Department of Information Technology
NBA Accredited

A.P. Shah Institute of Technology
G.B.Road, Kasarvadavli, Thane (W), Mumbai-400615
UNIVERSITY OF MUMBAI
Academic Year 2024-2025

A Project Presentation on

Dropout Defender: A Machine Learning Approach to Lower Dropout Rates

Submitted in partial fulfillment of the degree of
Bachelor of Engineering(Sem-8) in

INFORMATION TECHNOLOGY

By

Harmi Mathukiya (21104044)
Avantika More (21104033)
Sahil Mohite (21104099)
Atharva Mohape (21104121)

Under the Guidance of
Ms.Apeksha Mohite

1.Project Conception and Initiation

1.1 Abstract

- Dropout Defender: A Machine Learning Approach to Lower Dropout Rates is a comprehensive web application designed to predict and reduce student dropout rates by leveraging machine learning techniques. The system offers tailored dashboards for students, and mentors enabling each user role to interact with the platform according to their specific needs.
- The application is designed for deployment on scalable platforms, ensuring efficient and seamless user experience. The project utilizes Firebase for backend data storage, ensuring flexibility and scalability, while the front-end interface is built using modern web technologies.
- The system is designed to be user-friendly, providing role-based access to key features such as login, register, adding student details, predicting dropout status and viewing dashboards.

1.2 Objectives

- To create a model that can predict whether a student is likely to drop out of an academic program.
- To compare different ML classifiers (algorithms) to find the most accurate one for predicting dropouts.
- To create different dashboards such as students, and teacher to view and upload the progress of the student.
- To implement login and registration for students, and mentors to maintain secure access.
- To ensure easy navigation between dashboards, grade reports, and communication tools.

1.3 Literature Review

Title	Authors	Findings
Predicting Student Dropout in Higher Education Using Machine Learning Tech-niques	O. B. D. Ayoob, R. Shah, A. Ali (2022)	This paper presents a comparative analysis of various machine learning algorithms, including decision trees, random forest, and support vector machines, to predict student dropouts. The study highlights the effectiveness of ensemble methods in improving prediction accuracy, with a focus on real-time data analysis and intervention strategies in higher education systems.
Analyzing Student Dropout Factors Using Predictive Analytics	C. T. Munoz, J. A. Asensio, M. A. Gonz´alez (2023)	This research introduces a predictive model based on longitudinal data analysis, combining academic, social, and behavioral factors. The authors employ recurrent neural networks to detect early signs of disengagement, proposing a real-time alert system for educators to intervene before students drop out.

1.3 Literature Review

Title	Author	Findings
Machine Learning Approaches for Student Dropout Prediction: A Systematic Review	P. M. Sharma, A. B. Rani, N. K. Gupta (2023)	This paper provides a comprehensive analysis of the different machine learn- ing (ML) techniques applied to predict student dropout rates, emphasizing the role of ML models in educational settings.The study categorizes various ML methods, and examines their effectiveness, comparing approaches like decision trees, logistic regression, neural networks, support vector machines, and ensemble methods.
Early Detection of At-Risk Students: A Machine Learning Perspective	F. A. Taleb, L. P. Trivedi, J. K. Singh (2022)	This research finds that machine learning models, particularly decision trees and random forests, can effectively predict students at risk of dropping out based on academic and behavioral data. It highlights that early identification allows for timely interventions, significantly improving student retention rates. The study also emphasizes the importance of using a combination of academic records, attendance, and engagement metrics to achieve higher prediction accuracy.

1.4 Problem Definition

- This project aims to solve is the high number of students who drop out of college. The project wants to find these students early on, before they decide to leave, and offer them help to stay in college.
- Therefore, the study focuses on developing a model that uses machine learning algorithms to forecast dropout rates based on historical student data.

1.5 Scope

- Can predict student dropouts using machine learning models by analyzing academic performance, assignments and grades.
- Can compare multiple ML classifiers to determine the most accurate algorithm for predicting dropouts.
- Can provide interactive dashboards for students and mentors to track progress,predict dropout status, and receive insights.
- Can implement secure login and registration to ensure role-based access for students and mentors.
- Can facilitate easy navigation between dashboards, grade reports, and communication tools for an improved user experience.

