International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(7s) 2026 pp. 949-956
https://doi.org/10.70917/ijcisim-2026-3166
Article

An Educational Data Intelligence Model for Early Identification of Slow Learners in BCA Programs Using Multi-Factor Analysis and Explainable Machine Learning

Santosh P. Nalawade[1], Rajendra S. Pujari[2]

[1]Bharati Vidyapeeth (Deemed to be University), Institute of Management and Rural Development Administration, Sangli, Maharashtra, India
Email:ID: nalawadesantosh011@gmail.com;
[2]Department of Computer Applications, Bharati Vidyapeeth (Deemed to be University), Institute of Management and Rural Development Administration, Sangli, Maharashtra, India
Email:ID:rajendraspujari@gmail.com

Abstract: — In recent years, higher education institutions have witnessed an unprecedented growth in educational data due to increased student enrollment, digital learning platforms, Learning Management Systems (LMS), and continuous assessment mechanisms. Technology-oriented programs such as the Bachelor of Computer Applications (BCA) generate multidimensional data related to students' academic, behavioral, psychological, and technical performance. However, traditional evaluation systems rely on delayed assessments, leading to late identification of learning difficulties and ineffective interventions. This paper proposes an Educational Data Intelligence (EDI) model for the early identification of slow learners in BCA programs using multi-factor analysis and explainable machine learning. Primary data collected from 2,028 BCA students through a structured questionnaire were analyzed. After correlation analysis, chi-square validation, and expert review, 20 significant educational intelligence factors were identified. An interpretable decision-tree-based REPTree algorithm was employed to classify students into risk categories. The proposed EDI model provides transparent, actionable insights for early intervention, academic planning, and student retention. The study demonstrates that EDI functions not only as a predictive model but also as an effective academic decision-support system.
Keywords: — Educational Data Intelligence, Slow Learners, BCA, Explainable Machine Learning, Learning Analytics, REPTree

1. INTRODUCTION

The rapid expansion of student enrollment, digital learning environments, and Learning Management Systems (LMS) has resulted in an exponential increase in educational data within higher education institutions. [1] Programs such as the Bachelor of Computer Applications (BCA), which are highly technology-intensive, generate large volumes of academic, behavioral, psychological, and technical data. Despite this abundance, institutional decision-making continues to rely primarily on delayed indicators such as end-semester examinations and cumulative results. [2]

Traditional assessment practices are reactive in nature and often identify learning difficulties only after students experience significant academic decline. Consequently, timely educational interventions become difficult, leading to loss of confidence, poor performance, and increased dropout rates. [3] To address these limitations, Educational Data Intelligence (EDI) has emerged as a data-driven paradigm that converts raw educational data into actionable insights using statistical analysis, data mining, and machine learning techniques. [4]

This study proposes an EDI-based Model to enable early identification of slow learners in BCA programs. By shifting from delayed assessment to proactive intelligence, the proposed model supports early intervention, informed decision-making, and improved student retention. [5]

2. RELATED WORK

A. Educational Data Mining and Learning Analytics

Educational Data Mining (EDM) and Learning Analytics focus on extracting meaningful patterns from student data such as attendance, academic performance, assignment submission, LMS usage, and engagement levels. [6] Prior studies have demonstrated the potential of data-driven insights for educational decision-making; however, most analyses remain descriptive and retrospective, limiting their utility for early intervention. [7]

B. Machine Learning-Based Slow Learner and Dropout Prediction

Recent research has employed machine learning algorithms such as Naïve Bayes, Logistic Regression, Decision Trees, Support Vector Machines, and Random Forests to predict academic risk and dropout probability. [8] Although these models often report high predictive accuracy, many function as black-box systems, providing limited interpretability for educators and academic administrators.

C. Research Gap

The literature indicates a lack of interpretable, factor-driven Educational Data Intelligence systems. Most existing models classify students as “slow” or “not slow” without explaining the contributing academic, behavioral, or psychological factors. This gap underscores the need for transparent EDI Models that support early warning and targeted intervention.

3. DATA INTELLIGENCE MODEL

A. Data Acquisition

Primary data were collected using a structured 30-item questionnaire administered to 2,028 students enrolled in the BCA program. The questionnaire captured multidimensional information encompassing academic background, behavioral patterns, psychological aspects, socio-economic conditions, and technical competencies. This approach enabled a holistic understanding of student learning behavior.

