data Collection and Preparation
- Data Preprocessing and Feature Engineering
- Predictive Model Development
- Dashboard and Visualization Development
- System Evaluation and Validation

The system architecture integrates learning analytics, machine learning, and social computing into a unified platform capable of monitoring student engagement and identifying at-risk learners. The overall workflow begins with dataset preprocessing, followed by predictive analytics, social interaction analysis, and dashboard visualization.

3.4 Data Preprocessing and Feature Engineering

Data preprocessing was conducted to ensure the quality, consistency, and reliability of the dataset prior to predictive modeling. Missing values, duplicate entries, and inconsistent records were identified and handled through data cleaning procedures. Categorical variables were encoded into numerical representations suitable for machine learning algorithms.

Feature engineering was then performed to derive meaningful indicators associated with student engagement and academic performance. Engineered features included:

- Login frequency;
- Assessment submission timeliness;
- Average assessment scores;
- Interaction frequency;
- Engagement duration;
- Forum participation;
- Resource access behavior.

Temporal engagement patterns were also analyzed to identify early warning indicators associated with academic risk. Feature scaling and normalization techniques were applied to improve model convergence and predictive performance.

3.5 Predictive Modeling Techniques

The study implemented multiple machine learning algorithms to classify students into Pass, Fail, Withdraw, and Distinction categories. The selected algorithms were chosen based on their established effectiveness in educational data mining and learning analytics research.

The predictive models utilized in the study included:

3.5.1 Logistic Regression

Logistic Regression was used as a baseline classification model due to its interpretability and suitability for educational prediction tasks. The model estimates the probability of student outcomes based on demographic, assessment, and engagement variables.

3.5.2 Random Forest

Random Forest was implemented as an ensemble learning technique capable of handling nonlinear relationships and complex feature interactions. The model generates multiple decision trees and aggregates predictions to improve accuracy and reduce overfitting.

3.5.3 Neural Network

A multilayer neural network architecture was utilized to capture hidden nonlinear relationships within the dataset. The neural network model processed engagement and assessment features to identify complex learning behavior patterns associated with academic outcomes.

3.5.4 Extreme Gradient Boosting (XGBoost)

XGBoost was implemented due to its efficiency, scalability, and strong predictive performance in classification problems. The model utilizes gradient boosting mechanisms to iteratively improve prediction accuracy through sequential tree optimization.

3.6 Model Training and Evaluation

The dataset was divided into training and testing subsets to evaluate model generalization performance. Cross-validation techniques were employed to minimize bias and ensure model reliability.

The predictive models were evaluated using the following performance metrics:

- Accuracy;
- Precision;
- Recall;
- F1-Score;
- Area Under the Curve (AUC).

These metrics provided a comprehensive assessment of predictive capability, especially in identifying at-risk students. Comparative analysis was performed to determine which algorithm achieved the highest predictive performance.

3.7 Social Computing and Social Network Analysis

To enhance contextual understanding of learner engagement, the study incorporated social computing techniques through Social Network Analysis (SNA). Social interaction metrics were derived from student communication and engagement patterns.

The following SNA metrics were utilized:

- Degree Centrality;

- Betweenness Centrality;
- Clustering Coefficient;
- Interaction Frequency.

These metrics enabled the identification of socially isolated learners, collaborative clusters, and highly connected students. The integration of social interaction data provided a more comprehensive understanding of academic persistence and learner engagement beyond assessment performance alone.

3.8 Dashboard Design and System Interface

The proposed intelligent class record system includes an interactive analytics dashboard designed for educators. The dashboard was developed to provide intuitive visualization of predictive outputs, engagement trends, and intervention recommendations.

The dashboard features include:

- Student risk classification;
- Engagement trend visualization;
- Assessment performance monitoring;
- Social interaction analytics;
- Automated intervention recommendations;
- Class-level and student-level analytics.

Visual elements such as charts, graphs, progress indicators, and risk flags were incorporated to improve interpretability and usability. The interface was designed to support evidence-based instructional decision-making while minimizing information overload.

3.9 System Integration

The proposed intelligent class record system follows an integrated architecture in which educational data flows through interconnected analytical components. Student demographic, assessment, and engagement data are first collected and preprocessed to ensure quality and consistency. The processed dataset is then transmitted to machine learning models responsible for predictive classification. Prediction outputs are subsequently combined with social computing metrics to generate comprehensive student risk profiles. Finally, all analytical outputs are integrated into an interactive dashboard interface that enables educators to monitor performance trends, interpret risk indicators, and initiate evidence-based interventions.

The system integration process consists of the following interconnected modules:

- OULAD Data Integration Layer
- Data Preprocessing and Feature Engineering Pipeline
- Machine Learning Prediction Module
- Social Computing and SNA Module
- Dashboard Visualization Layer
- Intervention Recommendation Engine

This integrated structure ensures seamless data flow between analytical processes and instructional visualization components.

3.10 Explainable Artificial Intelligence Integration

To improve transparency and interpretability, the study incorporated Explainable Artificial Intelligence (XAI) mechanisms within the predictive framework. Feature importance analysis was utilized to identify the variables contributing most significantly to predictive outcomes.