1.6 Technology stack

1. HTML
2. CSS
3. JAVASCRIPT
4. Firebase Database
5. Machine Learning Algorithms
6. Visual Studio Code

1.7 Benefits for environment & Society

- To reduce students dropout rates that are been increased.
- Personal attentions to particular student by the teacher as well as the parent.
- Many extra classes and improving progress for the students future sake.

2. Project Design

2.1 Proposed System

- The proposed machine learning model analyzes students dropout or not dropout system using various ML algorithms.
- It also includes a website where it gives proper guidance to the dropout student with the help of various study materials.

2.2 Design(Flow Of Modules)

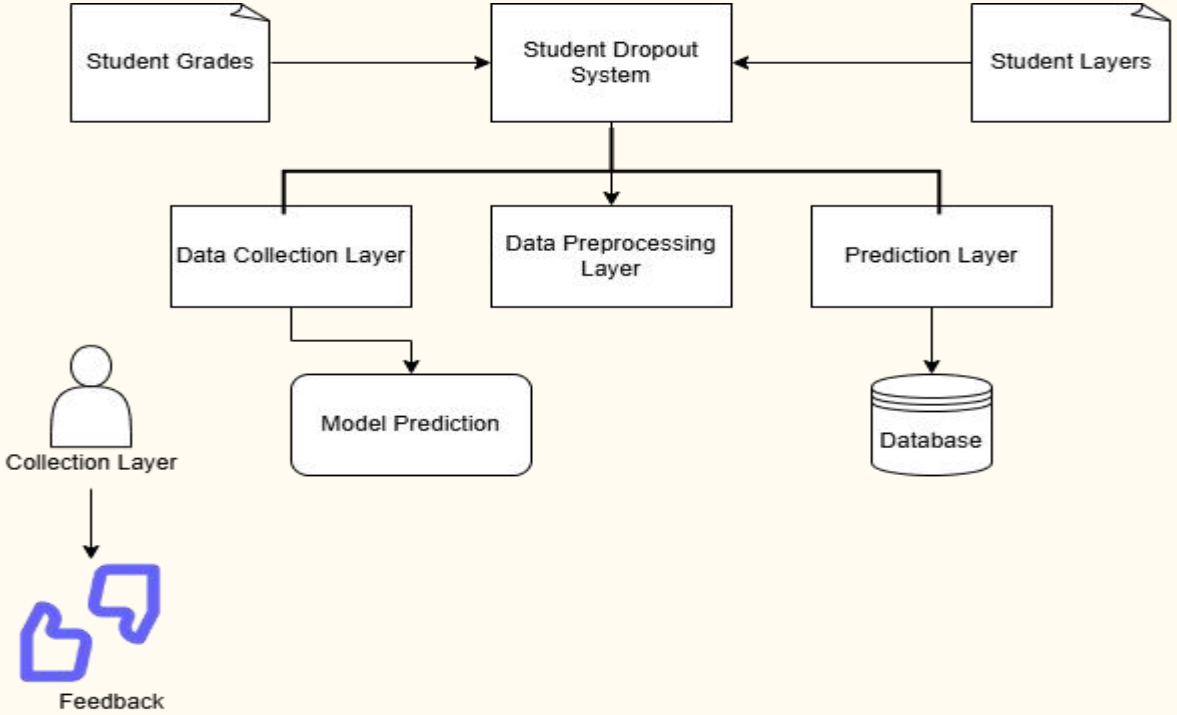


Fig 2.1 Machine Learning Analysis

2.2 Design(Flow Of Modules)

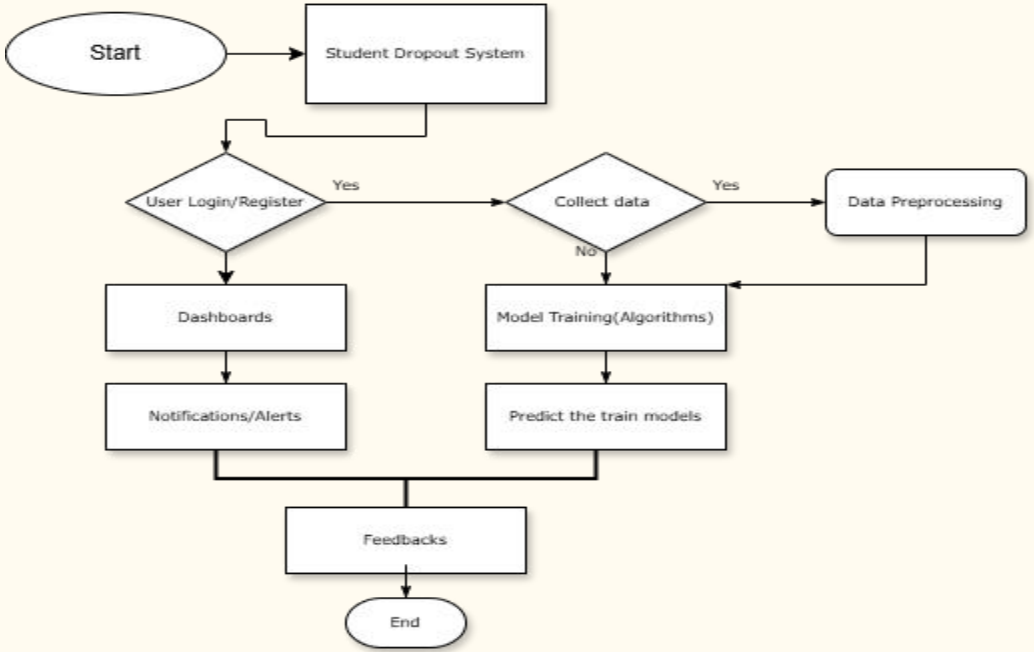


Fig 2.2 Workflow of the System

2.3 Description Of Use Case

- The Use case diagram showcases the admin involved and their interactions with the system. Each use case represents a specific functionality, ensuring that all user needs are addressed in the system's design.
- Use Case Diagrams are vital in project design as they depict the interactions between users and the system, facilitating better under standing of user requirements and system functionalities.