B. Feature Intelligence and Reduction

A multi-stage feature selection process was applied to identify the most influential educational intelligence factors:

A. Correlation Analysis to examine relationships between factors and academic performance;

B. Chi-square validation to identify statistically significant categorical relationships; and

C. Expert validation by academic mentors and subject experts.

Based on these validations, 20 educational intelligence factors were finalized for model development.

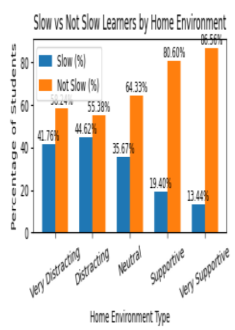
Table 1: “Finalized 20 Educational Intelligence Factors”

1. Home Environment	2. 12th Percentage	3. Daily Study Hours	4. Exam Preparation Strategy	5. Practice of Programming Outside Class
6. Use of LMS Tools (Moodle, Google Classroom)	7. Availability of Teaching Infrastructure	8. Part-Time Job	9. First Semester Percentage	10. Mathematics Proficiency

11. Basic Computer Knowledge	12. Basic Programming Concepts	13. Class Attendance	14. Timely Assignment Submission	15. Students' Attention Level
16. Exam Stress Frequency	17. Confidence During Oral Presentations	18. Internal Class Test Performance	19. Improvement in Programming Concepts	20. Use of Technology in Teaching

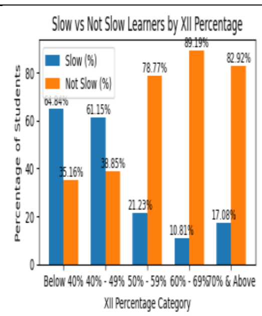
4. FACTOR-WISE EDUCATIONAL INTELLIGENCE ANALYSIS

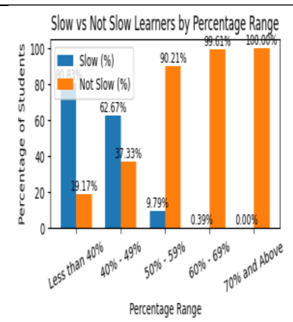
A. Environmental Intelligence (HOME_ENV)



Analysis reveals that a supportive home environment significantly enhances learning speed, whereas disruptive environments are strongly associated with slow learning. This insight supports interventions such as extended lab access and mentoring programs.

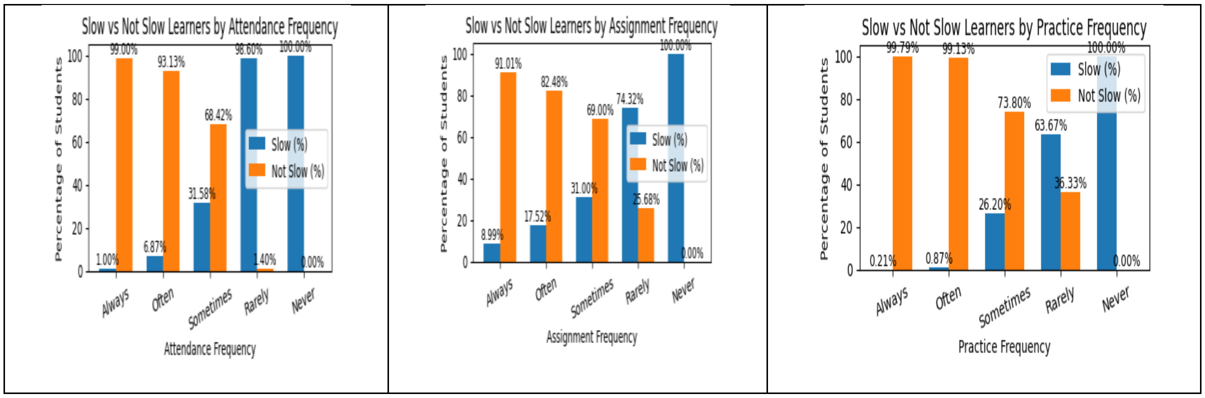
B. Academic Intelligence (XII_PER, SEM1_PER)