Explainability techniques enabled educators to understand:

- Why students were classified as at-risk;
- Which engagement indicators influenced predictions;
- How behavioral patterns affected academic outcomes.
- This approach strengthened trust in predictive outputs and supported more informed instructional interventions.

3.11 System Evaluation

The developed system was evaluated using the System Usability Scale (SUS) among selected educators. SUS is a standardized usability assessment tool widely used for evaluating interactive systems and educational technologies.

Participants evaluated the system based on:

- Ease of use;
- Interface clarity;
- Learnability;
- System responsiveness;
- Overall usability;
- Perceived usefulness.

Responses were analyzed quantitatively to determine the overall usability score of the developed intelligent class record system.

3.12 Statistical Treatment of Data

The collected data and model evaluation results were analyzed using descriptive and inferential statistical techniques. Descriptive statistics such as frequency, percentage, mean, and standard deviation were used to summarize dataset characteristics and usability responses. Comparative analysis of predictive models was conducted using performance metrics including accuracy, precision, recall, F1-score, and AUC values. Cross-validation results were also analyzed to determine model stability and generalization performance.

3.13 Ethical Considerations

The study adhered to ethical standards in handling educational data and predictive analytics. The OULAD dataset used in the study is publicly available and anonymized, ensuring that no personally identifiable information was exposed during analysis.

The study also considered ethical concerns related to predictive learning analytics, including:

- Data privacy;
- Responsible AI usage;
- Transparency of predictive outputs;
- Avoidance of algorithmic bias;
- Ethical interpretation of at-risk classifications.

These safeguards ensured that the proposed system supports responsible and student-centered educational analytics practices.

4. Results and Discussion

Figure 1. Log In Page

Figure 1. The login page is the entrance point for the proposed social computing-based intelligent class record system. It is intended to enable secure and role-based access to authorized users, namely educators and system administrators. This module allows users to validate their credentials before accessing the analytics dashboard and system features. The login process contributes to privacy, secrecy, and safeguarding of critical academic information. The solution prevents illegal access while also ensuring the integrity of predictive analytics outputs and student records.

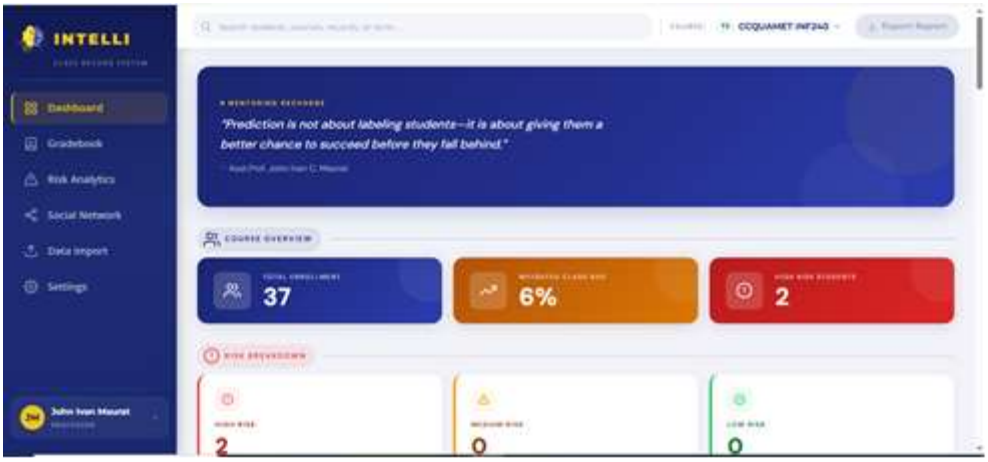

Figure 2. Dashboard

Figure 2. The dashboard functions as the central interface of the system, presenting summarized analytics and real-time insights regarding student performance, engagement, and risk distribution. It provides educators with an overview of important indicators such as at-risk student counts, engagement trends, predictive classifications, and intervention alerts. The dashboard utilizes visual analytics components such as charts, graphs, and summary cards to transform complex educational data into interpretable information. Through this interface, educators can quickly monitor class performance and make timely, evidence-based instructional decisions.

