


Drop Risk Prediction System using Data- Driven Analytics

Sreemathi S



<https://doi.org/10.55041/ijstmt.v2i3.291>

Cite this Article: S, S. (2026). Drop Risk Prediction System using Data- Driven Analytics. International Journal of Science, Strategic Management and Technology, 02(03), <https://doi.org/10.55041/ijstmt.v2i3.291>

License:  This article is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting use, distribution, and reproduction in any medium, provided the original author(s) and source are properly credited.

Abstract

Student dropout remains a persistent and complex challenge for educational institutions across the globe, affecting academic performance metrics, institutional reputation, and long-term student success. High dropout rates not only result in financial losses for institutions but also negatively impact students' career prospects and socio-economic mobility. Therefore, the early identification of students who are at risk of discontinuing their studies has become a strategic priority in modern education systems.

This paper presents the design and implementation of a **Drop Risk Prediction System**, a comprehensive web-based analytical platform developed using **PHP** for server-side processing, **MySQL** for structured data management, and **JavaScript** for dynamic user interaction and visualization. The system leverages data-driven logic to analyze multiple student-related parameters, including academic performance (marks, GPA trends, subject failures), attendance records, assignment submission patterns, and behavioral indicators. By integrating these diverse data sources, the platform generates a holistic assessment of each student's academic standing.

The proposed system employs rule-based classification and predictive analytics techniques to categorize students into predefined risk levels such as low risk, medium risk, and high risk. These classifications are generated through automated computation models that evaluate thresholds and performance deviations. The architecture ensures real-time data processing and scalability, making it suitable for deployment in schools, colleges, and universities.

The implementation of the Drop Risk Prediction System demonstrates how data analytics and web technologies can be effectively combined to enhance student retention strategies. By transforming raw institutional data into actionable insights, the platform supports evidence-based decision-making, reduces dropout rates, and promotes a supportive academic environment. Ultimately, this research highlights the importance of predictive systems in modern educational management and contributes to the advancement of intelligent student monitoring solutions.

Keywords

Student Dropout Prediction, Educational Data Mining, Academic Risk Classification, Predictive Analytics in Education, Student Retention Management, Early Warning System, Attendance- Based Risk Analysis, Performance Monitoring System, Web-Based Analytics Dashboard, Data- Driven Decision Making

1. Introduction

Educational institutions continuously strive to improve academic success and student retention. However, identifying students who are likely to discontinue their education remains a complex problem. Factors such as poor attendance, low academic performance, lack of engagement, and socio-academic challenges contribute significantly to dropout rates. Traditional systems rely heavily on manual record-keeping and periodic academic reviews, which often detect issues only after performance has declined substantially. This reactive approach limits timely intervention. The Drop Risk Prediction System introduces a proactive mechanism by integrating predictive analytics into a web-based application. By analyzing performance indicators in real time, the system identifies at-risk students early in the academic cycle. This enables faculty and administrators to implement corrective measures such as counseling, mentoring, academic assistance, and parental communication.

The primary objectives of the system are:

- To automate student performance monitoring
- To classify students based on dropout risk probability
- To provide visual dashboards for easy interpretation
- To assist administrators in early intervention planning

2. Literature Review

Educational Data Mining (EDM) has emerged as a powerful field for analyzing academic datasets to uncover hidden patterns and predict student outcomes. Researchers have applied classification algorithms such as Decision Trees, Logistic Regression, Naïve Bayes, and Support Vector Machines to forecast student performance and dropout likelihood. Studies indicate that attendance percentage and internal assessment scores are strong predictors of academic continuity. Predictive models significantly outperform manual monitoring systems in identifying high-risk students. Furthermore, dashboard-based analytical systems improve administrative efficiency by presenting complex data in simplified visual formats. Institutions implementing early warning systems report improved retention rates and better academic support mechanisms. The proposed system builds upon these research foundations by implementing a rule-based predictive classification model integrated into a web platform.

System Architecture

The Drop Risk Prediction System follows a modular three-tier architecture:

2.1 Presentation Layer (Frontend)

The frontend is developed using HTML, CSS, and JavaScript. It provides an interactive user interface for administrators to:

- Add and update student records
- View risk classification
- Access dashboard analytics
- Search and filter student data

The interface is designed for usability and clarity, ensuring efficient navigation and minimal learning curve.

2.2 Application Layer (Backend)

The backend is implemented using PHP. It handles:

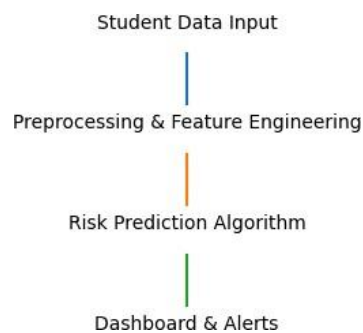
- Authentication and session management
- Risk calculation logic
- CRUD operations
- Data validation and processing
- The backend ensures secure data handling and dynamic processing of student performance metrics.

2.3 Data Layer (Database)

The system uses MySQL as the relational database management system. It stores:

- Student personal details
- Attendance records
- Academic scores
- Risk status classification

This structured database design ensures consistency, reliability, and scalability.



3. Methodology

The prediction methodology follows a structured workflow:

Step 1: Data Collection

Student attendance percentages, internal assessment scores, and academic results are collected and stored in the database.