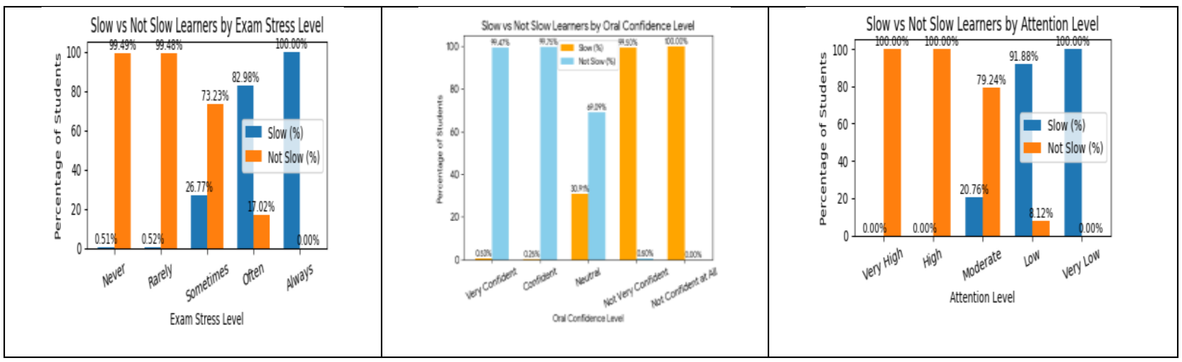
Students with strong academic foundations in higher secondary education and the first semester exhibit faster learning progression. Conversely, weak foundational performance indicates higher academic risk, highlighting the importance of bridge courses and remedial instruction.

C. Behavioral Intelligence (Attendance, Assignments, Practice)



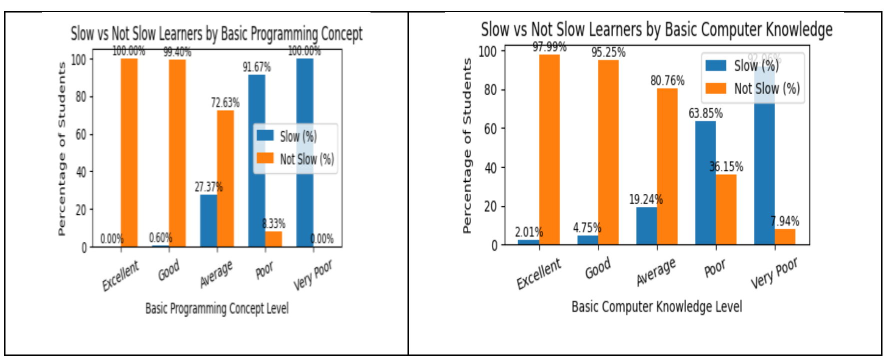
Regular attendance, timely assignment submission, and consistent programming practice are strong indicators of fast learners. Irregular engagement patterns are key predictors of slow learning.

D. Psychological Intelligence (Stress, Confidence, Attention)



High stress levels, low confidence, and poor attention are closely associated with slow learners, emphasizing the need for counseling and stress-management interventions.

E. Technical Intelligence (Programming and Computer Knowledge)



Strong programming fundamentals and basic computer knowledge are essential for success in BCA programs. Weak technical skills consistently correlate with slow learning patterns.

This factor-wise intelligence analysis makes it clear that BCA student learning is not dependent on a single factor, but is based on the combined effects of academic, behavioral, psychological, environmental, and technological factors. [9] This interpretable insight provides a strong basis for early identification, targeted intervention, and effective educational planning through Educational Data Intelligence.

5. Educational Data Intelligence–Based Model Development

V. I Tool Ecosystem:

This research used reliable, academically acceptable and research-oriented tools to implement the Educational Data Intelligence (EDI) concept. [10]

• **Python:** Python was used for data preprocessing, analysis and initial experiments. This made it possible to handle large amounts of educational data and find meaningful patterns. [11]

• **Jupyter Notebook:** Jupyter Notebook was used for step-by-step documentation of the entire data intelligence process, visual observation and reproducible experiments. [12]

• **WEKA:** WEKA was used as the main tool to implement, compare and obtain clear rule-based results of machine learning models. [13]

This tool ecosystem is research-oriented and can be easily adopted in educational institutions. [14]

V. II Explainable ML Selection – REPTree

According to the EDI approach, interpretability and educational utility were given priority over accuracy alone while selecting the model. For these reasons, the REPTree (Reduced Error Pruning Tree) algorithm was chosen.[15]

• **Transparent Rules:** The decision-tree structure generated by REPTree is clear and easy to understand, making it easy to understand the reasoning behind each decision.