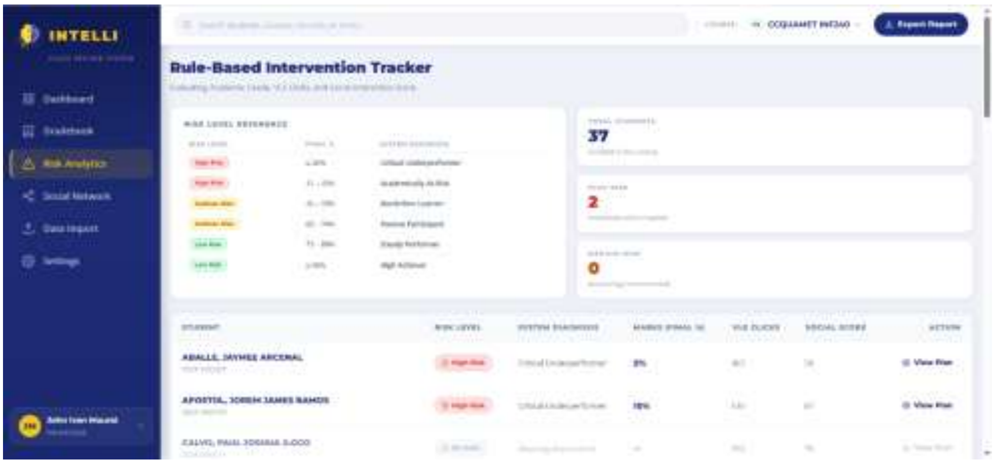

Figure 3. Student Risk Profile

Figure 3. The Student Risk Profile module offers thorough data for individual students using predictive modeling and engagement analysis. This module presents data on academic performance trends, engagement metrics, prediction classifications (Pass, Fail, Withdraw, or Distinction), and computed risk levels classified as Low, Moderate, or High Risk. Explainable AI outputs, such as feature importance indicators, are also included to help instructors comprehend the aspects that influence prediction results. The module enables instructors to identify difficult students early and conduct targeted interventions before academic deterioration worsens.

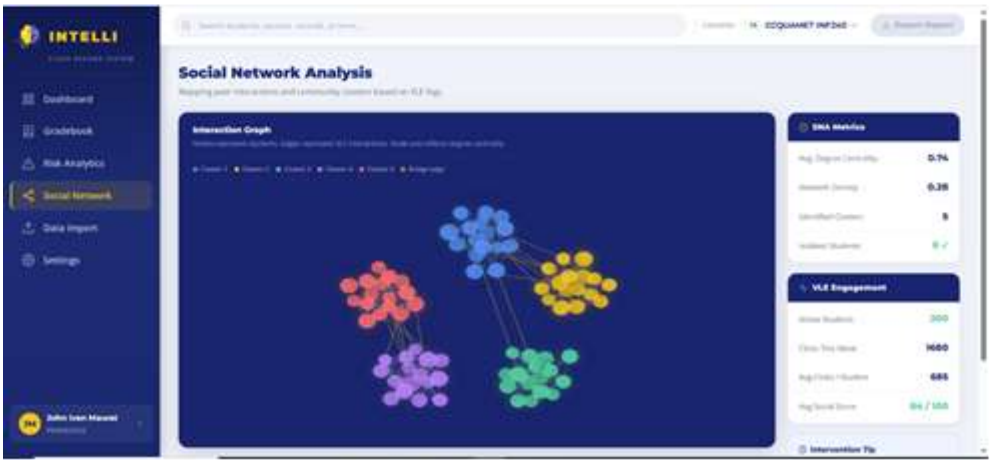

Figure 4. Social Cluster Visualization

Figure 4. The Social Cluster Visualization module uses Social Network Analysis (SNA) to show patterns of peer contact and engagement among students. The system uses graphical representations such as node-link diagrams and interaction clusters to identify highly linked learners, isolated students, and collaborative learning communities. Metrics such as centrality, clustering, and interaction frequency are utilized to analyze social participation within the learning environment. This module supports educators in understanding how social engagement influences academic performance and persistence, allowing them to encourage collaboration and improve student integration within learning communities.

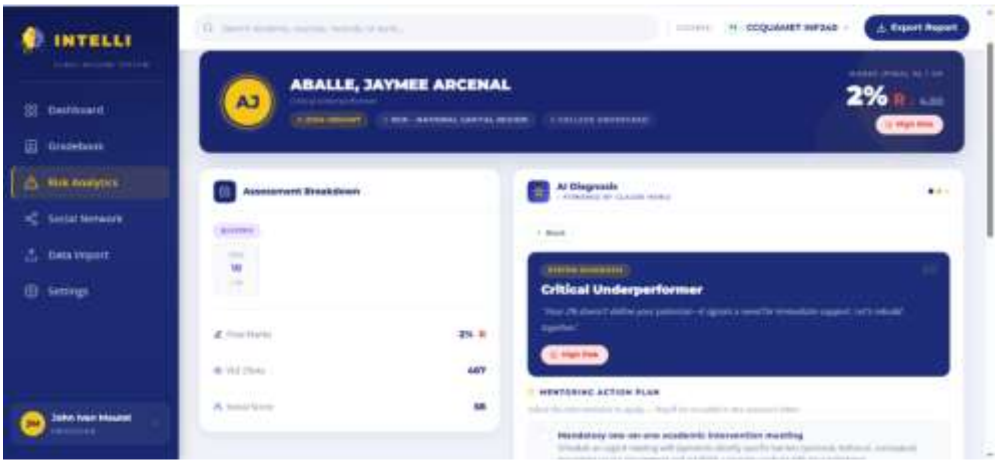

Figure 5. Intervention Module

Figure 5. The Intervention Module generates automated, rule-based recommendations intended to support at-risk students identified through predictive analytics. Low engagement levels, poor assessment performance, inconsistent participation, and social isolation indicators all play a role in the recommendations. Academic counseling, enhanced monitoring, tutoring support, consulting meetings, and engagement activities are some of the suggested approaches. The module is designed to assist educators in implementing proactive and data-informed support strategies while ensuring that final intervention decisions remain under educator supervision and professional judgment.

