

Student Wellbeing Dropout Prediction System

A Vasavi Sujatha¹, Janapala Ruchitha², Neela Sathvika³, Perala Srija⁴

¹Assistant Professor; Department Of Information Technology Bhoj Reddy Engineering College For Women
Hyderabad India

^{2,3,4}B.Tech Student's; Department Of Information Technology Bhoj Reddy Engineering College For Women
Hyderabad India

Mail Id's; janapalaruchitha@gmail.com, sathvikanela07@gmail.com, peralaserija@gmail.com

Abstract

The Student Wellbeing Dropout Prediction System is an integrated platform that combines academic tracking, attendance monitoring, and health management within a unified application. Developed using a Flask-based API and enhanced with Machine Learning techniques, the system enables institutions to monitor, analyze, and support student wellbeing more effectively. The platform aggregates data on academic performance, attendance, lifestyle habits, and physical and emotional health to provide a holistic view of each student. Machine Learning models process this data to generate risk scores and classify students into low- and high-risk categories. High-risk cases trigger automated alerts for educators, enabling timely intervention. An interactive dashboard visualizes trends, risk levels, and key insights, while report generation features support documentation and follow-up. By integrating predictive analytics with real-time monitoring, the system facilitates early identification of at-risk students and supports proactive decision-making. Additionally, the platform helps uncover patterns and correlations between academic outcomes and wellbeing factors, enabling institutions to design targeted preventive strategies. Overall, the system aims to reduce dropout rates and foster a more supportive and data-driven educational environment.

Keywords :

Student Dropout Prediction, Academic Performance Monitoring, Wellbeing Analytics, Machine Learning, Attendance Tracking, Risk Assessment, Flask API, Dashboard Visualization, Predictive Analytics

Introduction :

Student dropout remains a significant challenge in higher education institutions, affecting both academic outcomes and institutional performance. Students discontinue their studies due to a variety of factors, including poor academic performance, irregular attendance, psychological stress, health-related issues, and personal circumstances. Among these, health and wellbeing factors—such as inadequate sleep, emotional stress, and physical conditions—often have a substantial impact on academic performance, yet they are rarely integrated into traditional academic

monitoring systems. This lack of integration limits the ability of institutions to identify at-risk students at an early stage. To address this gap, the Student Wellbeing Dropout Prediction System is proposed as an integrated solution that combines academic records, attendance data, and basic health-related information within a single platform. The system leverages Machine Learning techniques to analyze multi-dimensional student data and generate risk scores, enabling the classification of students into different risk categories. This data-driven approach reduces reliance on manual observation and supports more accurate and timely identification of students who may be at risk of dropping out. Furthermore, the system includes an alert mechanism that notifies educators when students are identified as high-risk, allowing for prompt intervention. A user-friendly dashboard presents key metrics, trends, and risk indicators, facilitating easier interpretation and decision-making. By consolidating diverse student-related factors into a unified framework, the proposed system enhances monitoring capabilities, supports proactive intervention strategies, and contributes to improving student retention and overall academic success.

Related Work

Survey-Based Analysis

To evaluate the practical need for early identification of at-risk students, a structured survey was conducted among key stakeholders in higher education, including undergraduate and postgraduate students, faculty members, and academic administrators. The objective was to understand existing challenges in monitoring student wellbeing and to assess the demand for an integrated predictive system. The findings indicate that a significant proportion of students experience challenges that negatively influence their academic performance, such as low grades, irregular attendance, stress, and health-related concerns. However, only a limited number reported receiving timely institutional support, with most relying on self-management or informal assistance. A strong majority of respondents expressed the need for a unified system that integrates academic, attendance, and wellbeing data, indicating that such a platform would be highly beneficial for continuous monitoring and early intervention. Faculty

and administrative participants highlighted the limitations of manual tracking methods, noting that these approaches are time-intensive and often ineffective in identifying at-risk students at an early stage. Most respondents showed clear interest in a system capable of generating risk scores, issuing automated alerts, and presenting insights through visual dashboards. Additionally, participants emphasized the importance of incorporating lifestyle and emotional wellbeing indicators—such as sleep patterns, stress levels, and nutrition—to enable more personalized and accurate assessment. Overall, the survey results reveal a clear gap in current monitoring practices and validate the necessity of an integrated Student Wellbeing and Dropout Prediction System. The findings support the development of a data-driven platform that enables early detection, targeted intervention, and improved student retention.

