**COMPREHENSIVE ACADEMIC PERFORMANCE MONITORING AND ANALYSIS SYSTEM FOR ENHANCED LEARNING OUTCOMES**

**Hemalatha M1, Janarthanan S2**

*1PG Student of Rathinam College of Arts & Science, Coimbatore, Tamil Nadu, India*

*2Assistant Professor, Department of Computer Science, Rathinam College of Arts & Science,*

*Coimbatore, Tamil Nadu, India*

***ABSTRACT:- In today’s educational landscape, monitoring and improving student performance is essential for academic success. This paper presents a Comprehensive Academic Performance and Analysis System that leverages machine learning and data visualization to predict student outcomes and provide actionable insights. The system utilizes a Random Forest Classifier, trained on a dataset containing key academic and behavioral indicators such as Math Score, Reading Score, Writing Score, Attendance Rate, Study Hours, Parent Education, Economic Status, and Extra Tutoring, to predict whether a student will pass or fail.*** ***The proposed system is implemented as an interactive web application using Streamlit, enabling educators, students, and parents to visualize performance trends, compare individual progress with class averages, and receive personalized recommendations for improvement.***

***Keywords: Student Performance, Machine Learning, Random Forest Classifier, Predictive Analytics, Educational Data Mining, Streamlit, Data Visualization.***

**I.INTRODUCTION**

The rapid advancements in data science and machine learning have revolutionized the way student performance is analysed and improved. This research introduces the Student Performance Analyzer, a Streamlit-powered web application that leverages Random Forest Classifier to predict student success based on a combination of academic scores, behavioural factors, and socio-economic conditions. By analysing key parameters such as Math, Reading, and Writing scores, Attendance Rate, Study Hours, Parental Involvement, and Classroom Environment, this system provides a data-driven approach to early intervention and personalized recommendations. Unlike conventional evaluation methods that rely solely on past academic performance, this model identifies hidden patterns that contribute to student outcomes. The interactive dashboard enables educators to track individual progress, compare student performance against peers, and receive actionable insights to improve learning outcomes. This work aims to empower educators and institutions with a scalable, AI-driven solution for better student assessment and targeted academic interventions, ultimately fostering a more inclusive and effective learning environment.

**II.OBJECTIVE**

The objective of this research is to develop an AI-driven Student Performance Analyzer that integrates machine learning and data visualization to assess, predict, and enhance student outcomes. By utilizing a Random Forest Classifier, the system predicts whether a student will pass or fail based on academic scores, attendance, and socio-economic factors. Additionally, it identifies key influences such as study habits, parental involvement, economic background, and school resources, helping educators understand their impact on student success. The project further aims to provide personalized improvement strategies, offering tailored recommendations to enhance learning based on individual performance. Through an interactive Streamlit dashboard, educators can gain real-time insights, visualize trends, and detect at-risk students for early intervention. Furthermore, the system supports comparative analysis, allowing students and teachers to benchmark individual performance against class averages to identify learning gaps. Ultimately, this project bridges the gap between data science and education, equipping students and educators with intelligent, data-driven tools to optimize academic success.

**III.PROBLEM STATEMENT**

The increasing complexity of student performance assessment poses significant challenges for educators in identifying struggling students and providing timely interventions. Traditional evaluation methods often fail to capture the multifaceted nature of academic success, which is influenced by various factors such as subject scores, attendance, study hours, parental involvement, and socio-economic background. Without a structured analytical approach, institutions lack the necessary insights to predict student outcomes accurately and implement data-driven strategies for improvement. This project addresses this issue by developing a Student Performance Analyzer using machine learning (Random Forest Classifier) and Streamlit. The system predicts whether a student will pass or fail based on key academic and behavioral indicators while providing educators with an interactive dashboard to visualize performance trends and receive personalized improvement suggestions. By integrating predictive analytics with real-time data visualization, this solution empowers educators to make informed decisions, enhance student learning experiences, and minimize academic failures.