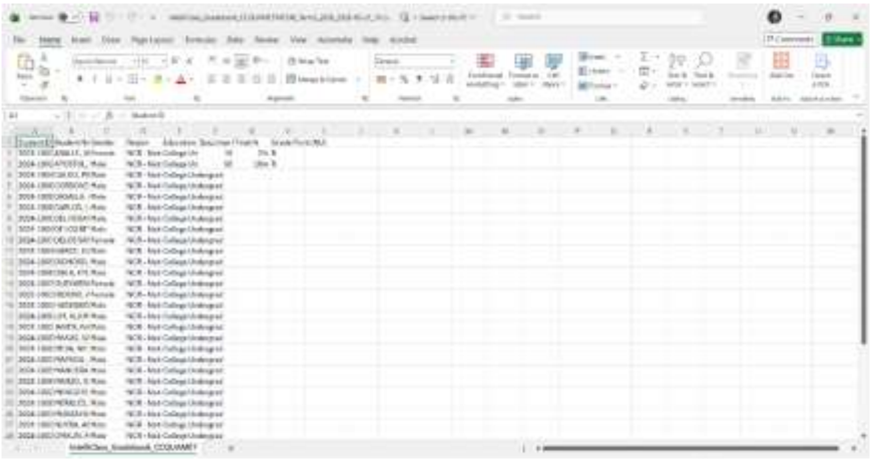

Figure 6. Report Export

Figure 6. The Report Export module enables educators and administrators to generate and download analytical reports from the system in formats such as PDF or CSV. Reports may include predictive classifications, engagement summaries, risk analytics, intervention records, and dashboard visualizations. This functionality supports documentation, academic monitoring, institutional reporting, and data sharing for decision-making purposes. By allowing exportable analytics outputs, the system enhances accessibility, record management, and communication of student performance information within the institution.

Figure 7. Engagement Trends Visualization

Figure 7. presents the engagement trends visualization generated by the system based on student interaction data and learning activity patterns. The dashboard displays changes in student engagement over time, including login frequency, assessment participation, and activity completion behavior. Students with declining engagement trends are visually highlighted, enabling educators to identify potential academic risks early. Visualization supports proactive monitoring by allowing instructors to observe behavioral patterns that may indicate disengagement or reduced academic participation.

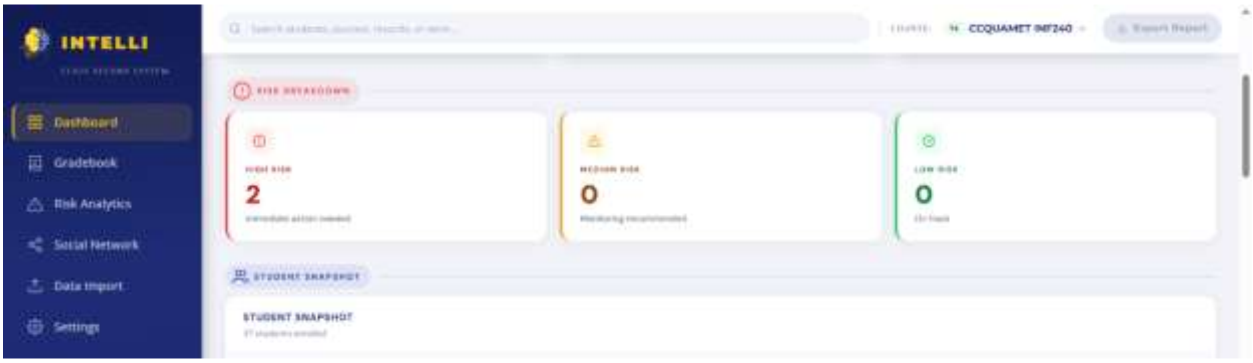

Figure 8. Risk Distribution Analytics

Figure 8. depicts the distribution of students according to expected risk classifications given by machine learning models. The dashboard divides learners into low-risk, moderate-risk, and high-risk categories based on assessment performance, engagement metrics, and behavioral indicators. Visualization enables instructors to easily determine the proportion of students that require academic support and intervention. The presence of high-risk learners emphasizes the value of predictive analytics in assisting with early diagnosis and retention efforts.