Requirement Analysis

Functional Requirements

The proposed system is designed as a multi-user platform that supports students, teachers, and administrators through role-based access control. Each user interacts with a customized interface aligned with their responsibilities, ensuring both security and usability. Students are able to input and review their academic and wellbeing information, while teachers can monitor student performance, track risk levels, and receive alerts for critical cases. Administrators are responsible for managing user data, overseeing system operations, and generating analytical reports. The system enables secure authentication for all users and provides administrators with the ability to manage student records efficiently. It supports data collection related to academic performance, attendance, and health indicators, which are processed using Machine Learning algorithms to compute risk scores. Based on these scores, students are categorized into risk levels, and automated notifications are sent to teachers when high-risk conditions are detected. Additionally, the platform allows for report generation and download, facilitating documentation and follow-up actions.

Non-Functional Requirements

The system is designed to ensure high performance, security, scalability, reliability, availability, and usability. It provides a seamless and intuitive user experience, allowing users to perform tasks such as

data entry, monitoring, and report generation efficiently. The platform prioritizes data protection by implementing secure authentication mechanisms and encrypted communication protocols, ensuring the confidentiality of sensitive academic and health information. From a performance perspective, the application must respond quickly to user interactions, with minimal latency in data processing and visualization. Scalability is addressed through a modular system design that can accommodate increasing numbers of users and data without degradation in performance. Reliability is ensured by maintaining stable system behavior under varying workloads, minimizing failures during critical operations such as risk analysis and alert generation. Usability is a key consideration, with a simple and consistent interface that accommodates users with varying levels of technical expertise. The system is designed to be continuously available, ensuring uninterrupted access to essential services such as login, dashboards, and reporting features. Together, these requirements ensure that the platform remains efficient, secure, and dependable in real-world academic environments.

Computational Resources

The implementation of the system requires both software and hardware resources to support development, deployment, and execution. On the software side, the front-end is developed using standard web technologies, while the backend is built using a Flask-based API framework. Data is stored and managed using a relational database system, and Machine Learning functionalities are implemented using widely adopted data science libraries for analysis, modeling, and visualization. Development and testing are carried out using an integrated code editor environment. From a hardware perspective, the system requires a computing environment with a modern multi-core processor, sufficient memory capacity to handle data processing tasks, and adequate storage for maintaining datasets and application files. These resources ensure smooth system performance and support future scalability.

