International Journal of Combined Research & Development (IJCRD)
eISSN:2321-225X; pISSN:2321-2241 Volume: 14; Issue: 12; December -2025

# An Intelligent Student Progress Monitoring and Early Warning System Using CIE and SEE Performance Analytics

Achutha JC
Asst. Professor,Department of Master of Computer Applications
AMC Engineering College, Bengaluru, India
achutha.sir@gmail.com

Abstract—Monitoring student academic progress traditionally depends on periodic examination results and manual identification of students who require additional support. This paper presents an intelligent student progress monitoring and early warning framework that integrates Continuous Internal Evaluation (CIE) and Semester End Examination (SEE) performance data. The proposed system consolidates assessment records, computes subject-wise and student-wise performance indicators, identifies changes in performance, and generates early warning categories for students who may require academic intervention. The framework combines descriptive analytics with predictive modeling so that CIE performance can be used as an early indicator of possible SEE outcomes. It provides dashboards and intervention-oriented reports for students, faculty members, mentors, and academic administrators. The proposed methodology includes data preprocessing, feature engineering, risk classification, model evaluation, and continuous updating as new CIE assessments become available. Since predictive results depend on an institution's historical dataset, this paper defines an experimental methodology rather than fabricating numerical results. The framework can support data-driven mentoring, targeted remedial instruction, and continuous academic quality improvement.

***Keywords***—CIE, SEE, student progress monitoring, early warning system, learning analytics, academic performance prediction, educational data mining, student risk prediction.

## I. INTRODUCTION

Higher education institutions collect substantial amounts of academic information through internal assessments, assignments, laboratory examinations, attendance records, and semester-end examinations. Continuous Internal Evaluation (CIE) provides multiple observations of student performance during a semester, while the Semester End Examination (SEE) provides a major end-point assessment. However, these data are frequently analyzed independently or only after the semester has concluded.

A major challenge for academic departments is identifying students who are progressively falling behind early enough for intervention. A student who performs poorly in the first CIE may improve after mentoring, whereas a student whose marks decline across successive assessments may require immediate support. Consequently, a monitoring system should consider both current achievement and the direction of performance change.

This paper proposes an intelligent framework that integrates CIE and SEE data for continuous progress monitoring and early warning. The principal contributions are: (1) a unified student-performance data model; (2) CIE trend and SEE outcome analytics; (3) a risk classification mechanism for early intervention; (4) an experimental machine-learning framework for predicting SEE performance; and (5) dashboards and reports designed for academic mentoring.

The remainder of the paper discusses related work, the problem definition, proposed architecture, methodology, feature engineering, prediction framework, evaluation strategy, discussion, limitations, future work, and conclusion.

## II. RELATED WORK

Educational Data Mining and Learning Analytics have been used to discover patterns in student interaction and academic performance. Romero and Ventura surveyed educational data mining techniques and highlighted classification, clustering, prediction, and visualization as important applications [1]. Siemens and Baker described learning analytics and educational data mining as complementary approaches for understanding learning processes and supporting educational decision making [2]. Early warning systems attempt to identify students who are at risk of poor academic outcomes before final results are available. Such systems commonly use assessment performance, attendance, learning-management-system activity, and historical academic records as predictors. The effectiveness of these systems depends on the quality, timeliness, and relevance of available features.

The proposed framework focuses specifically on the relationship between CIE and SEE performance. CIE assessments provide repeated observations during the semester, allowing trend-based features such as improvement, decline, average performance, variance, and consistency to be derived. These features can be combined with historical academic indicators to support risk prediction.

## III. PROBLEM STATEMENT AND OBJECTIVES

The research problem is to develop a system that continuously monitors student performance using CIE data and uses the available information to estimate academic risk before SEE results are available. The system should support both descriptive monitoring and predictive early warning while providing interpretable indicators for faculty intervention.

| Objective | Description |
|---|---|
| Continuous monitoring | Track performance after each CIE assessment. |
| Trend analysis | Measure improvement, decline, stability and variation. |
| SEE prediction | Estimate the likelihood of satisfactory or unsatisfactory SEE performance. |
| Early warning | Classify students into actionable risk categories. |
| Intervention support | Provide information that helps mentors prioritize students. |
| Visualization | Present student, subject, class and cohort-level dashboards. |
| Evaluation | Measure prediction accuracy, recall, precision and calibration. |

## IV. PROPOSED SYSTEM ARCHITECTURE

The proposed system consists of five major layers: data acquisition, preprocessing, analytics, prediction, and intervention reporting. The data acquisition layer imports CIE and historical academic records from institutional sources. The preprocessing layer validates marks, handles missing values, standardizes assessment scales, and creates a consistent student-subject representation.