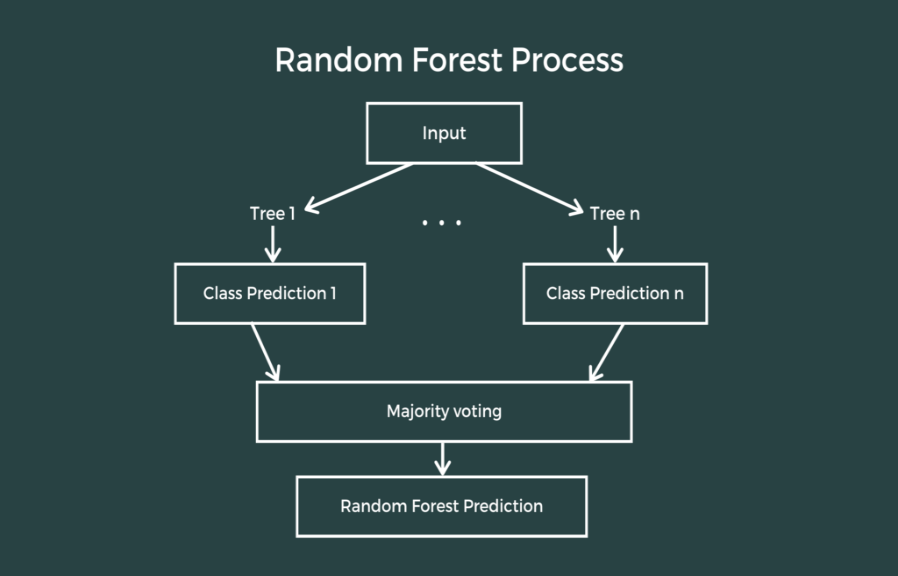
**IV.METHODOLOGY**

**A . MACHINE LEARNING**

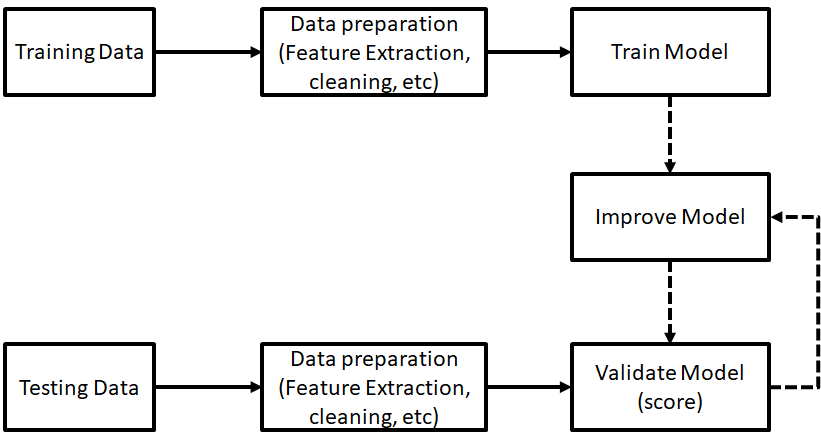
Machine learning is a branch of artificial intelligence that enables computers to identify patterns and make predictions based on data. In the Student Performance Analyzer, machine learning techniques are applied to predict student outcomes and provide actionable insights for academic improvement. The methodology follows a structured approach to ensure accurate analysis and prediction. (1) Data Collection: The dataset includes student performance attributes such as Math Score, Reading Score, Writing Score, Attendance Rate, Study Hours, Parent Education, Economic Status, and Class Size. The data is collected from educational records and processed to ensure consistency. (2) Data Analysis: Exploratory Data Analysis (EDA) is conducted to understand data distribution, identify trends, and detect missing values or anomalies. Statistical methods and visualizations are used to extract meaningful insights from student performance metrics. (3) Feature Engineering & Model Selection: Key features influencing student outcomes are selected, and categorical variables are encoded. The Random Forest Classifier is chosen as the primary machine learning model due to its high accuracy and robustness in handling complex relationships within the data. (4) Model Training and Evaluation: The dataset is split into training and testing sets, and the Random Forest model is trained to classify students as pass or fail. Model performance is evaluated using accuracy, precision, recall, and F1-score, ensuring optimal predictive capability. (5) Deployment and Visualization: The trained model is deployed in an interactive Streamlit-based web application, allowing users to upload student data, visualize performance trends, and generate predictive insights. The application provides a user-friendly dashboard for educators and administrators to make informed decisions and support students effectively. By leveraging machine learning and data-driven approaches, this project aims to enhance academic performance monitoring and assist in early intervention strategies for struggling students.

**B. RANDOM FOREST CLASSIFIER:**

The Random Forest Classifier is a powerful ensemble learning algorithm widely used for classification tasks due to its robustness, accuracy, and ability to handle high-dimensional data. It operates by constructing multiple decision trees during training and aggregating their outputs to improve predictive performance. Each tree in the forest is trained on a random subset of the dataset, and its final classification is determined by a majority vote among the trees. This reduces the risk of overfitting, making the model more generalizable to unseen data. The Random Forest Classifier is particularly effective when dealing with datasets that contain both numerical and categorical features, as it can automatically handle missing values and capture complex relationships between variables. In the Student Performance Analyzer, Random Forest is used to predict whether a student will pass or fail based on factors such as Math Score, Reading Score, Writing Score, Attendance Rate, Study Hours, Parent Education, and Economic Status. Its ability to provide feature importance scores also helps educators identify the most influential factors affecting student outcomes, enabling targeted interventions for academic improvement.