• **Fast Inference:** Due to its low computational complexity, REPTree produces fast results even on large academic data, which is useful for actual academic use.

• **Academic Usability:** The decision-tree-based rules are easy for teachers, mentors, and academic administrators to understand and can be used directly for intervention planning.

This model is not just a "predictive" tool, but also acts as an academic decision-support system. Instead of providing ambiguous results like a black-box model, this EDI-based model explains the academic reasoning behind each decision and enables timely, targeted interventions. [16]

6. Workflow of the proposed Educational Data Intelligence (EDI) system

A coherent and step-by-step system workflow is required to effectively implement the Educational Data Intelligence (EDI) approach. The EDI system in the present research represents the entire process from data collection to educational intervention in a unified manner.

Stage I: Questionnaire Design

The system begins with a structured questionnaire of 30 items. The most effective 20 educational intelligence items are selected from these through the data intelligence process, which form the foundation of the EDI system.

Stage II: Feature Intelligence Validation

The selected items are validated based on correlation analysis, chi-square test, and expert opinions, which ensures that the data is educationally meaningful.

Stage III: GUI-based data entry

A GUI-based system has been developed using Python and Django, which makes it easy for teachers and administrators to record standardized and error-free data.

Stage IV: REPTree-based Classification

The recorded data is classified using the REPTree algorithm. This decision-tree based process explains the reasons behind the classification, making the system interpretable.

Step V: Student Risk Classification

Based on the results of REPTree, students are classified into groups such as Low Risk, Moderate Risk and High Risk (Slow Learner).

Step VI: Early Educational Intervention

According to the risk level, early interventions such as remedial classes, guidance, counseling and technical training can be implemented, which represents the ultimate goal of the EDI system.

Summary

The proposed EDI workflow is organized into clear steps such as

Fig. a. EDI-Based Early Intervention

Therefore, the system is useful for EDI-centric journals as a data-based, transparent and practical educational decision-support solution.

7. Results and Educational Insights:

Instead of restating model accuracy or technical metrics, this section explains the meaningful educational insights derived from Educational Data Intelligence (EDI). These insights are directly useful for institutional decision-making and early intervention. [17]

7. 1 Dominant Educational Intelligence Factors:

The analysis found that certain factors consistently influence student learning speed. In particular, class attendance, understanding of basic programming concepts, academic performance in the semester, test stress, and regular practice appear prominently in most decision-tree paths. [18] These factors emerge as key decision-makers in the EDI system and provide clear indications of the risk level of students.

7.2 Emerging Learner Profiles:

Different types of learner profiles emerge clearly from EDI-based analysis: [19]

• **Academically Weak but Motivated Students:** Students with high attendance and regular practice despite a low academic foundation, who can progress rapidly with appropriate guidance.

• **Technically Weak Students:** Students who are deficient in computer and programming knowledge, but have relatively well other behavioral factors.

• **Stressed Students:** Students who are in the slow learner group despite having good academic ability due to high test stress, low self-confidence and lack of focus.

• **Students in the persistent risk group (Slow Learner Profile):** Students with low attendance, weak programming fundamentals and high test stress, who are consistently classified as slow learners.

7.3 Early Institutional Action:

This EDI system allows institutions to act early instead of reacting late. The following solutions based on the rules derived from the decision-trees are possible: [20]

• Immediate academic consultation and monitoring if low attendance and low practice are detected.

• Bridge courses and additional lab sessions if programming fundamentals are weak.

• Psychological counseling and stress management programs for students with high stress and low self-confidence.

Exemplary Insights: Students who show low attendance, weak programming fundamentals and high test stress are consistently found in slow learners. This insight suggests that not only academic, but also combined educational and psychological interventions are essential for such students. [21]

These results demonstrate that Educational Data Intelligence is not just a classification system, but also provides a deeper understanding of students' learning patterns. Based on such insights, higher education institutions can make more effective, targeted and timely educational decisions.

8. Discussion (Educational Data Intelligence – EDI Approach)

The findings of this research demonstrate that the learning process of students is multidimensional and is influenced by a combination of academic, behavioral, psychological, technological and environmental factors. While traditional assessment methods ignore this complexity, Educational Data Intelligence (EDI) effectively identifies the interrelationships between these dimensions.

A major strength of EDI is its explainability. Explainable models like REPTree provide a clear understanding of why students are classified into at-risk groups, allowing teachers and administrators to plan appropriate interventions with confidence in the findings.