Design

System Architecture

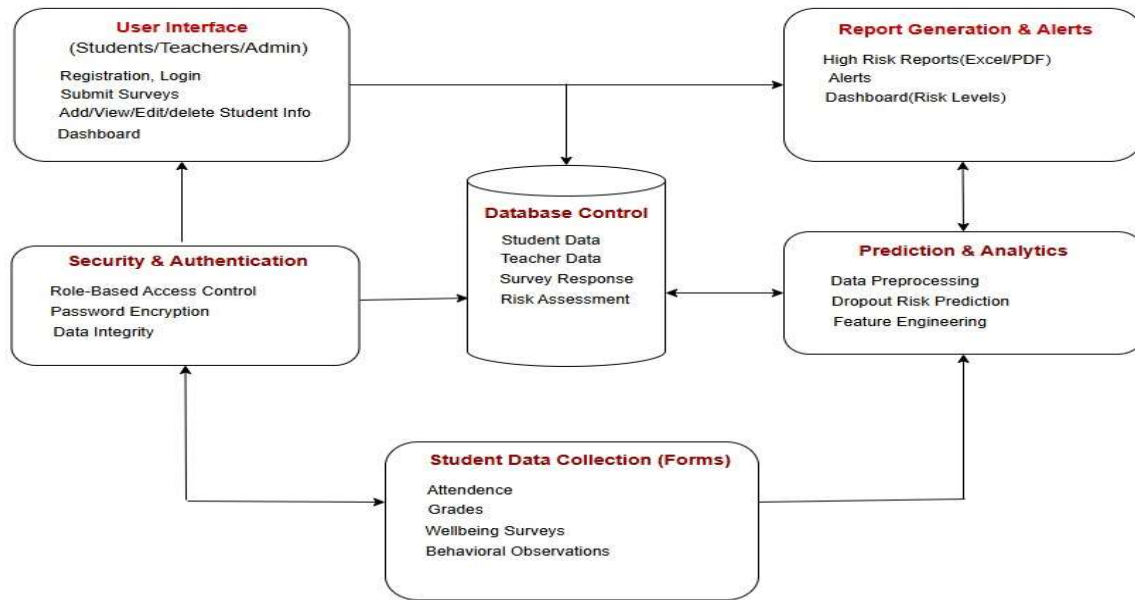


Fig. 1 System Architecture

The architecture of the Student Wellbeing Dropout Prediction System is designed as a modular and data-driven framework that integrates user interaction, data processing, and predictive analytics into a unified platform. The system follows a layered architecture consisting of the presentation layer, application layer, and data layer. The presentation layer provides user interfaces for students, teachers, and administrators, enabling role-based access and interaction with the system. The application layer, implemented using a Flask-based API, handles business logic, data processing, and communication between components. The data layer manages storage and retrieval of student-related information, including academic

records, attendance, and health indicators. The overall workflow begins with data collection from students and teachers through structured forms. This data is stored in a centralized database and later combined into a unified dataset. The system processes the collected information to compute risk-related features, which are then used by Machine Learning models to classify students into risk categories. Based on the results, the system generates alerts and displays insights through an interactive dashboard. This architecture ensures scalability, maintainability, and efficient data flow across all system components.

Technical Architecture

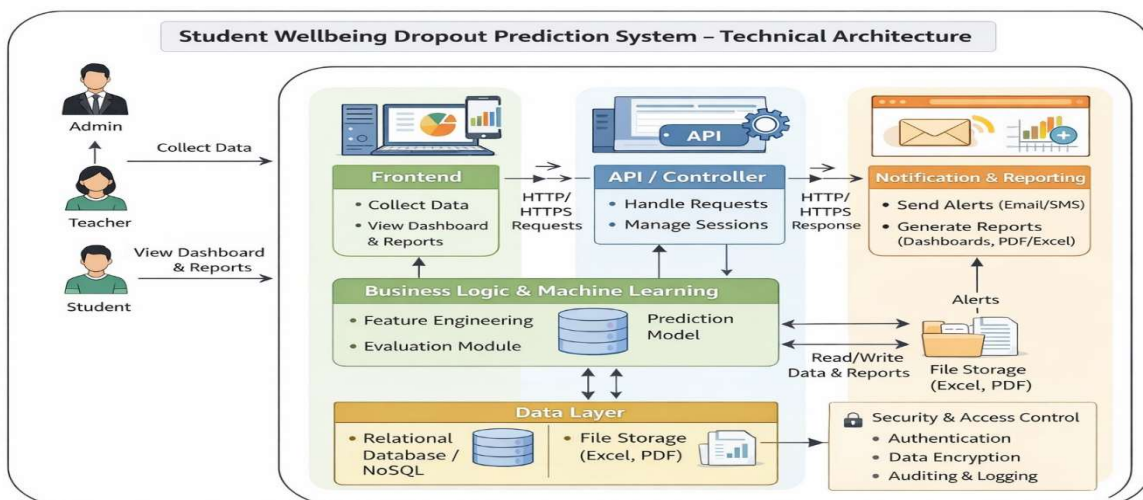


Fig. 2 Technical Architecture

The technical architecture defines the implementation environment and interaction between software components. The system is developed using a web-based technology stack, where the front-end is built with standard web technologies to provide an accessible and responsive interface. The backend is implemented using the Flask framework, which manages routing, session handling, and API communication. Data is stored in structured files and managed using data processing libraries, enabling efficient handling of student records. Machine Learning functionalities are integrated using established libraries for data preprocessing, model training, and prediction. The system also incorporates authentication mechanisms and role-based access control to ensure secure usage. At a high level, the architecture supports seamless interaction between the user interface, backend logic, and predictive model. This design enables real-time data processing, risk evaluation, and visualization, ensuring that users receive timely and meaningful insights.

Implementation

System Implementation Overview

The system is implemented as a web application using the Flask framework, integrating user management, data collection, and predictive analytics into a single platform. It supports multiple user roles, including students and teachers, each with specific permissions and functionalities. Authentication is handled through a login mechanism, where user credentials are validated and session management is used to maintain access control. Students can register and submit personal, academic, and health-related information through structured forms. Teachers are provided with access to academic data entry and monitoring features. All collected data is stored and processed to generate meaningful insights related to student performance and wellbeing.