**V. SYSTEM ARCHITECTURE**



**VI. IMPLEMENTATION**

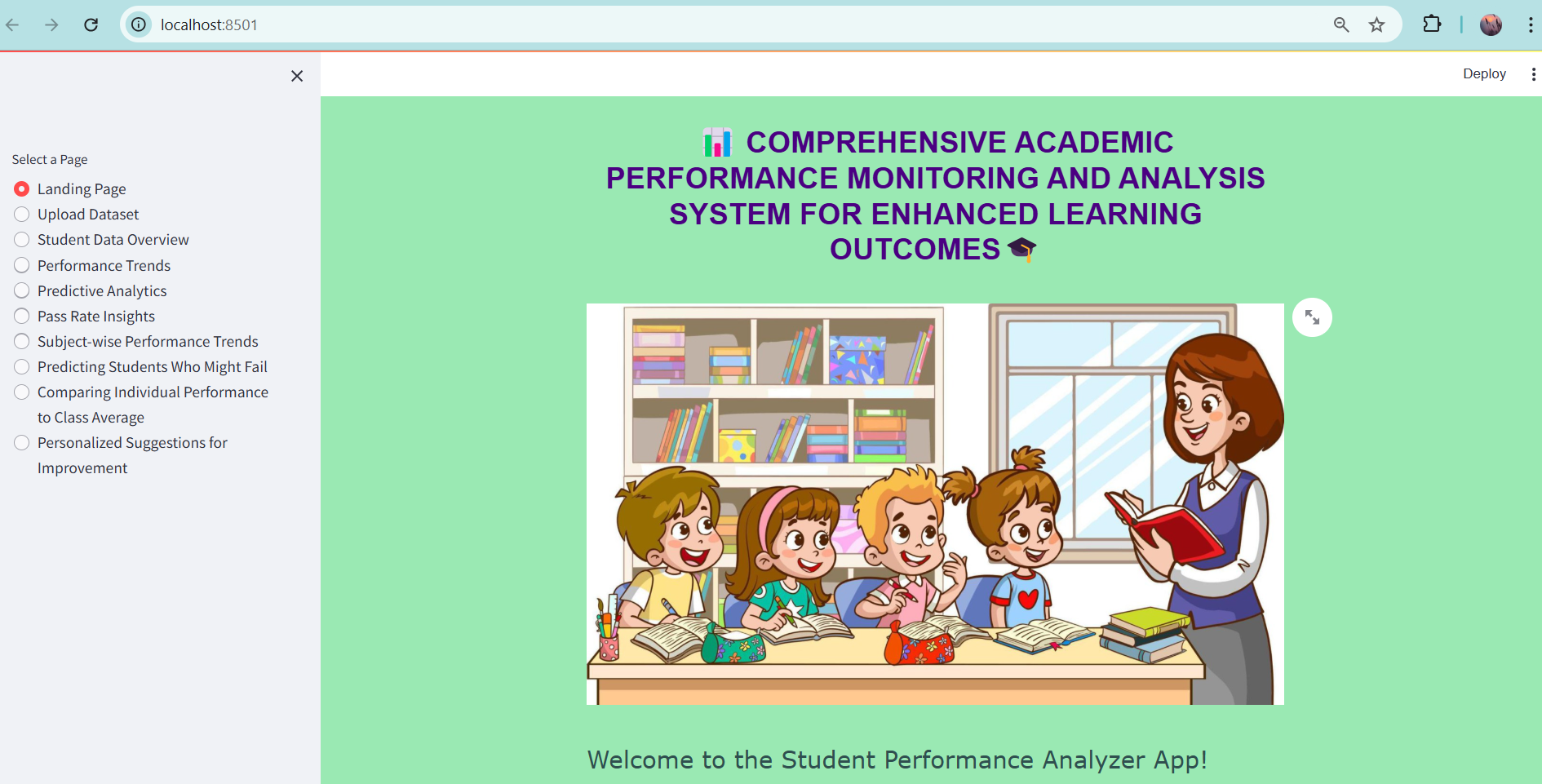


Fig 1. Home page

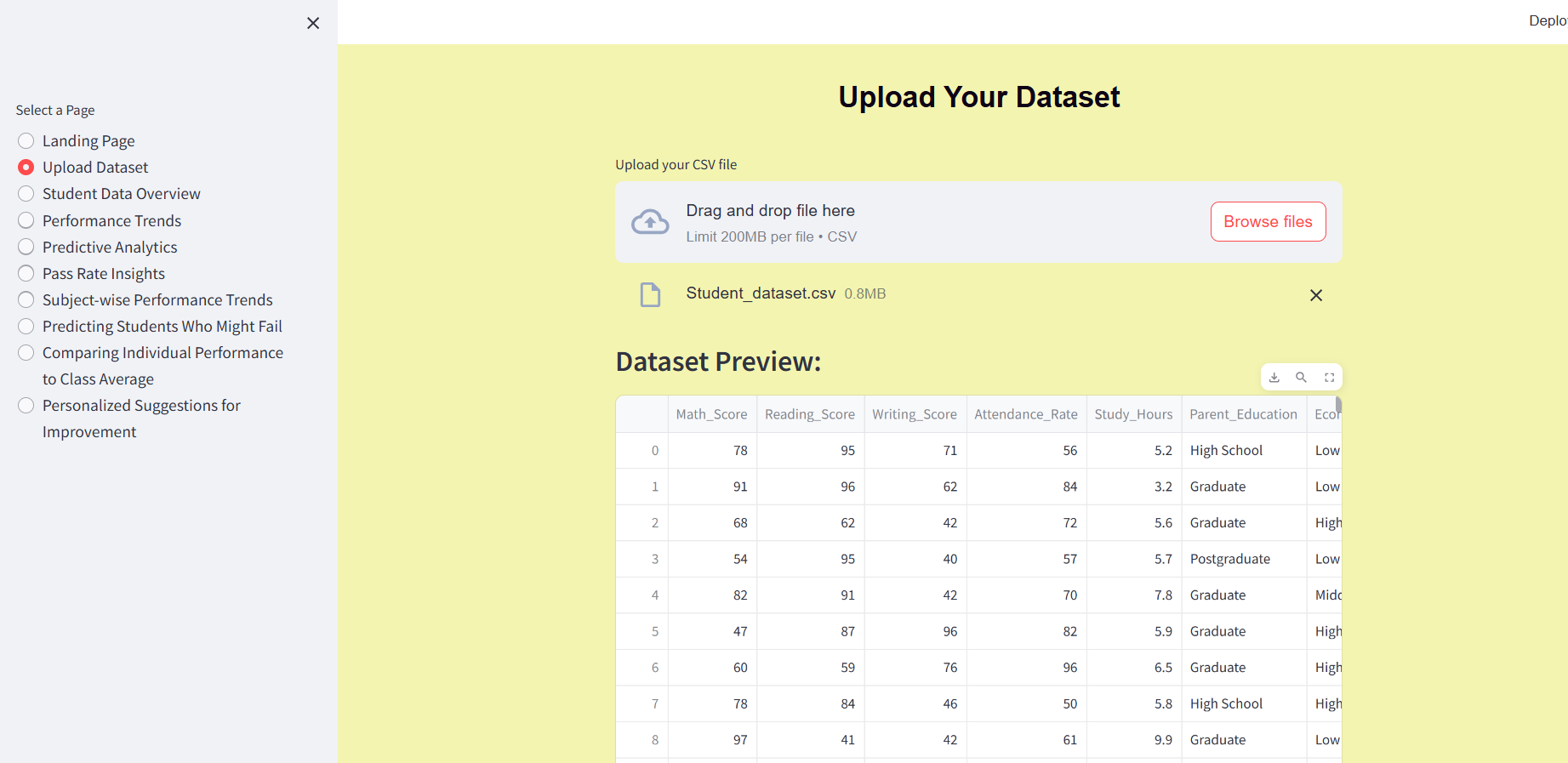


Fig 2. Uploading dataset

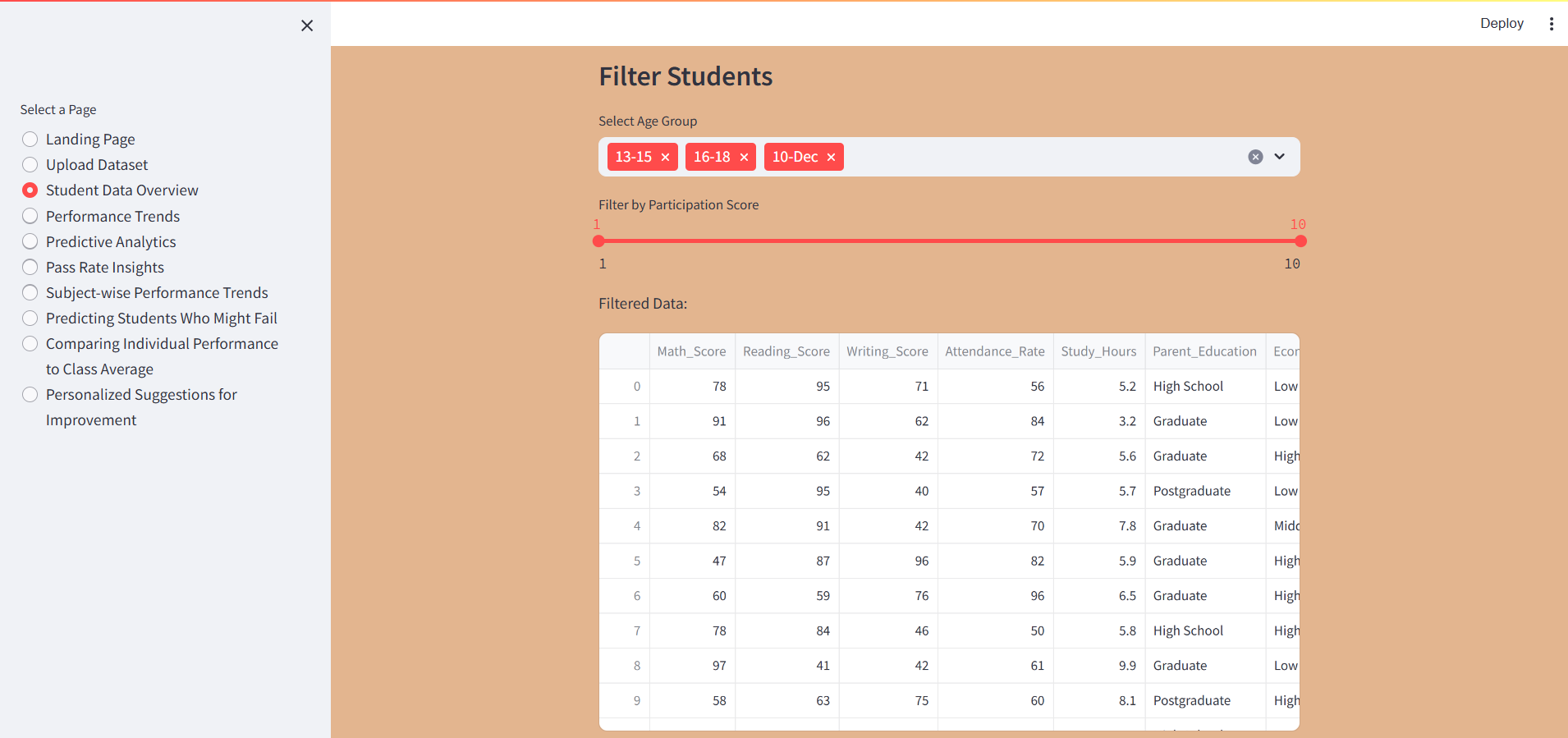


Fig 3. Overview of student data

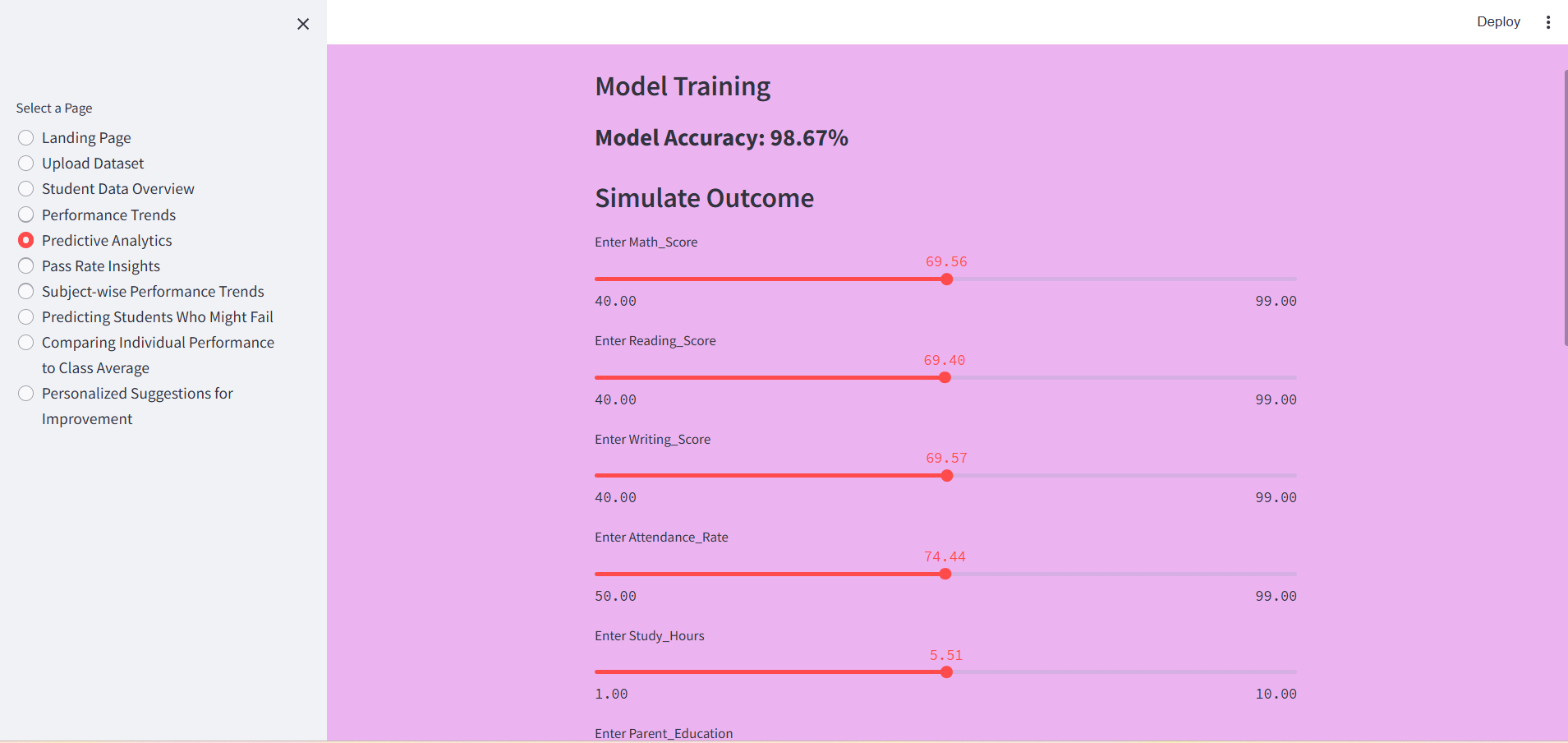


Fig 4. Model training and Accuracy

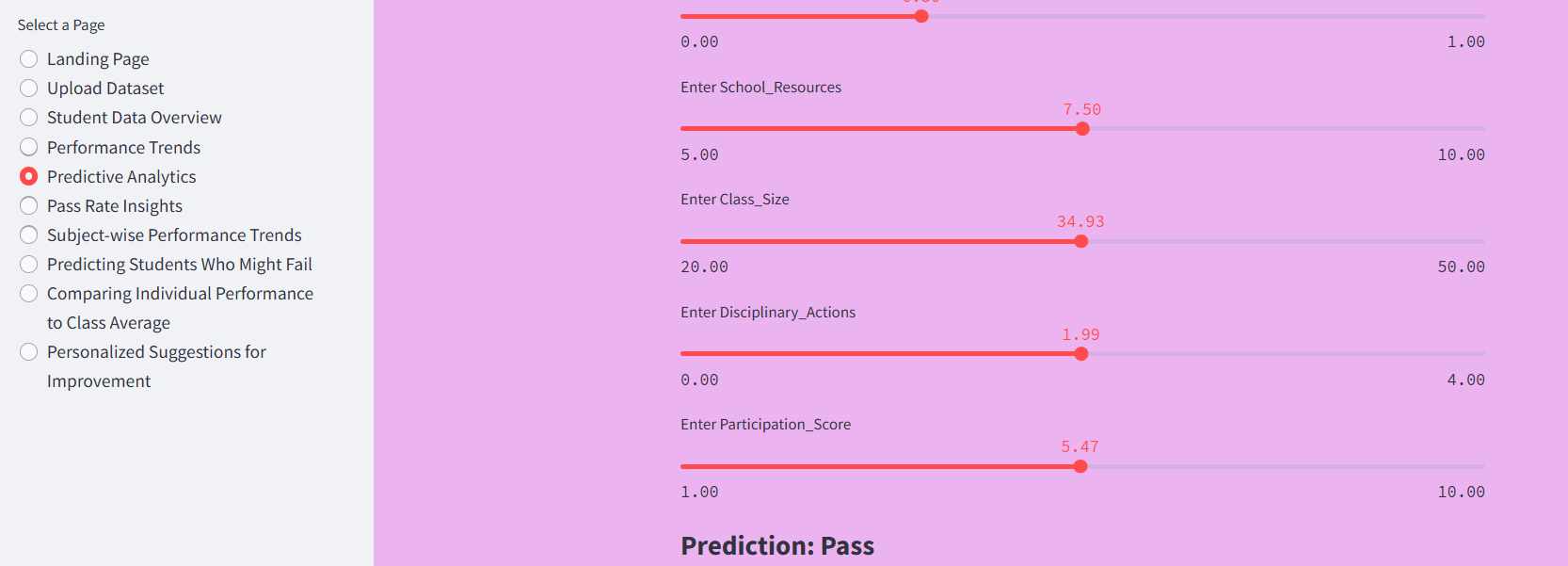


Fig 5. Predictive Analytics

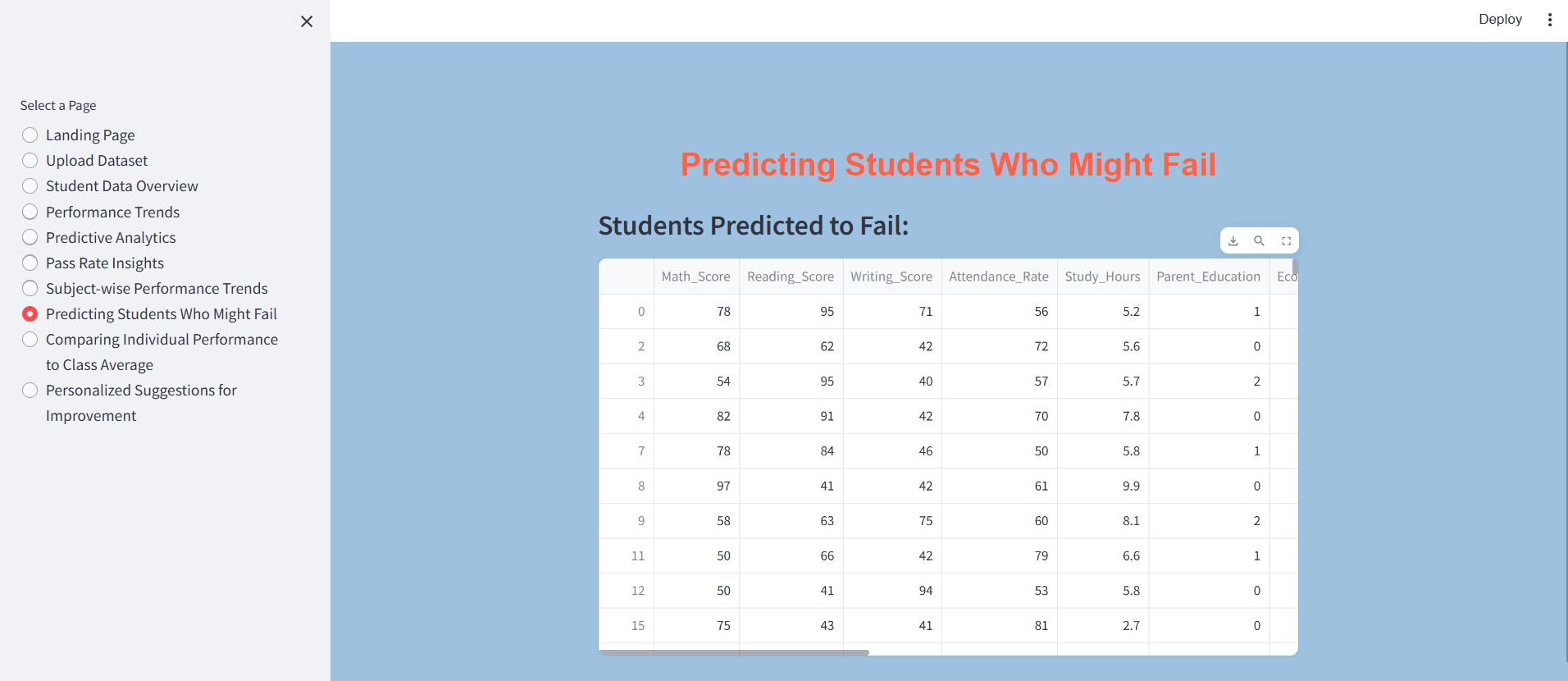