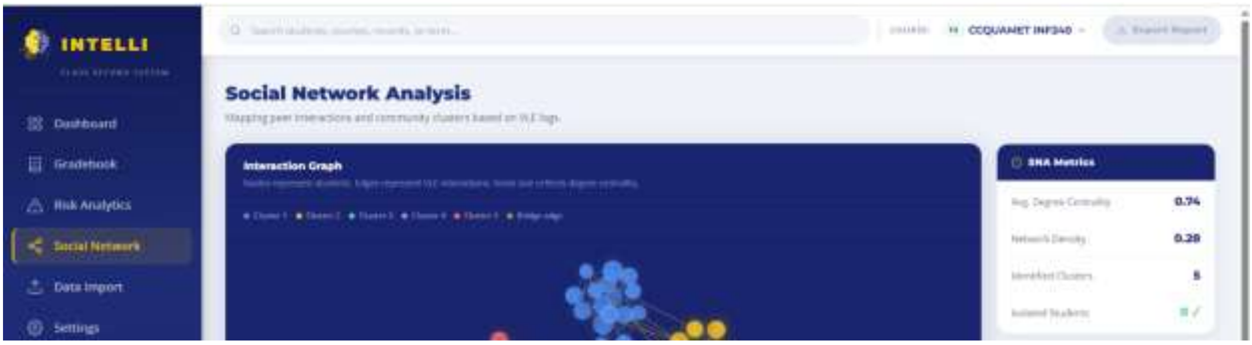

Figure 9. Social Network Graph Visualization

Figure 9. presents the Social Network Analysis (SNA) visualization generated by the system. The graph illustrates interaction patterns among students based on engagement and collaborative activities within the learning environment. Connected nodes represent active peer interactions, while isolated nodes indicate students with limited participation or weak social integration. Visualization enables educators to identify learning clusters, highly engaged students, and socially disconnected learners who may require additional academic or motivational support.

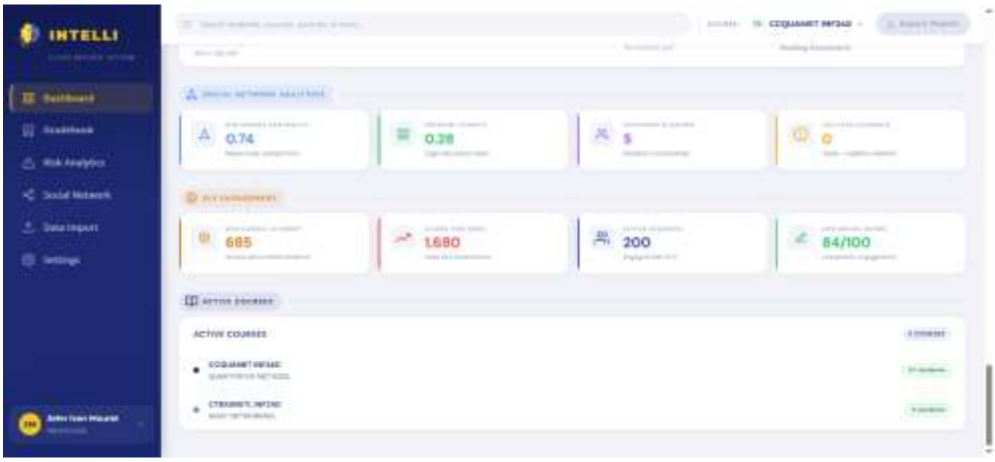

Figure 10. Performance Analytics Dashboard

Figure 10. shows the performance analytics dashboard of the developed system. The dashboard summarizes student academic performance through assessment scores, engagement indicators, predictive classifications, and performance trends. Analytical visualizations provide educators with a comprehensive overview of class standing and student progression. By consolidating predictive and behavioral analytics into a unified interface, the system supports data-driven instructional monitoring and intervention planning. Based on the parameters utilized in the study, at-risk students commonly demonstrated characteristics such as low assessment scores, declining engagement trends, inconsistent participation in learning activities, limited interaction within social learning networks, and reduced virtual learning environment activity. Students identified as socially isolated or exhibiting weak peer connectivity also showed increased likelihood of academic risk.

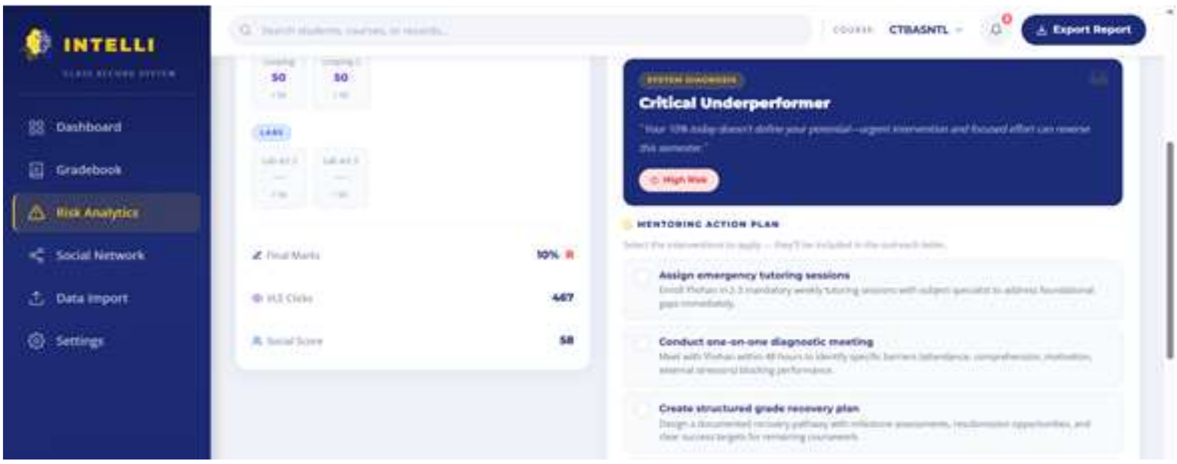

Figure 11. Sample Intervention Recommendation Output

Figure 11. presents a sample intervention recommendation generated by the proposed Social Computing-Based Intelligent Class Record System. The output displays the student's profile information, predicted risk classification, engagement indicators, and the corresponding recommended intervention generated by the system based on predictive analytics and behavioral analysis.