Data Processing and Feature Engineering

The implementation includes preprocessing techniques to ensure data consistency and usability. Input values are validated and converted into appropriate numerical formats where required. Missing values are handled using statistical methods, such as mean substitution, to maintain dataset integrity. Multiple features are considered for risk prediction, including academic performance indicators

(such as CGPA and attendance), behavioral factors (such as study hours and screen time), and wellbeing attributes (such as stress level and sleep duration). These features are combined to form a comprehensive dataset used for analysis.

Risk Prediction Model

The system employs a Machine Learning approach to classify students based on their risk of dropout. A Random Forest classifier is used due to its robustness and ability to handle diverse feature types. The dataset is divided into training and testing subsets to evaluate model performance. A rule-based scoring mechanism is initially applied to generate labels, which are then used to train the model. The trained model predicts whether a student falls into a high-risk or low-risk category. These predictions are further mapped into user-friendly outputs that can be easily interpreted by educators.

Dashboard and Visualization

An interactive dashboard is implemented to present student data, risk levels, and analytical insights. The dashboard displays categorized student information, summary statistics, and risk distribution. This visualization enables teachers to quickly identify high-risk students and monitor overall trends. Additionally, the system provides functionality to generate and download reports, particularly for high-risk students. These reports support documentation and facilitate timely intervention by educators.

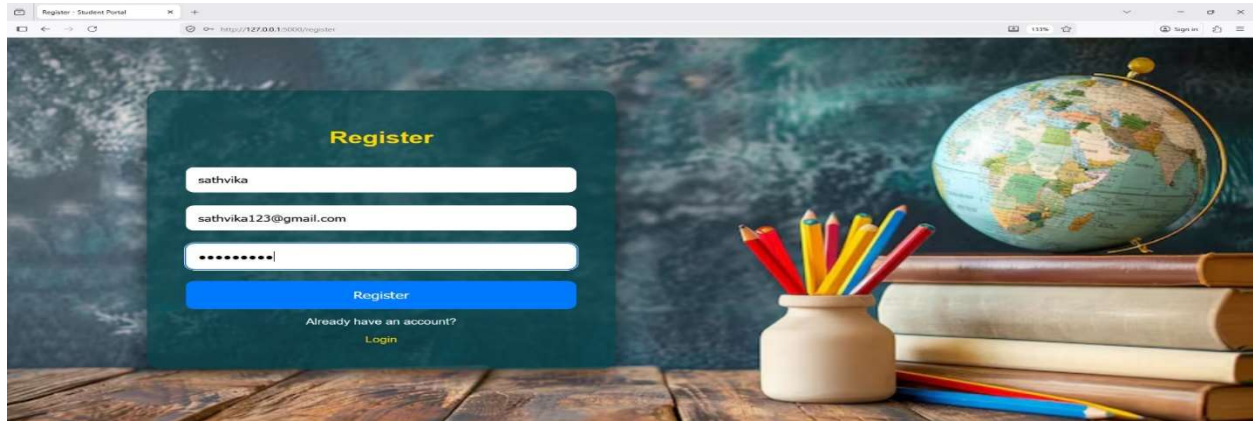
Alert and Notification System

The system includes an automated notification mechanism that identifies high-risk students and generates alerts for teachers. This feature ensures that critical cases are brought to attention without delay, allowing for early intervention. Notifications are dynamically generated based on prediction results and displayed through the system interface.