Fig 6. Fail prediction

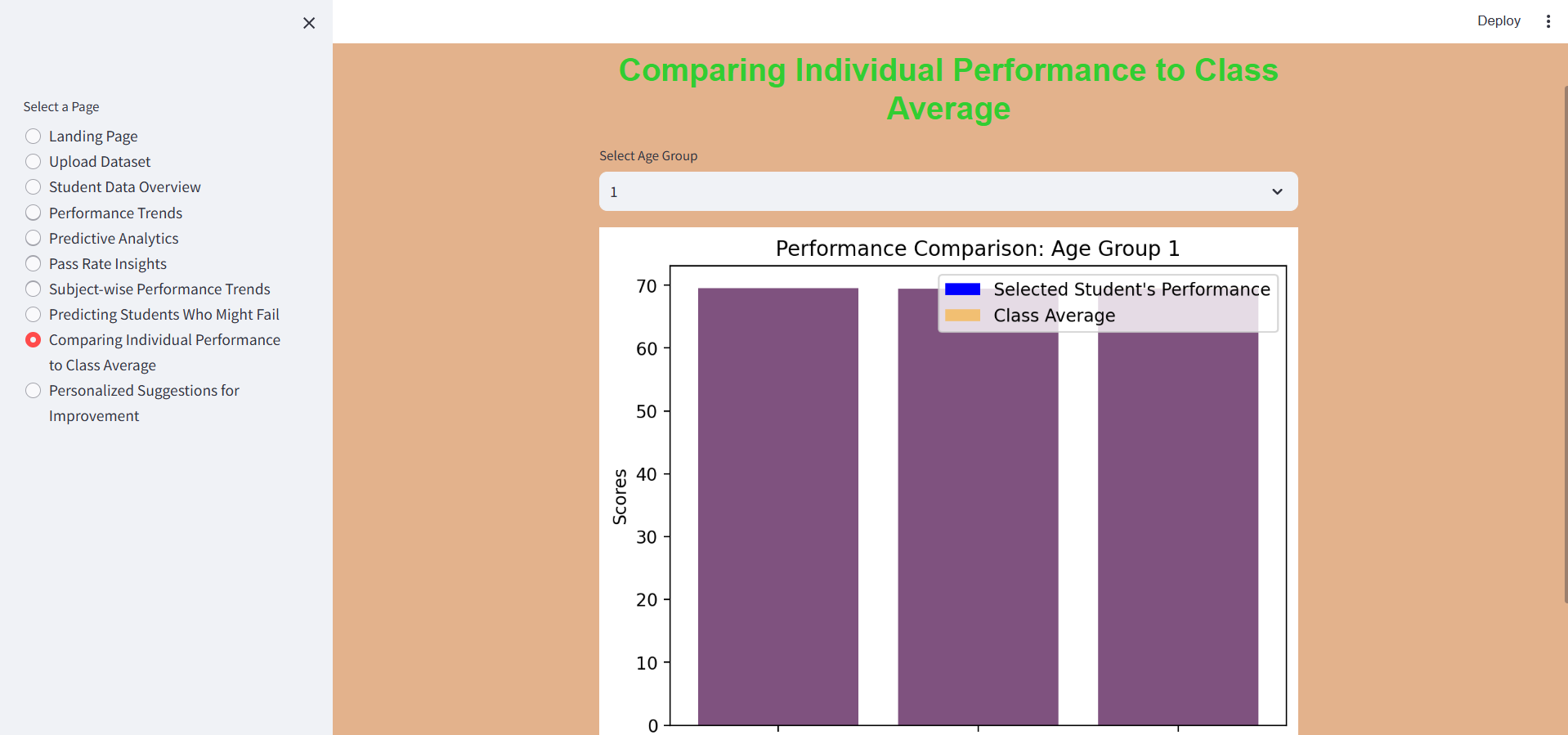


Fig 7. Comparison of Individuals performance to class Average

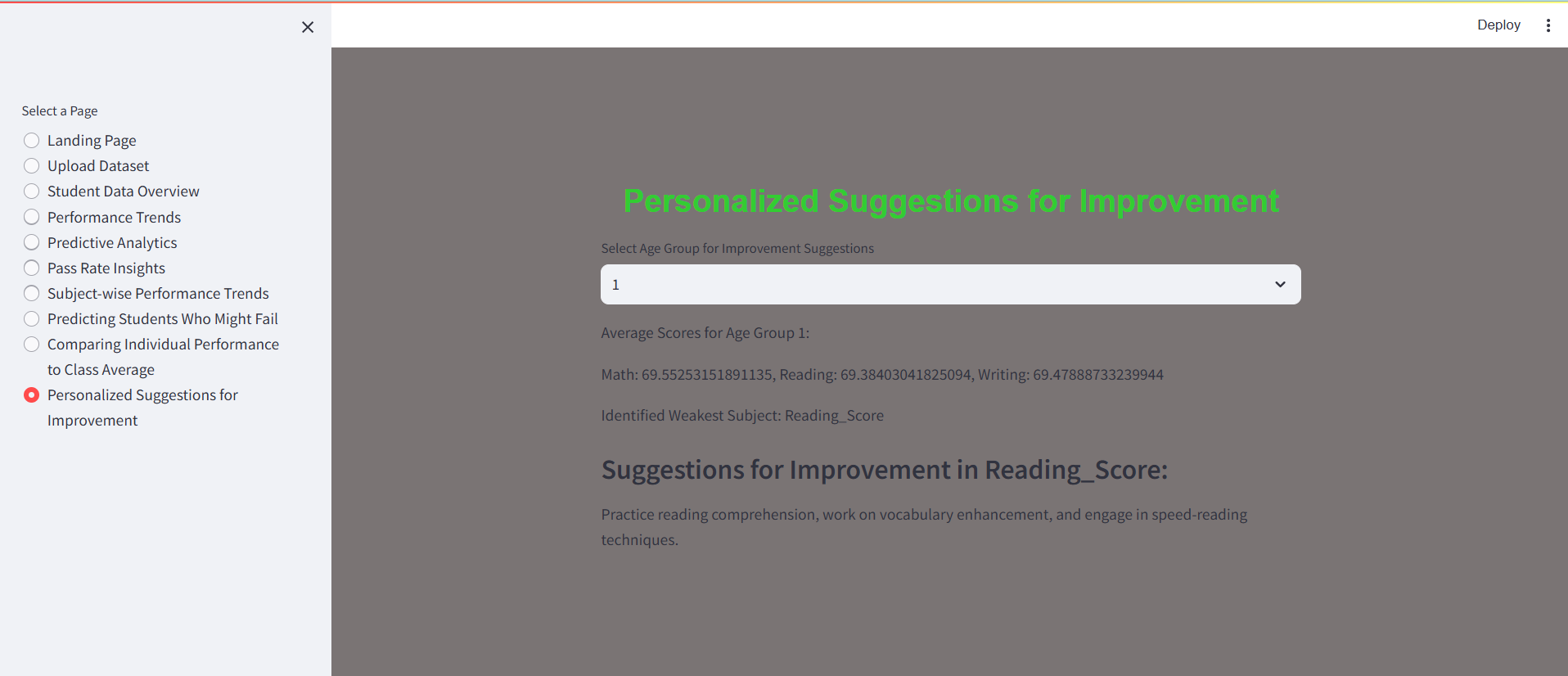


Fig 8. Future Improvement

**VII.CONCLUSION**

By leveraging the Random Forest Classifier, the model demonstrates high accuracy (98.67%) in identifying students at risk of failing based on factors such as academic scores, attendance, study hours, and socio-economic conditions. The