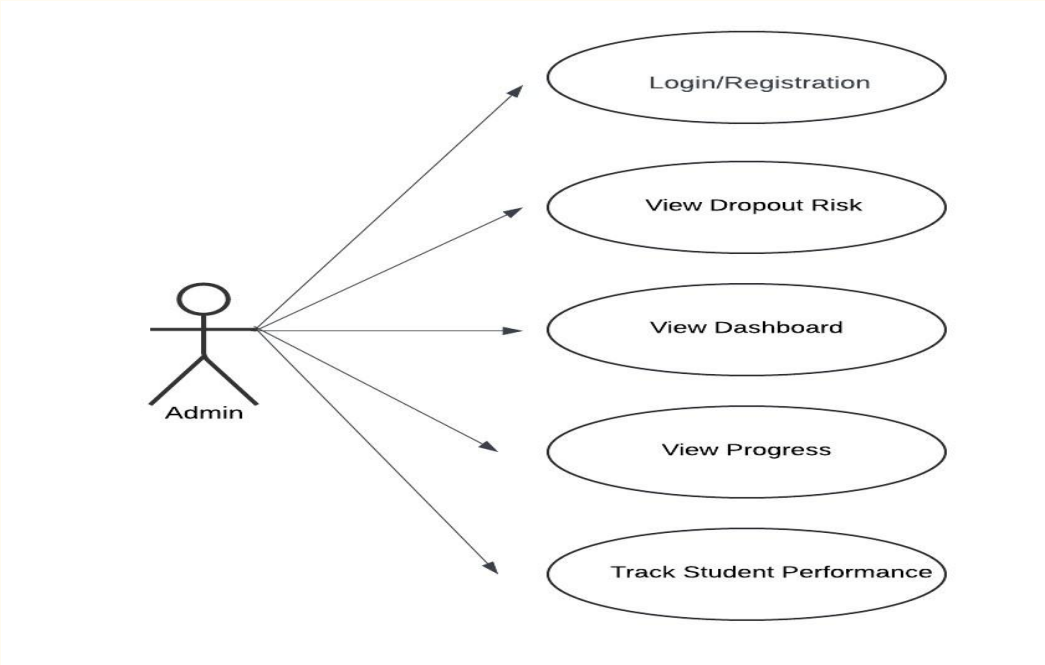


Fig 2.3 Use Case Diagram

2.4 Activity diagram

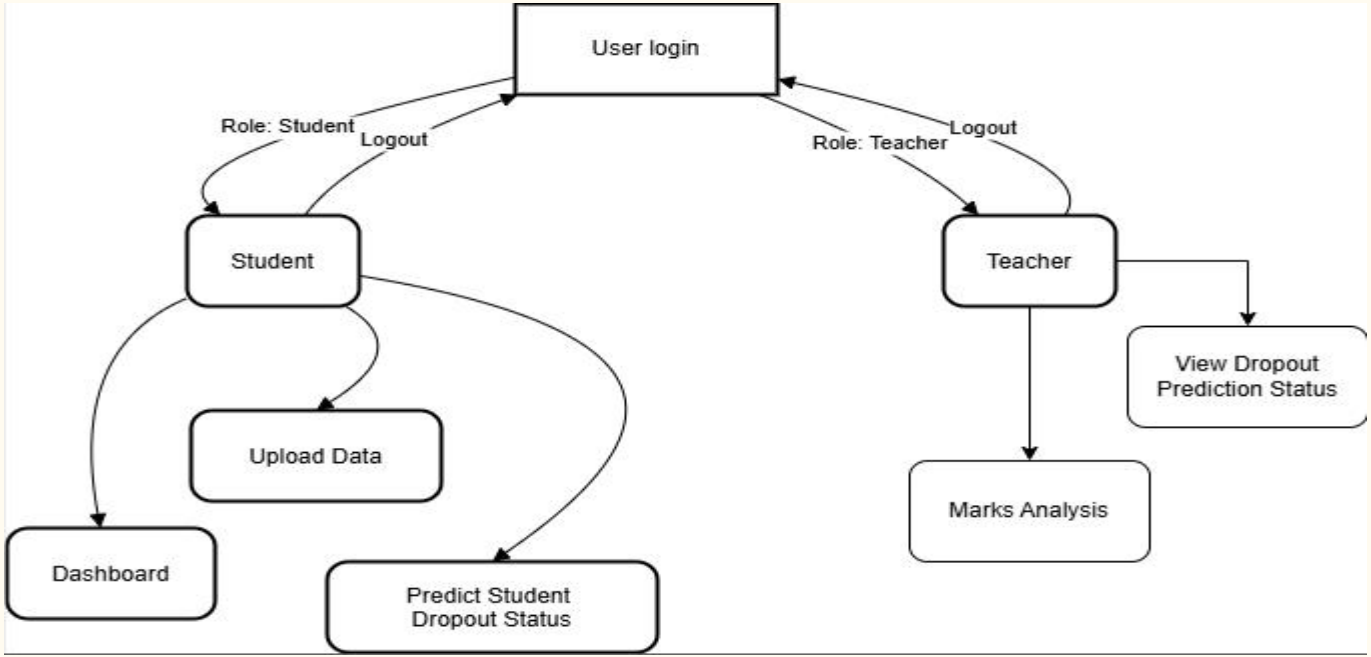


