

AI: Dropout Prediction System

Anuja Nandurkar¹, Harshwardhan Talware², Amruta Ranotkar³, Shantanu Hate⁴, Sakshi Rathod⁵, Dr. H. N. Datir⁶

^{1,2,3,4,5} *Sipna Collage of Engineering and Technology*

Abstract- Student dropout is a major challenge in educational institutions due to academic, financial, and behavioural factors. Traditional systems fail to detect at-risk students early. This research proposes AI, a machine learning-based system that predicts student dropout risk using historical and real-time data. The system uses algorithms like Random Forest and XGBoost to analyze student performance and generate risk scores. It also automates counsellor assignment and provides dashboards for monitoring. This approach enables proactive intervention and improves student retention.

I. INTRODUCTION

Student retention is an important issue in modern education. Many students leave their studies due to poor academic performance, financial problems, or personal issues. Traditional systems are reactive and take action only after problems become serious.

With the help of Machine Learning, institutions can predict dropout risk early. AI uses data-driven techniques to identify at-risk students and provide timely support.

II. PROBLEM STATEMENT

Current systems face several problems:

- Data is stored in separate departments (no integration)
- No early prediction of dropout
- Manual and slow intervention
- Difficulty in analyzing multiple factors
- Poor communication with students

This creates a need for an intelligent automated system.

III. OBJECTIVES

- To predict student dropout using Machine Learning
- To identify at-risk students early

- To automate counsellor assignment
- To improve student retention
- To provide real-time monitoring dashboards

IV. LITERATURE REVIEW

Earlier methods used statistical models and manual tracking, which were not accurate.

Modern approaches use:

- Educational Data Mining (EDM)
- Machine Learning algorithms
- Ensemble models like Random Forest and XGBoost

These methods provide better accuracy and can handle complex data.

V. PROPOSED SYSTEM

AI consists of the following modules:

5.1 Data Collection Module

- Academic data (marks, attendance)
- Behavioral data
- Financial data

5.2 Prediction Module

- Uses ML models to calculate dropout risk

5.3 Counsellor Assignment Module

- Automatically assigns counsellor based on issue

5.4 Dashboard Module

- Admin dashboard
- Teacher dashboard
- Student dashboard

VI. METHODOLOGY

Step 1: Data Collection

- Collect student academic, financial, and behavioural data

Step 2: Data Preprocessing

- Clean and normalize data
- Handle missing values

Step 3: Model Training

- Train models using dataset
- Use Random Forest and Boost

Step 4: Prediction

- Generate risk score (Low / Medium / High)

Step 5: Intervention

- Assign counsellor automatically
- Provide support to students

VII. ALGORITHMS USED

1. Random Forest

- Handles large datasets
- Provides high accuracy

2. XGBoost

- Fast and efficient
- Handles missing data

These algorithms help in accurate prediction of dropout risk.

VIII. RESULTS AND DISCUSSION

- High prediction accuracy
- Early identification of at-risk students
- Improved intervention process
- Reduced dropout rates

The system provides real-time monitoring and decision-making support.

IX. ADVANTAGES

- Early prediction of dropout
- Automated system
- Saves time and effort
- Improves student success
- Scalable for large institutions

X. DISADVANTAGES

- Requires quality data
- Initial setup complexity
- Needs technical knowledge

XI. APPLICATIONS

- Colleges and universities
- Online learning platforms
- Educational analytics systems

XII. FUTURE SCOPE

- Integration with LMS/ERP systems
- Mobile application
- AI-based personalized learning
- NLP for sentiment analysis

XIII. CONCLUSION

AI provides an intelligent solution for predicting student dropout and improving retention. By using Machine Learning and automation, the system helps institutions take proactive actions and support students effectively. It transforms traditional systems into data-driven and efficient platforms.

REFERENCES

- [1] Putra L G R, Prasetya D D, Mayadi, Student Dropout Prediction Using Random Forest and XGBoost Method International Journal of Information Systems, 2024
- [2] Nti I K, Ramanayake S, Explainable Machine Learning for Student Dropout Prediction and Intervention, Discover Artificial Intelligence, 2026
- [3] Dasi H, Kanakala S, Student Dropout Prediction Using Machine Learning Techniques, International Journal of Intelligent Systems, 2022
- [4] Villar A, Andrade C, Supervised Machine Learning Algorithms for Predicting Student Dropout, Springer, 2024
- [5] Alhardi A, Alan S, Predicting Student Dropout Using XGBoost Algorithm, International Conference on Engineering Technologies, 2024
- [6] Maulana M, Hariyadi M, Predicting Student Dropout Risk Using Random Forest, Journal of Information Systems, 2025