project not only provides predictive insights but also enables educators to make data-driven decisions for student improvement through visual analysis and personalized recommendations.

**VIII.REFERENCES**

[1] D.Solomon, S. Patil, and P. Agrawal, Predicting performance and potential difficulties of university student using classi cation: Survey paper, Int. J. Pure Appl. Math, vol. 118, no. 18, pp. 27032707, 2023.

[2] E. Alyahyan and D. Dütegör, Predicting academic success in higher education: Literature review and best practices, Int. J. Educ. Technol. Higher Educ., vol. 17, no. 1, Dec. 202.

[3] V. L. Miguéis, A. Freitas, P. J. V. Garcia, and A. Silva, Early segmentation of students according to their academic performance: A predictive modelling approach, Decis.SupportSyst.,vol.115,pp. 3651,Nov.2018

[4] Y.Zhang,Y.Yun,H.Dai,J.Cui,andX.Shang, Graphsregularizedrobust matrix factorization and its application on student grade prediction, Appl. Sci., vol. 10, p. 1755, Jan. 2020.

[5] H.Aldowah,H.Al-Samarraie,and W.M.Fauzy, Educationaldatamining and learning analytics for 21st century higher education: A review and synthesis, Telematics Informat., vol. 37, pp. 1349, Apr. 2019.

[6] K. L.-M. Ang, F. L. Ge, and K. P. Seng, Big educational data & analytics: Survey, architecture and challenges, IEEE Access, vol. 8, pp. 116392116414, 2020.

[7] A. Hellas, P. Ihantola, A. Petersen, V. V. Ajanovski, M. Gutica, T. Hynninen,A.Knutas,J.Leinonen,C.Messom,andS.N.Liao, Predictingacademic performance: A systematic literature review, in Proc. 23rd Annu. Conf. Innov. Technol. Comput. Sci. Educ., Jul. 2018, pp. 175199.

[8] L. M. Abu Zohair, Prediction of students performance by modelling small dataset size, Int. J. Educ. Technol. Higher Educ., vol. 16, no. 1, pp. 18, Dec. 2019, doi: 10.1186/s41239-019-0160-3.