Fig 2.4 Activity Diagram

3. Implementation

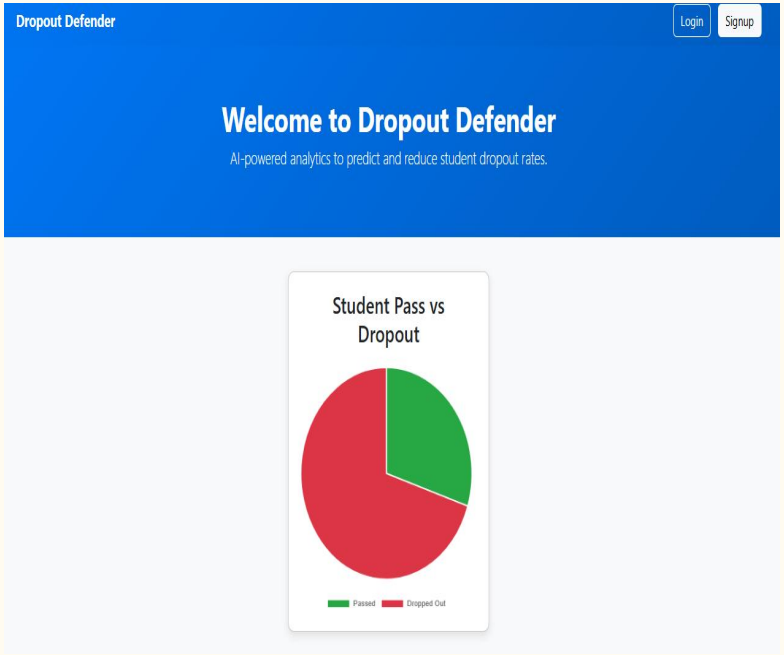


Figure 3.1 Home Page

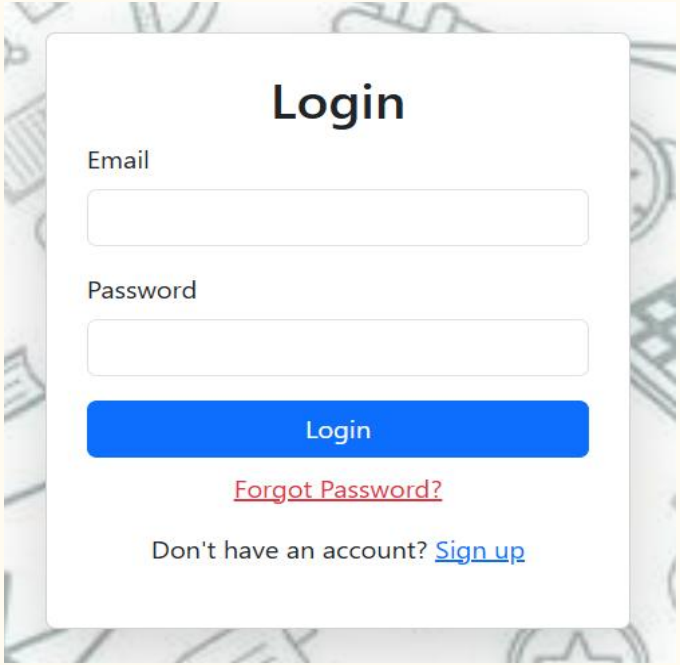