The visualization includes the student's identified risk level, categorized as Low Risk, Moderate Risk, or High Risk, together with supporting indicators such as assessment performance, engagement trends, and participation behavior. Based on these analytics outputs, the system automatically generates evidence-based intervention recommendations intended to assist educators in providing timely academic support.

For students classified under the High-Risk category, the system may recommend immediate academic counseling, instructor consultation, personalized monitoring, or additional learning support activities. Moderate-risk students may receive recommendations related to progress monitoring, supplemental exercises, and periodic feedback sessions, while low-risk students are encouraged to maintain consistent academic participation and performance.

Figure 12. Bangor SUS Grade Chart and Percentile Visualization

Figure 12. presents the Bangor System Usability Scale (SUS) grade interpretation and percentile visualization of the developed system. The chart indicates an overall SUS score of 91.25, which corresponds to a Grade A+ under the Bangor grading scale. This score falls within the "Excellent" adjective rating and lies in the top 10% percentile of systems evaluated using the SUS framework.

The figure further illustrates the acceptability ranges, where the system is positioned within the "Acceptable" category, signifying a high level of usability and user satisfaction. The grade scale and adjective ratings reinforce that the system exceeds standard usability benchmarks, achieving near-optimal performance. Additionally, the percentile rank visualization highlights that the system outperforms most comparable systems, emphasizing its strong usability characteristics.

This study developed and evaluated a social computing-based intelligent class record system meant to discover at-risk kids early on by combining learning analytics, machine learning, and social computing techniques. The study addressed the constraints of standard class record systems, which primarily serve as grade and attendance records and lack predictive, analytical, and intervention-support capabilities.

Based on learning analytics theory, educational data mining, Tinto's student retention theory, and Vygotsky's social constructivism, the study offered an intelligent system capable of converting raw educational data into usable pedagogical insights. The system used the Open University Learning Analytics Dataset (OULAD), which contains student demographic information, assessment records, and virtual learning environment engagement data, to predict academic performance and identify risk-related behaviors.

The study followed a developmental and experimental research design that included the creation, implementation, and evaluation of a functional system prototype. Multiple machine learning methods, such as logistic regression, random forest, neural network, and XGBoost, were combined and tested to establish the most effective predictive method for detecting at-risk students. The results showed that XGBoost had the highest prediction accuracy of the models tested, demonstrating the efficacy of ensemble-based machine learning techniques in dealing with multidimensional educational datasets.

In addition, the study incorporated Social Network Analysis (SNA) to examine peer interaction patterns, engagement behaviors, and learning communities that may influence academic outcomes. The developed system was implemented through an interactive dashboard capable of visualizing student risk levels, engagement trends, social interaction clusters, and automated intervention recommendations intended to support proactive instructional decision-making.

To assess usability and acceptability, 30 IT specialists and educators were given the System Usability Scale (SUS) and supplemental evaluation measures. The results showed a high usability rating, with an overall SUS score of 91.25 and strong internal consistency, as shown by Cronbach’s alpha value of 0.93. These findings show that the designed system is extremely useable, dependable, and suitable for possible educational implementation.

5. Conclusion and Recommendations

CONCLUSIONS

Based on the results of the investigation, the following conclusions were drawn:

1. The developed Social Computing-Based Intelligent Class Record System successfully combined learning analytics, machine learning, social computing, and dashboard visualization to create a unified educational decision-support platform capable of identifying at-risk students using predictive analytics and engagement monitoring.
2. The integration of machine learning algorithms demonstrated that predictive analytics can effectively support early detection of at-risk learners. Among the evaluated algorithms, XGBoost achieved the highest predictive accuracy, indicating that ensemble-based learning approaches are highly effective for analyzing multidimensional educational datasets.
3. The incorporation of Social Network Analysis (SNA) provided additional insights into student engagement and peer interaction patterns by identifying learning communities, highly engaged learners, and socially isolated students who may require academic support and intervention.
4. The developed dashboard analytics and visualization modules successfully transformed complex predictive outputs into interpretable and educator-friendly insights, thereby improving accessibility, instructional monitoring, and evidence-based decision-making.
5. The automated intervention recommendation module demonstrated the potential of predictive systems to support proactive academic interventions by generating risk-based recommendations aligned with student behavioral and engagement indicators.

6. The usability evaluation results revealed that the developed system met high standards of usability, dependability, and user acceptance. The derived SUS score of 91.25 and Cronbach's alpha value of 0.93 indicate that the system is intuitive, efficient, and suitable for possible use in educational settings.
7. Overall, the findings support the effectiveness of integrating predictive analytics and social computing into an intelligent class record system to aid in early detection, academic monitoring, student engagement analysis, and proactive intervention strategies aimed at improving student retention and academic success.