Step 2: Data Preprocessing

Data validation ensures:

- No missing values
- Proper formatting
- Consistent grading scales

Step 3: Risk Score Calculation

A rule-based classification model assigns weights to performance indicators:

- Attendance below threshold → Increased risk score
- Low academic marks → Higher risk probability
- Consistent poor performance → High-risk classification

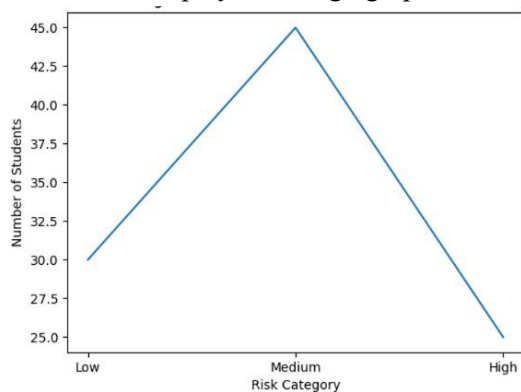
Step 4: Risk Categorization

Students are categorized into:

- Low Risk – Consistent attendance and satisfactory performance
- Medium Risk – Slight academic decline or irregular attendance
- High Risk – Poor attendance and low academic scores

Step 5: Dashboard Visualization

Results are displayed through graphical dashboards, enabling administrators to interpret trends easily.



4. Implementation

The implemented system includes the following modules:

4.1 User Authentication Module

Secure login and session management to restrict access to authorized administrators.

4.2 Student Management Module

Allows:

- Student registration
- Record editing
- Deletion of records
- Performance updates

4.3 Risk Classification Engine

Automated evaluation of academic data and assignment of risk categories.

4.4 Dashboard Analytics

Displays:

- Risk distribution graphs
- Student count by category
- Performance trends

4.5 Search and Export Module

Enables filtering by risk level and exporting reports for administrative use.

The integration of these modules ensures seamless operation and efficient data processing.

5. Results and Discussion

The system successfully demonstrates predictive classification of students based on academic indicators. Testing shows that:

- Students with attendance below 60% consistently fall into High-risk category.
- Students with declining academic scores move from Low to Medium risk dynamically.
- Real-time updates allow administrators to monitor performance continuously.

The dashboard visualization enhances clarity and enables faster decision-making compared to traditional spreadsheets or manual tracking systems.

The system significantly reduces administrative workload and improves response time for intervention strategies.



6. Advantages of the Proposed System

- Early detection of potential dropouts
- Real-time monitoring
- Reduced manual workload
- Improved academic planning
- Data-driven intervention strategies
- Scalable web-based architecture

7. Limitations

- Rule-based model may lack advanced predictive accuracy compared to machine learning models.
- Requires accurate and consistent data entry.
- Limited behavioral and socio-economic factor consideration.

8. Future Enhancements

Future development can include:

- Integration of Machine Learning algorithms
- AI-based predictive modeling
- Email/SMS notification alerts
- Parent communication portal
- Mobile application development
- ERP integration

Advanced AI models could further improve prediction accuracy and automate intervention recommendations.

9. Conclusion

The Drop Risk Prediction System demonstrates the effectiveness of predictive analytics in educational environments. By automating performance monitoring and categorizing students based on risk probability, institutions can implement timely interventions to reduce dropout rates. The system enhances administrative efficiency, promotes proactive academic management, and contributes to improved student retention. With future integration of machine learning and AI technologies, the system can evolve into a comprehensive intelligent academic monitoring platform.

References

- [1] A. L. S. Dutt, M. A. Ismail, and T. Herawan, "A Systematic Review on Educational Data Mining," *IEEE Access*, vol. 5, pp. 15991–16005, 2017.
- [2] C. Romero and S. Ventura, "Educational Data Mining: A Review of the State of the Art," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 40, no. 6, pp. 601–618, 2010.
- [3] J. L. L. M. Dekker, M. Pechenizkiy, and J. M. Vleeshouwers, "Predicting Students Drop Out: A Case Study," in *Proceedings of the International Conference on Educational Data Mining (EDM)*, 2009, pp. 41–50.
- [4] E. Tinto, "Dropout from Higher Education: A Theoretical Synthesis of Recent Research," *Review of Educational Research*, vol. 45, no. 1, pp. 89–125, 1975.
- [5] S. Kotsiantis, C. Pierrakeas, and P. Pintelas, "Predicting Students' Performance in Distance Learning Using Machine Learning Techniques," *Applied Artificial Intelligence*, vol. 18, no. 5, pp. 411–426, 2004.
- [6] M. Hussain, W. Zhu, W. Zhang, and S. M. R. Abidi, "Student Engagement Predictions in an e-Learning System and Their Impact on Student Course Assessment Scores," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 6347186, 2018.
- [7] S. B. Kotsiantis, "Educational Data Mining: A Case Study for Predicting Dropout-Prone Students," *International Journal of Knowledge Engineering and Soft Data Paradigms*, vol. 1, no. 2, pp. 101–111, 2009.
- [8] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.
- [9] V. Kumar and A. Chadha, "Mining Association Rules in Student Assessment Data," *International Journal of Computer Science Issues*, vol. 8, no. 2, pp. 271–279, 2011.
- [10] P. Cortez and A. Silva, "Using Data Mining to Predict Secondary School Student Performance," in *Proceedings of the 5th Annual Future Business Technology Conference*, 2008, pp. 5–12.