Fig 3.2 Login Page

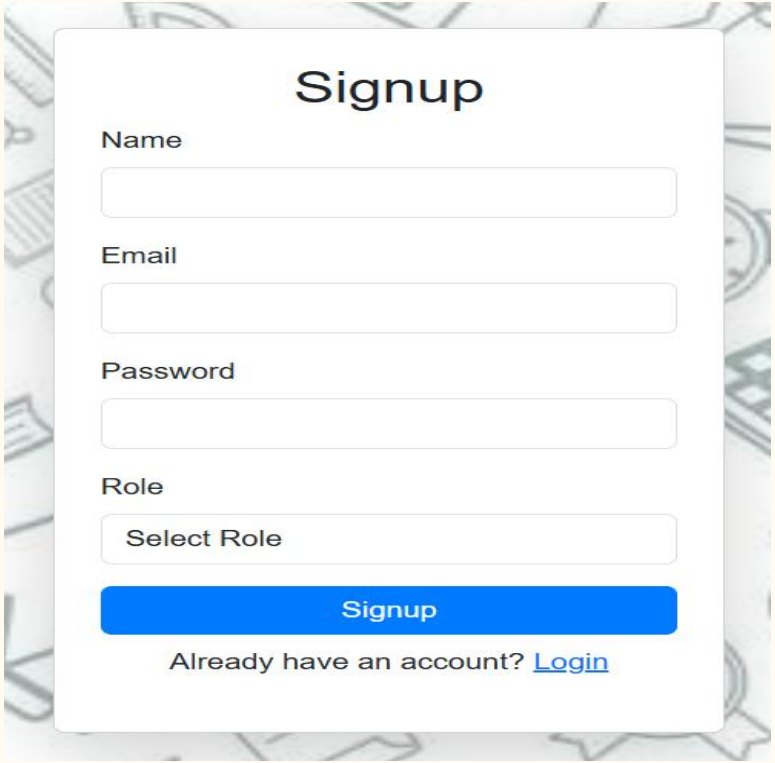


Figure 3.3 SignUp Page

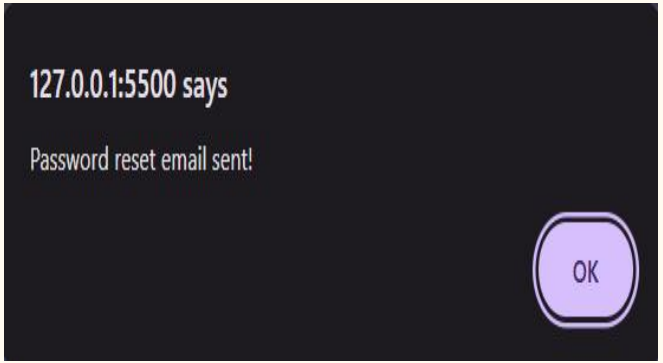