The analytics layer calculates descriptive indicators such as CIE average, maximum, minimum, standard deviation, assessment-to-assessment change, subject-level performance, and cohort percentile. The prediction layer uses selected features to estimate SEE risk. The intervention layer presents risk categories and supporting indicators through dashboards and reports.

| Component | Function |
|---|---|
| Data Acquisition | Collect CIE, SEE, subject and historical academic data. |
| Data Preprocessing | Validate, clean, normalize and transform records. |
| Feature Engineering | Create averages, trends, changes, variability and historical indicators. |
| Analytics Engine | Generate descriptive student and cohort performance measures. |
| Prediction Engine | Classify or predict SEE academic risk. |
| Dashboard | Display progress, trends and risk indicators. |
| Intervention Module | Generate mentor lists and recommended follow-up actions. |

## V. DATA MODEL AND FEATURE ENGINEERING

A student-subject assessment record can be represented using a student identifier, semester, subject, assessment number, assessment score, maximum score, and assessment date. SEE marks can be used as the target variable for retrospective model training. Student identifiers should be pseudonymized during analytical processing.

The following features are useful for modeling:

| Feature | Description |
|---|---|
| CIE Mean | Average normalized CIE score for the subject. |
| CIE Trend | Rate of change between successive CIE assessments. |
| CIE Variance | Variation in performance across CIE assessments. |
| Latest CIE | Most recent CIE score available. |
| Best CIE | Highest CIE score achieved. |
| Consistency | Measure of stability across assessments. |
| Subject Risk Count | Number of subjects in which performance is below a configured threshold. |
| Historical Performance | Previous-semester academic indicators, when available. |
| CIE Coverage | Number of assessments completed relative to expected assessments. |

Marks from different assessment schemes should be normalized before aggregation. For example, if assessments have different maximum marks, each score can be converted to a percentage. Thresholds should be defined by institutional academic policy and should not be presented as universal standards.

## VI. EARLY WARNING METHODOLOGY

The system operates periodically after each CIE event. Newly available assessment records are incorporated into the student's progress profile. A rule-based layer can provide transparent baseline alerts, while a machine-learning layer can provide retrospective prediction when sufficient historical labeled data are available.

A simple baseline risk score can combine normalized CIE performance and trend. For example, the system may assign low, moderate, or high risk according to institutionally approved thresholds. The exact thresholds should be calibrated using historical outcomes and reviewed by academic experts.

For machine learning, the historical dataset is divided into training and test sets without allowing future information to leak into the training features. Candidate models may include Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and gradient-boosting methods. The target can be defined as a binary outcome such as satisfactory/at-risk SEE performance, or as multiple performance categories.

Algorithm 1: Early warning workflow
1. Import the latest CIE records.
2. Validate student, subject and assessment fields.
3. Normalize scores to a common scale.
4. Calculate student-subject performance features.
5. Calculate CIE trends and consistency measures.
6. Apply transparent academic risk rules.
7. If an institutionally validated predictive model is available, generate SEE-risk probabilities.
8. Assign an early-warning category.
9. Update the student and class dashboards.
10. Record interventions and subsequent outcomes for evaluation.

## VII. PREDICTION MODELS AND EVALUATION

The predictive component should be evaluated using historical CIE and SEE data. The model must be trained only on information that would have been available at the time the prediction is intended to be made. This prevents data leakage, particularly the accidental use of SEE marks or post-SEE information as input features.

| Metric | Purpose |
|---|---|
| Accuracy | Overall proportion of correct classifications. |
| Precision | Proportion of predicted at-risk students who are actually at risk. |
| Recall | Proportion of actual at-risk students correctly identified. |
| F1-score | Balance between precision and recall. |
| ROC-AUC | Ability to distinguish risk classes across thresholds. |
| Confusion Matrix | Detailed distribution of correct and incorrect classifications. |
| Calibration | Agreement between predicted risk probabilities and observed outcomes. |

For an early warning system, recall for the at-risk class is particularly important because missing a genuinely struggling student can prevent timely intervention. However, very high recall accompanied by excessive false alarms may overload mentors. Therefore, threshold selection should consider intervention capacity and the cost of false positives and false negatives.
Actual model performance should be reported only after experiments on the institution's dataset. This paper deliberately does not invent accuracy, precision, recall, or F1 values.

## VIII. DASHBOARD AND INTERVENTION DESIGN

The student dashboard can display CIE marks by assessment, subject-wise trends, current average, areas of improvement, and comparison with the student's previous performance. Faculty dashboards can summarize class distributions, subjects with low performance, and students requiring attention.
An early warning report should provide more than a risk label. For each flagged student, it should show the factors contributing to the alert, such as repeated low CIE scores, declining trend, high variability, or multiple weak subjects. This improves interpretability and enables targeted intervention.

| Risk Category | Possible Academic Response |
|---|---|
| Low Risk | Continue normal monitoring and encourage sustained performance. |
| Moderate Risk | Mentor discussion, additional practice and progress review. |
| High Risk | Priority mentoring, remedial support and frequent follow-up. |
| Critical/Exceptional | Immediate academic review according to institutional policy. |

Risk categories and interventions should be configurable and approved by the institution. The system should support recording the date, type and outcome of interventions so that future research can evaluate whether early warnings lead to improved academic outcomes.