Compared to traditional black-box machine learning models, EDI is not just a predictive but a decision-support system. By facilitating the data → insight → action journey, EDI helps in achieving early identification, effective intervention and sustainable educational improvement. Therefore, EDI approach is crucial for a data-centric yet human-centric education system.

9. Conclusion and Future Scope:

This study demonstrates that the concept of Educational Data Intelligence (EDI) transforms raw educational data into meaningful and actionable insights. By analyzing academic, behavioral, psychological and environmental factors together, it is possible to identify early risks among BCA students. This enables higher education institutions to proactively provide timely academic and psychological support to slow learners.

This research also demonstrates that EDI-based approach is not limited to classification alone but acts as an effective decision-support system. Clear and understandable rules enable teachers and administrators to plan appropriate interventions, which improves academic performance and helps reduce dropout.

In the future, EDI systems can be further enhanced using real-time learning analytics, data integration across online platforms, and explainable deep learning models. Educational Data Intelligence will thus become a data-driven and student-centric solution for higher education.

References:

1. M. Yağcı, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," Smart Learning Environments, vol. 9, 2022.
2. L. P. Macfadyen and S. Dawson, "Mining LMS data to develop an early warning system for educators," Computers & Education, vol. 54, no. 2, pp. 588–599, 2010.
3. M. L. Bernacki, M. M. Chavez, and P. M. Uesbeck, "Predicting achievement and providing support before students begin to fail," Computers & Education, 2020.
4. B. Albreiki, N. Zaki, and H. Alashwal, "A systematic literature review of student performance prediction using machine learning techniques," Education Sciences, vol. 11, no. 9, 2021.
5. "Explainable artificial intelligence in educational data mining for student performance prediction," Telematics and Informatics Reports, 2025.
6. R. S. Baker and K. Yacef, "The state of educational data mining in 2009: A review and future visions," Journal of Educational Data Mining, vol. 1, no. 1, pp. 3–17, 2009.
7. L. P. Macfadyen and S. Dawson, "Mining LMS data to develop an early warning system for educators," Computers & Education, vol. 54, no. 2, pp. 588–599, 2010.
8. B. Albreiki, N. Zaki, and H. Alashwal, "A systematic literature review of student performance prediction using machine learning techniques," Education Sciences, vol. 11, no. 9, 2021.
9. R. S. Baker, "Data mining for education," International Encyclopedia of Education, 3rd ed., Elsevier, pp. 112–118, 2010.
10. R. S. Baker and K. Yacef, "Educational data mining and learning analytics," Journal of Educational Data Mining, vol. 1, no. 1, pp. 3–17, 2009.
11. J. VanderPlas, Python Data Science Handbook, O'Reilly Media, 2016.
12. T. Kluyver et al., "Jupyter Notebooks – a publishing format for reproducible computational workflows," Positioning and Power in Academic Publishing, IOS Press, 2016.
13. I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, 4th ed., Morgan Kaufmann, 2017.
14. G. Siemens and R. S. J. d. Baker, "Learning analytics and educational data mining: Towards communication and collaboration," Proceedings of the 2nd International Conference on Learning Analytics, 2012.
15. J. R. Quinlan, "Simplifying decision trees," International Journal of Man-Machine Studies, vol. 27, no. 3, pp. 221–234, 1987.
16. F. Dosilovic, M. Brcic, and N. Hlupic, "Explainable artificial intelligence: A survey," 2018 41st International Convention on Information and Communication Technology, IEEE, 2018.

17. G. Siemens and R. S. J. d. Baker, “Learning analytics and educational data mining: Towards communication and collaboration,” Proceedings of the 2nd International Conference on Learning Analytics, 2012.
18. L. P. Macfadyen and S. Dawson, “Mining LMS data to develop an early warning system for educators,” Computers & Education, vol. 54, no. 2, pp. 588–599, 2010.
19. M. L. Bernacki, M. M. Chavez, and P. M. Uesbeck, “Predicting achievement and providing support before students begin to fail,” Computers & Education, 2020.
20. B. Albreiki, N. Zaki, and H. Alashwal, “A systematic literature review of student performance prediction using machine learning techniques,” Education Sciences, vol. 11, no. 9, 2021.
21. R. S. Baker, “Data mining for education,” International Encyclopedia of Education, 3rd ed., Elsevier, 2010.