Figure 3.4 Alert Message

Fig 3.5 Password Reset Email

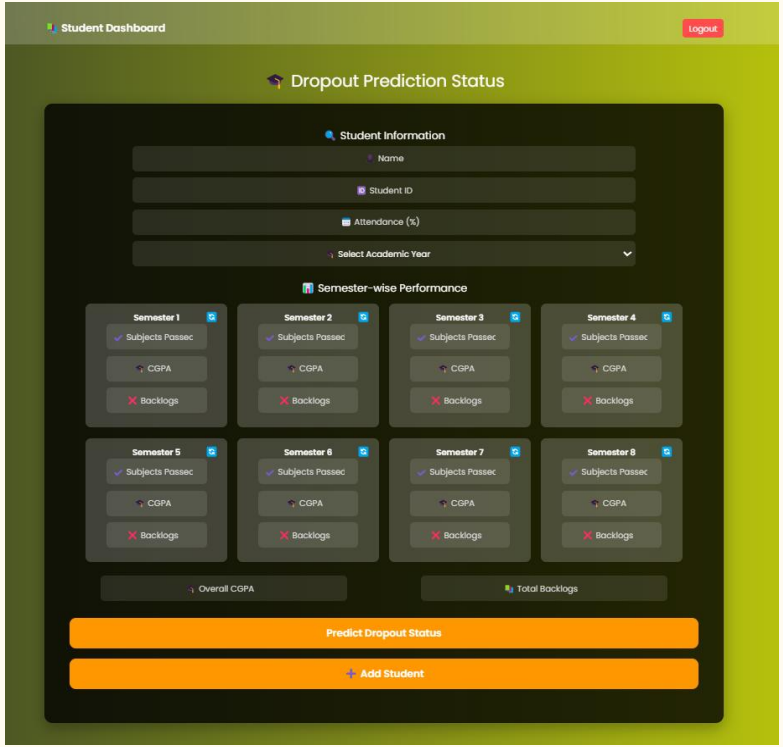
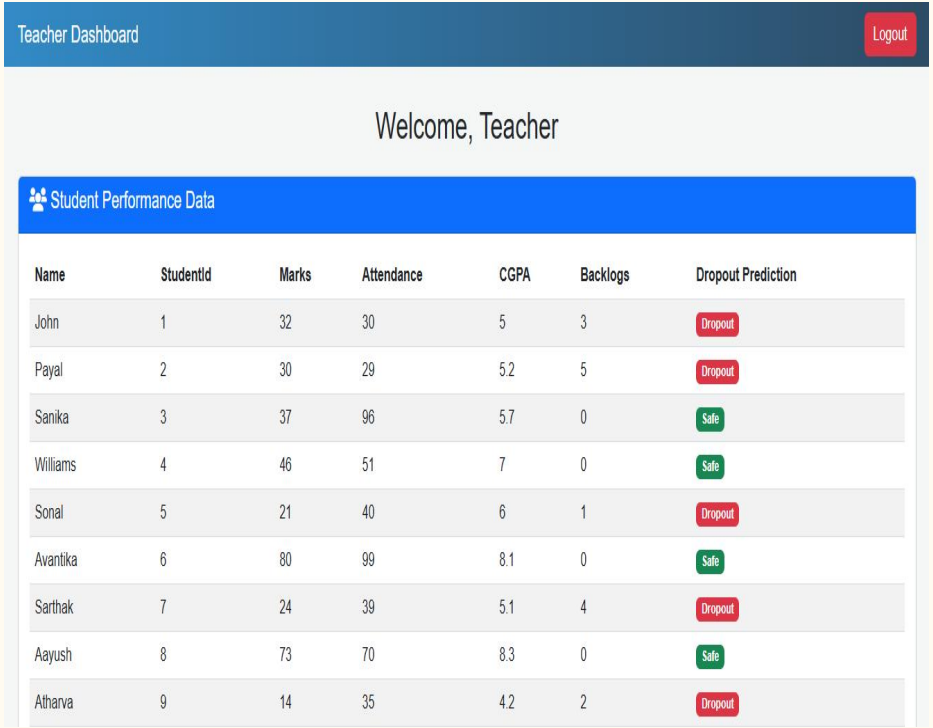