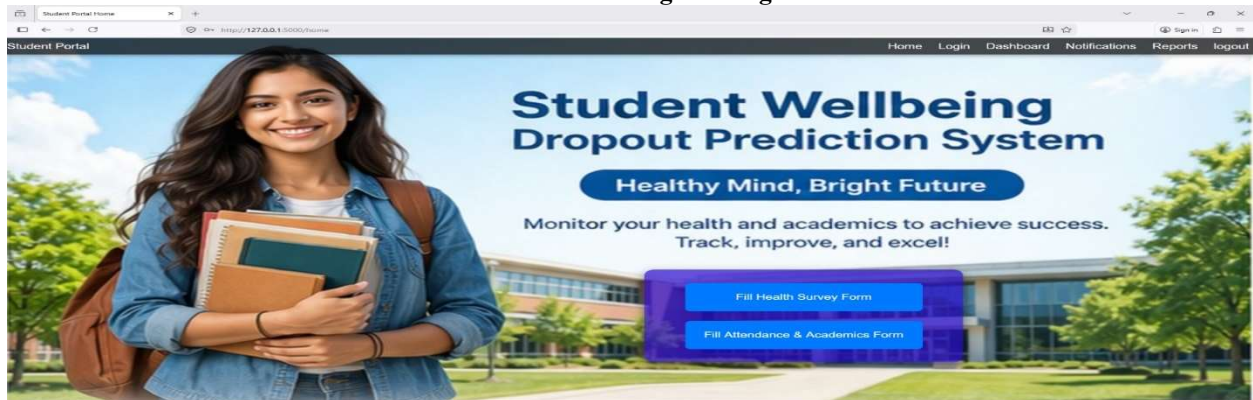
System Security and Session Management

Security is implemented through authentication, session tracking, and controlled access to system features. Password validation mechanisms ensure strong credentials, and user roles restrict access to sensitive operations. Data handling practices are designed to protect user information and maintain system integrity.