## IX. IMPLEMENTATION CONSIDERATIONS

A practical implementation can use Python with pandas for data preparation, scikit-learn for machine-learning models, a relational database for persistence, and a web framework such as Flask or Django for dashboards and APIs. Excel or CSV files may be supported as import sources when institutional systems do not provide APIs.
Data preprocessing should include duplicate detection, missing-value handling, range validation, subject-code validation, and assessment-scale normalization. Access control should restrict individual student information to

authorized faculty and administrators. Analytical datasets should use pseudonymized identifiers whenever possible.
The system should maintain versioned models and configuration values. Whenever a prediction is generated, the model version, feature timestamp, and available CIE assessment number should be recorded. This provides reproducibility and allows administrators to understand how a warning was generated.

## X. EXPERIMENTAL DESIGN

A rigorous experiment should use historical cohorts for which both CIE and SEE outcomes are available. A time-aware split is preferable to a random split when the goal is to simulate deployment on future cohorts. Models should be compared against a transparent rule-based baseline.

| Experiment | Purpose |
|---|---|
| Baseline rules vs ML | Determine whether predictive models improve over transparent thresholds. |
| Assessment-by-assessment | Measure how early reliable prediction becomes possible. |
| Subject-wise models | Test whether prediction differs across subjects. |
| Cohort-wise validation | Evaluate generalization to subsequent student batches. |
| Feature ablation | Identify the contribution of trend, consistency and historical features. |
| Threshold analysis | Study the trade-off between recall and false alarms. |

A suitable study can report results after CIE-1, CIE-2, and subsequent assessments separately. This allows researchers to answer an important question: how early can the system identify students who are likely to experience poor SEE performance while maintaining acceptable false-alarm rates?

## XI. DISCUSSION

Integrating CIE and SEE data provides a natural temporal structure for academic monitoring. Rather than waiting for the final examination, institutions can use repeated internal assessments to identify changes in performance. Trend information may be more informative than a single low mark because it distinguishes temporary difficulty from persistent decline.
The proposed framework also supports a shift from result reporting to intervention-oriented analytics. A warning is valuable only when it leads to an appropriate academic response. Consequently, intervention records should become part of the research dataset so that future studies can evaluate whether alerts and mentoring actually improve student outcomes.
There are important risks. A predictive model may reproduce historical biases, produce false alarms, or perform differently across subjects and cohorts. Academic staff should therefore treat predictions as decision-support information rather than automatic judgments about students. Explainability, human review, privacy, and periodic model validation are essential.

## XII. LIMITATIONS

The framework depends on the availability and quality of CIE records. Missing assessments, inconsistent marking schemes, subject changes, and small historical datasets can reduce predictive reliability. Models trained on one institution may not generalize to another institution because assessment patterns and academic policies differ.
Another limitation is that academic performance is influenced by factors that may not be represented in CIE data. Attendance, engagement, learning resources, and other contextual variables may improve prediction but also raise additional privacy and governance considerations. Only data that are necessary, authorized, and ethically appropriate should be incorporated.

## XIII. FUTURE WORK

Future work can investigate explainable machine learning to provide faculty with interpretable reasons for each warning. Longitudinal models can analyze performance trajectories across semesters. Personalized recommendation systems can suggest remedial resources based on identified weak subjects.
Further research can study intervention effectiveness using causal or quasi-experimental designs, fairness across student groups where legally and ethically appropriate, uncertainty-aware predictions, and model drift across academic years. Integration with institutional ERP, LMS, attendance, and examination systems through secure APIs can enable near-real-time monitoring.

## XIV. CONCLUSION

This paper proposed an intelligent student progress monitoring and early warning system that integrates CIE and SEE performance analytics. The framework combines continuous assessment data, trend analysis, descriptive dashboards, transparent risk rules, and machine-learning prediction to identify students who may require timely academic support.

The proposed system is designed as a decision-support mechanism rather than a replacement for faculty judgment. Its effectiveness should be established through institution-specific experiments measuring prediction quality, timeliness, false alarms, intervention outcomes, and student improvement. With appropriate governance and validation, CIE-driven early warning can provide a practical foundation for proactive academic mentoring and data-driven quality improvement.

**References.**

[1] C. Romero and S. Ventura, "Educational data mining: A review of the state of the art," IEEE Trans. Syst., Man, Cybern. C, Appl. Rev., vol. 40, no. 6, pp. 601–618, Nov. 2010.
[2] G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," in Proc. 2nd Int. Conf. Learning Analytics and Knowledge, Vancouver, BC, Canada, 2012, pp. 252–254.
[3] A. Peña-Ayala, "Educational data mining: A survey and a data mining-based analysis of recent works," Expert Syst. Appl., vol. 41, no. 4, pp. 1432–1462, 2014.
[4] R. S. J. d. Baker and K. Yacef, "The state of educational data mining in 2009: A review and future visions," J. Educ. Data Mining, vol. 1, no. 1, pp. 3–17, 2009.
[5] M. A. B. Pardo and G. Siemens, "Ethical and privacy principles for learning analytics," Brit. J. Educ. Technol., vol. 45, no. 3, pp. 438–450, 2014.
[6] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.
[7] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed. New York, NY, USA: Springer, 2009.
[8] S. D. O'Connor, "Early alert systems and student success: A review of predictive academic analytics," in Proc. Int. Conf. Learning Analytics, 2016.