Fig 3.6 Student Dashboard



Name	StudentId	Marks	Attendance	CGPA	Backlogs	Dropout Prediction
John	1	32	30	5	3	Dropout
Payal	2	30	29	5.2	5	Dropout
Sanika	3	37	96	5.7	0	Safe
Williams	4	46	51	7	0	Safe
Sonal	5	21	40	6	1	Dropout
Avantika	6	80	99	8.1	0	Safe
Sarthak	7	24	39	5.1	4	Dropout
Aayush	8	73	70	8.3	0	Safe
Atharva	9	14	35	4.2	2	Dropout

Fig 3.7 Teacher Dashboard

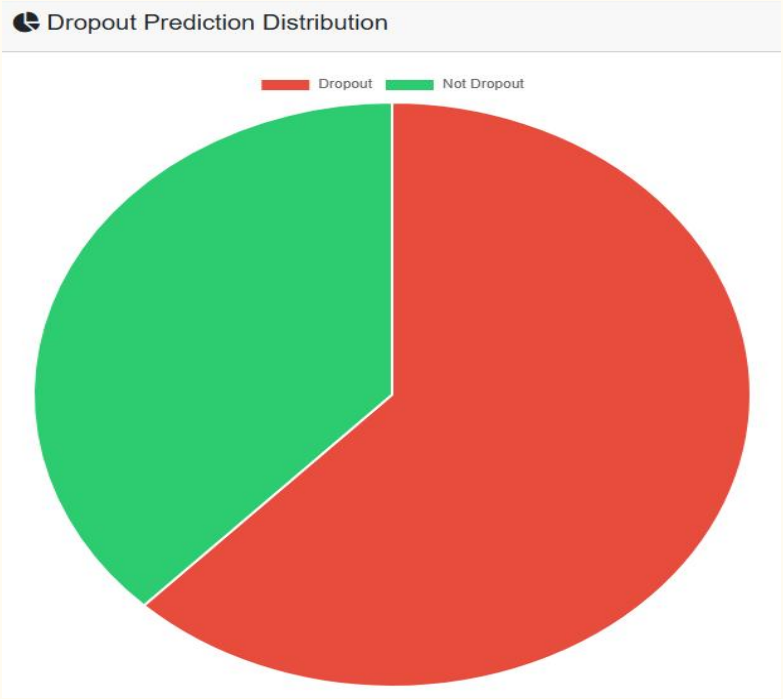


Fig 3.8 Dropout Prediction Status

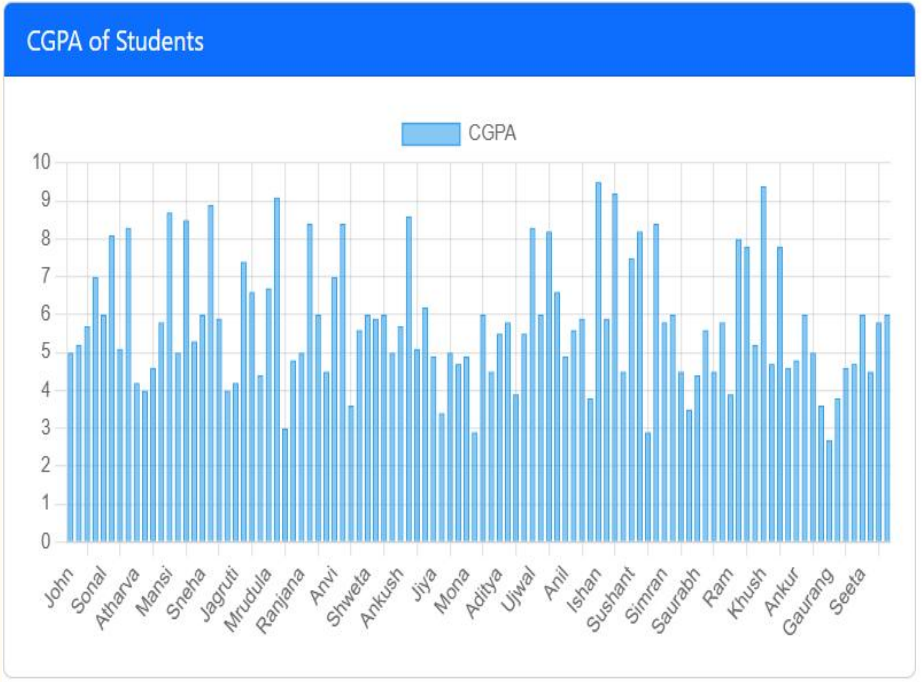


Fig 3.9 Student Marks Analysis

4. Testing

4.1 Software Testing

1. Software testing ensures that all components of the "Dropout Defender" system work together as expected.
2. It includes unit, integration, system, and regression testing to verify each part functions independently and as a whole.
3. Performance testing ensures the system can handle large datasets and multiple users without performance degradation.
4. Security and usability testing ensure that user data is protected and the interface is user-friendly.

4.2 Functional Testing

1. Functional testing verifies that the core features of the system perform as expected.
2. It includes testing the data preprocessing, model training, and accuracy calculations for dropout predictions.
3. User interactions, such as logging in, viewing dashboards, and logging out, are also tested for smooth operation.
4. The expected result is that the system provides accurate predictions and seamless user experience for students and teachers.

5. Result

5.1 Result

- In our project Random Forest Algorithm has more accuracy then other algorithm.
- Random Forest achieved the highest accuracy of 92%, making it the most reliable model for dropout prediction.
- It combines multiple decision trees, improving prediction accuracy and handling complex problems effectively.
- More trees in the model lead to better stability, reducing errors and improving problem-solving ability.
- With a precision of 90%, Random Forest minimizes false positives, ensuring accurate predictions.

Method	Accuracy	Precision
Random Forest	92%	90%
Decision Tree	89%	87%
Naive Bayes	86%	85%
Logistic Regression	85%	84%

6. Conclusion and Future Scope

6.1 Conclusion