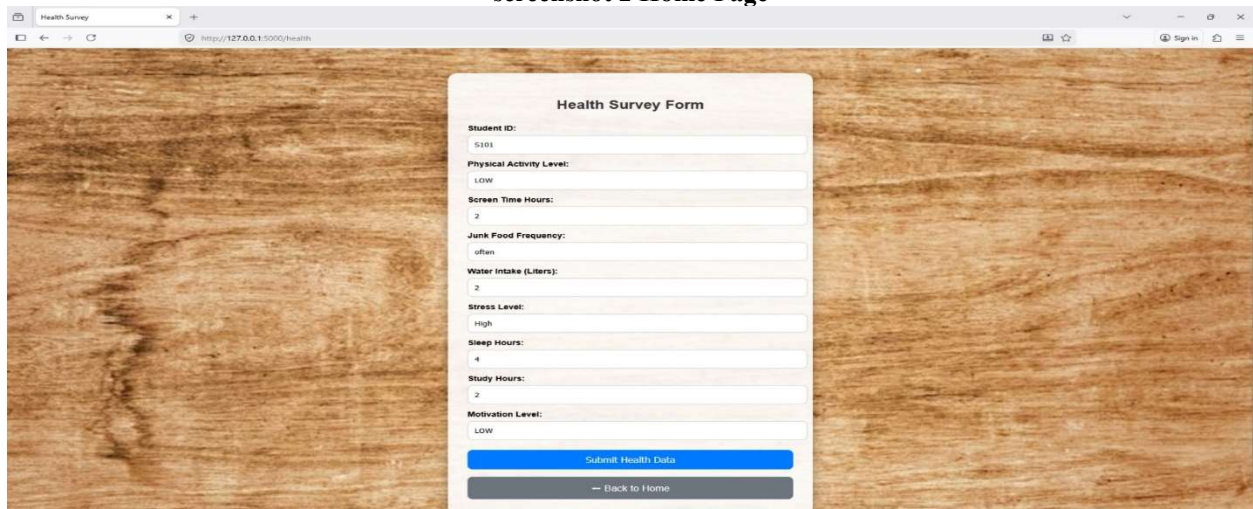
Screenshots



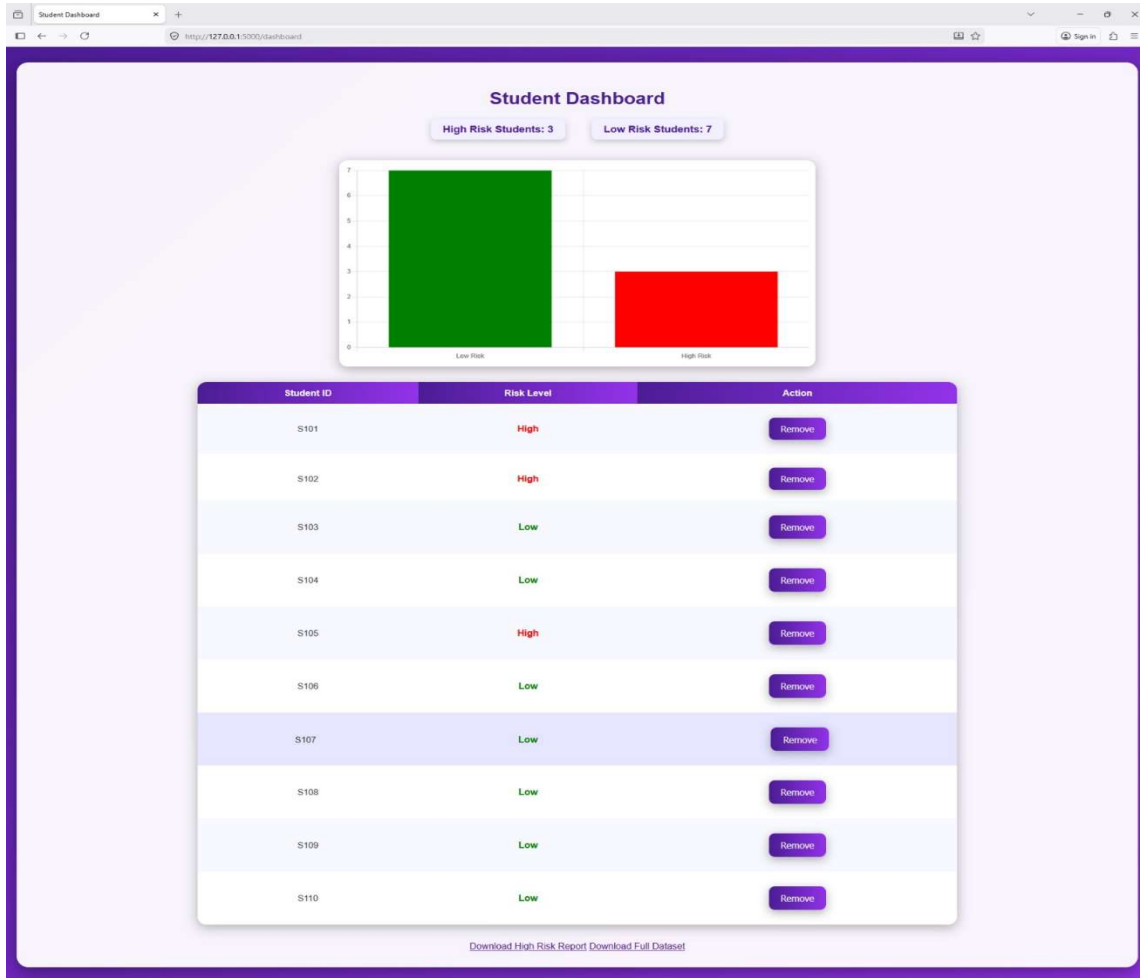
Screenshot 1 Register Page



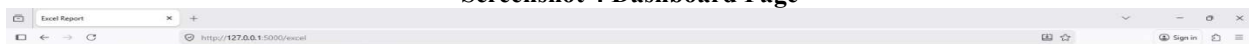
screenshot 2 Home Page



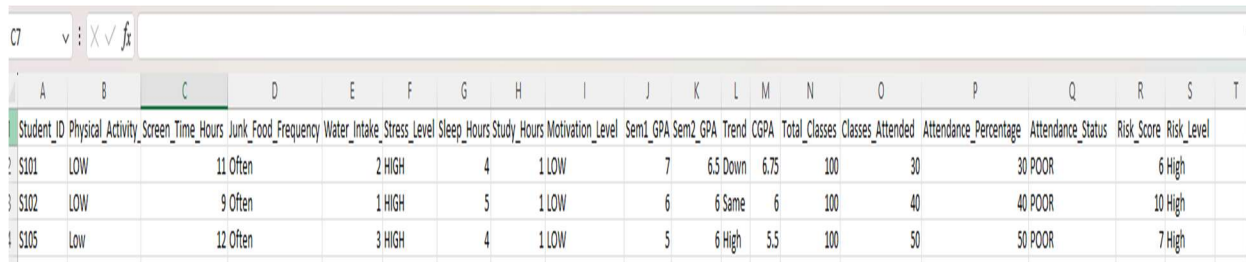
Screenshot 3 Health Survey Form Page



Screenshot 4 Dashboard Page



Screenshot 5 Reports Page



Student_ID	Physical_Activity	Screen_Time_Hours	Junk_Food_Frequency	Water_Intake	Stress_Level	Sleep_Hours	Study_Hours	Motivation_Level	Sem1_GPA	Sem2_GPA	Trend	CGPA	Total_Classes	Classes_Attended	Attendance_Percentage	Attendance_Status	Risk_Score	Risk_Level
S101	LOW	11	Often	2	HIGH	4	1	LOW	7	6.5	Down	6.75	100	30	30	POOR	6	High
S102	LOW	9	Often	1	HIGH	5	1	LOW	6	6	Same	6	100	40	40	POOR	10	High
S105	Low	12	Often	3	HIGH	4	1	LOW	5	6	High	5.5	100	50	50	POOR	7	High