The study's findings underline the necessity of incorporating predictive analytics and social computing into educational systems to enhance proactive and evidence-based teaching approaches. Early identification of disengagement and academic risk allows educators to execute timely interventions that promote student participation, retention, and academic performance. Furthermore, the use of social interaction analytics emphasizes the importance of collaboration and peer participation in determining positive learning outcomes in online and blended learning contexts.

RECOMMENDATIONS

Based on the findings and limitations of the study, the following recommendations are proposed:

1. For Educational Institutions

Institutions should consider adopting intelligent class record systems that integrate predictive analytics and social computing to enhance early warning mechanisms. Investment in data-driven tools can improve student retention strategies and academic support services.

2. For Educators

Educators are encouraged to utilize analytics dashboards as decision-support tools to monitor student engagement and performance. However, predictive output should be used in conjunction with professional judgment, ensuring that interventions remain personalized and context sensitive.

3. For System Development

- Future enhancements of the system may include:
- Integration with existing Learning Management Systems (LMS)
- Real-time data processing for continuous monitoring
- Improved explainable AI (XAI) features for better interpretability
- Mobile-friendly dashboard interfaces for accessibility

4. For Future Researchers

Future studies may:

- Conduct real-world implementation and longitudinal evaluation of the system.
- Explore additional datasets beyond OULAD to improve generalizability.
- Investigate the impact of intervention strategies triggered by predictive systems.
- Incorporate affective and behavioral data (e.g., motivation, sentiment analysis)

5. For Policymakers

Policymakers should promote the development and ethical use of learning analytics systems by establishing guidelines on:

- Data privacy and security
- Fairness and bias mitigation

- Responsible use of predictive models in education

6. For Ethical Considerations

- Future implementations must ensure:
- Transparency in how predictions are generated
- Protection of student data through anonymization
- Avoidance of labeling or stigmatization of at-risk students

Although the study used the Open University Learning Analytics Dataset (OULAD), the proposed framework has potential for use in a variety of educational contexts because it is based on common educational indicators such as assessment performance, engagement behavior, and participation metrics. However, institutional differences in curriculum structure, learning management systems, and student behavior patterns may have an impact on predictive performance; thus, retraining and contextual calibration with localized datasets are recommended prior to large-scale deployment.

References

1. Arnold, K., & Pistilli, M. (2012). Course signals at Purdue: Using analytics to increase student success. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK'12) (pp. 267–270). ACM. https://doi.org/10.1145/2330601.2330666
2. Blondel, V., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008
3. Brooks, M. T., & Thompson, S. T. (2018). Forum participation and retention: Social and cognitive presence. Journal of Online Learning and Teaching, 14(3), 212–229.
4. Brown, D. M. B. (2020). Hybrid models combining domain knowledge and ML for student performance prediction. Computers & Education, 150, 103850. https://doi.org/10.1016/j.compedu.2020.103850
5. Brownlee, J. (2016). A gentle introduction to gradient boosting. Machine Learning Mastery. https://machinelearningmastery.com/gentle-introduction-gradient-boosting/
6. Caballero, J., & Penna, M. (2019). Time-aware cross-validation for student outcome prediction. ACM Transactions on Learning Technologies, 12(2), 1–17. https://doi.org/10.1145/3304114
7. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16) (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785
8. Conijn, N. V., van den Beemt, A., & Cuijpers, P. (2017). Designing for teacher use: Dashboard affordances. IEEE Transactions on Learning Technologies, 10(2), 170–181. https://doi.org/10.1109/TLT.2016.2599522
9. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008
10. Dawson, G. M. (2017). A study of the relationship between student social networks and academic outcomes. Computers & Education, 106, 16–27. https://doi.org/10.1016/j.compedu.2016.11.001
11. Drachsler, P., & Greller, W. (2011). Privacy and learning analytics—It's a DELICATE issue. In Proceedings of the 1st International Conference on Learning Analytics and Knowledge (LAK'11) (pp. 89–98). ACM. https://doi.org/10.1145/2090116.2090129
12. Ferguson, R. (2012). The state of learning analytics in 2012: A review and future challenges (Tech.