- In conclusion, this research emphasizes the crucial role of machine learning in predicting student dropout rates and enabling timely interventions.
- The classification algorithms like decision trees, random forests, and SVM effectively identify at-risk students based on academic performance, attendance, and behavior.
- Integrating machine learning models with real-time data collection enhances prediction accuracy by considering key factors such as grades, attendance, and socio-economic status.
- For combining predictive techniques like logistic regression and neural networks helps create early warning systems that prevent dropouts.
- The choice of algorithm depends on dataset characteristics, with some models suited for smaller datasets and others excelling in complex data environments.

6.2 Future Scope

- The future of machine learning-based dropout prediction systems is highly promising as advancements in data collection and analysis continue.
- Expanding data sources to include social media activity, psychological assessments, and personalized learning experiences could enhance prediction accuracy. Incorporating advanced techniques like deep learning and reinforcement learning would improve the ability to analyze large, unstructured datasets, such as student feedback and interaction logs.
- The integration of real-time data streams and adaptive learning models could enable dynamic prediction systems that evolve with each student's progress. Ultimately, these advancements would lead to more proactive and personalized intervention strategies, significantly improving student retention in educational institutions.

References

1. Title: Predicting Student Dropout in Higher Education Using Machine Learning Techniques Authors: O. B. D. Ayoob, R. Shah, A. Ali Year: 2022
2. Title: Analyzing Student Dropout Factors Using Predictive Analytics Authors: C.T. Munoz, J. A. Asensio, M. A. Gonz´alez Year: 2023
3. Title: Machine Learning Approaches for Student Dropout Prediction: A Systematic Review Authors: P. M. Sharma, A. B. Rani, N. K. Gupta Year: 2023
4. Title: Early Detection of At-Risk Students: A Machine Learning Perspective Authors: F. A. Taleb, L. P. Trivedi, J. K. Singh Year: 2022

Thank You