Screenshot 6 Excel sheet

Testing and Validation

The system was evaluated using a set of functional test cases designed to verify the correctness, reliability, and robustness of each module. The testing process covered user management, academic data processing, health data handling, risk prediction, and reporting functionalities. In the user management module, registration and login operations were tested with both valid and invalid credentials. While successful authentication scenarios performed as expected, certain cases involving incorrect or missing credentials revealed validation gaps, indicating the need for stricter input handling and error messaging. The academic and health data modules were tested to ensure proper data submission and storage. Valid inputs were processed successfully; however, invalid data entries were not consistently rejected, highlighting the necessity for improved validation mechanisms at the input stage. The risk analysis module demonstrated correct classification behavior for both high-risk and low-risk scenarios based on predefined conditions and model predictions. Nevertheless, the system showed limitations in handling incomplete or null data, occasionally producing incorrect classifications instead of raising appropriate errors. In the analytics and reporting module, dashboard visualization, report generation, and notification features were tested. Most functionalities performed correctly, including real-time dashboard display and report downloads. However, inconsistencies in data validation and edge-case handling suggest areas for further refinement. Overall, the testing phase confirms that the system meets core functional requirements, while also identifying opportunities to enhance reliability and data validation.

Conclusion

The Student Wellbeing Dropout Prediction System presents an integrated approach to addressing the persistent issue of student attrition in educational institutions. By combining academic performance metrics, attendance records, and wellbeing-related factors within a single platform, the system enables a more comprehensive understanding of student

conditions. This holistic approach improves the ability to identify at-risk students at an early stage and supports timely intervention strategies. The platform incorporates role-based access for students and educators, ensuring secure and structured interaction with the system. Students can contribute personal and health-related data, while teachers can monitor academic progress, evaluate risk levels, and respond to alerts. The inclusion of an interactive dashboard facilitates clear visualization of trends and predictions, enabling informed decision-making. From a technical perspective, the system demonstrates effective integration of frontend and backend components, supported by Machine Learning techniques for predictive analysis. Key system qualities such as usability, performance, and reliability have been considered to ensure consistent operation. Features such as automated notifications and report generation further enhance the system's practical utility.

Future Scope

The proposed system offers significant potential for future enhancements aimed at improving intelligence, scalability, and user engagement. One possible direction is the integration of advanced Machine Learning and Artificial Intelligence techniques to provide personalized recommendations for at-risk students. These may include adaptive learning strategies, behavioral insights, and early counseling suggestions. The development of mobile-based applications could further increase accessibility, allowing students to update wellbeing data and receive feedback in real time. Integration with real-time communication systems, such as instant alerts to teachers and parents, would strengthen the intervention process. Additionally, enhanced data visualization and automated reporting tools could provide deeper insights into student trends and institutional performance. Future improvements may also include support for multiple institutions, multilingual interfaces, and cloud-based deployment for scalability. Incorporating mental health assessment tools and personalized support mechanisms could further extend the system's impact. With these advancements, the platform has the potential to evolve

into a comprehensive solution for student success, wellbeing management, and predictive educational analytics.

References

- [1]P. Goldstein and R. Katz, *Student Retention and Engagement Strategies*. Routledge, 2018.
- [2] V. Tinto, *Leaving College: Rethinking the Causes and Cures of Student Attrition*. University of Chicago Press, 1993.
- [3]G. D. Kuh et al., *Student Success in College: Creating Conditions that Matter*. Jossey-Bass, 2010.
- [4]J. Gray and C. McGuinness, *Data-Driven Decision Making in Education*. Springer, 2019.
- [5]A. Bowers and R. Sprott, "Predictive Analytics

- for Student Success," *EDUCAUSE Review*, 2012.
- [6]IBM, "SPSS Modeler," Available: <https://www.ibm.com/products/spss-modeler>
- [7]Microsoft, "Power BI," Available: <https://powerbi.microsoft.com>
- [8]Python Software Foundation, "Python Programming Language," Available: <https://www.python.org>
- [9]Pandas Documentation, "Python Data Analysis Library," Available: <https://pandas.pydata.org>
- [10]Scikit-learn, "Machine Learning in Python," Available: <https://scikit-learn.org>
- [11]Flask, "Flask Web Framework," Available: <https://flask.palletsprojects.com>
- [12] Jupyter Project, "Jupyter Notebooks," Available: <https://jupyter.org>