Rep. KMI-12-01). Knowledge Media Institute, Open University.
13. Ferguson, R., & Clow, D. (2017). Where is the evidence? A call to action for learning analytics. In Proceedings of the 7th International Conference on Learning Analytics and Knowledge (LAK'17) (pp. 56–65). ACM. https://doi.org/10.1145/3027385.3027396
14. Fortunato, A., & Hric, D. (2016). Community detection in networks: A user guide. Physics Reports, 659, 1–44. https://doi.org/10.1016/j.physrep.2016.09.002
15. Goggins, D. R., Jahnke, I., & Wulf, V. (2018). Social network analysis for learning analytics. IEEE Transactions on Learning Technologies, 10(3), 295–305. https://doi.org/10.1109/TLT.2016.2619697
16. Ifenthaler, D., & Yau, D. (2018). Utilizing learning analytics to support students' self-regulated learning: A systematic review. Educational Technology Research and Development, 66(5), 1275–1292. https://doi.org/10.1007/s11423-018-9578-x
17. Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilizing learning analytics dashboards for learner support: A systematic review. Educational Technology Research and Development, 68(2), 1–23.
18. Johnson, B. D., & Evans, R. L. (2020). External validation and portability of predictive models in education. Educational Data Science, 2(1), 45–63.
19. Joksimović, A., Gašević, D., Dawson, S., Kovanović, V., & Baker, R. S. (2015). Learning analytics to uncover interaction patterns in online discussions: A case study. Computers in Human Behavior, 47, 42–50. https://doi.org/10.1016/j.chb.2014.10.028
20. Kohn, A., & Hsieh, L. B. (2020). Addressing dataset shift in VLE predictive modeling. Pattern Recognition Letters, 132, 46–53. https://doi.org/10.1016/j.patrec.2020.03.007
21. Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., & Baker, R. S. (2017). Predicting student performance from VLE data: A sequence-aware approach. Journal of Computer Assisted Learning, 33(6), 1–12.
22. Kramer, S. J., Kizilcec, R. F., & Xu, B. (2019). Randomized evaluation of an early intervention informed by predictive analytics. Educational Evaluation and Policy Analysis, 41(2), 1–22. https://doi.org/10.3102/0162373719838317
23. Kuzilek, K., Hlosta, M., & Zdrahal, Z. (2017). Open University Learning Analytics Dataset. Scientific Data, 4(1), 170171. https://doi.org/10.1038/sdata.2017.171
24. Lakkaraju, H., Kleinberg, J., Leskovec, J., Ludwig, J., & Mullainathan, S. (2017). The selective labels problem: Evaluating algorithmic predictions in the presence of unobservables. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'17) (pp. 275–284). ACM. https://doi.org/10.1145/3097983.3098066
25. Lockyer, L., Heathcote, E., & Dawson, S. (2014). Informing pedagogical action: Aligning learning analytics with learning design. Australasian Journal of Educational Technology, 30(1), 1–17. https://doi.org/10.14742/ajet.482
26. Lundberg, S., & Lee, S. (2017). A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17) (pp. 4765–4774). ACM.
27. Macfadyen, K., & Dawson, S. (2010). Mining LMS data to develop an early warning system for educators: A proof of concept. Computers & Education, 54(2), 588–599. https://doi.org/10.1016/j.compedu.2009.09.008
28. National Center for Education Statistics. (2020). Undergraduate retention and graduation rates. U.S. Department of Education. https://nces.ed.gov/programs/coe/indicator/ctr

29. Perez-Sanchez, F., Romero, C., & Ventura, S. (2019). Feature engineering for student modeling: Temporal and content features. User Modeling and User-Adapted Interaction, 29(3), 501–528. https://doi.org/10.1007/s11257-019-09221-4
30. Powers, D. M. (2011). Evaluation: From precision, recall, and F-measure to ROC, informedness, markedness & correlation. Journal of Machine Learning Technologies, 2(1), 37–63.
31. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16) (pp. 1135–1144). ACM. https://doi.org/10.1145/2939672.2939778
32. Ribeiro, M. T., Singh, S., & Guestrin, C. (2018). Anchors: High-precision model-agnostic explanations. In Proceedings of the 32nd AAAI Conference on Artificial Intelligence (pp. 1527–1535). AAAI Press.
33. Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(6), 601–618. https://doi.org/10.1109/TSMCC.2010.2053532
34. Ross, A., Canosa, A., Williams, L., & Tomkins, S. (2021). Human-centered explainable AI for education: Challenges and opportunities. ACM Transactions on Computing Education, 21(4), 1–31. https://doi.org/10.1145/3442187
35. Sclater, N. (2015). Learning analytics in higher education. Journal of Learning Analytics, 2(1), 1–9.
36. Selwyn, N., & Facer, T. (2017). The ethics of learning analytics. IEEE Internet Computing, 21(2), 34–41. https://doi.org/10.1109/MIC.2017.40
37. Siemens, G., & Baker, R. S. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK'12) (pp. 252–254). ACM. https://doi.org/10.1145/2330601.2330661
38. Skiba, H., & Deacon, K. (2020). Instructor trust and adoption of learning analytics tools. Journal of Educational Computing Research, 58(7), 1297–1319. https://doi.org/10.1177/0735633120935903
39. Thompson, S. L., & Baker, J. T. (2019). Deep learning for clickstream sequence modeling in MOOCs. IEEE Transactions on Learning Technologies, 12(4), 481–492. https://doi.org/10.1109/TLT.2019.2919412
40. Tinto, V. (1993). Leaving college: Rethinking the causes and cures of student attrition. University of Chicago Press.
41. Tinto, V. (2012). Completing college: Rethinking institutional action. University of Chicago Press.
42. Verbert, S., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2014). Learning dashboards: An overview and future research opportunities. Personal and Ubiquitous Computing, 18(6), 1533–1547. https://doi.org/10.1007/s00779-013-0751-2
43. Viberg, O., Hatakka, M., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. Computers in Human Behavior, 89, 98–110.
44. Wise, T., & Jung, D. (2020). Intervention design best practices when using risk predictions. Learning Analytics Review, 12, 45–61.
45. Williamson, S., & Piattoeva, M. (2019). Objectivity as standardization in data-driven education governance. Research in Education, 103(1), 69–87. https://doi.org/10.1177/0034523719845003.