As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 12, Issue 07, pages 39-48.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue06.html
DOI: https://doi.org/10.5281/zenodo.20739591

Predictive Analysis of Student Academic Performance Using AI

U. Jenny Grace, Gummalla Aditya Rama Venkat, Kamisetti Siva Pavan Kumar, Sunkara Madhuri

Department of Information Technology , Swarnandhra College of Engineering & Technology, Seetharamapuram, Narsapur, Andhra Pradesh, INDIA.

To Cite this Article
U. Jenny Grace, Gummalla Aditya Rama Venkat, Kamisetti Siva Pavan Kumar & Sunkara Madhuri (2026). Predictive Analysis of Student Academic Performance Using AI. International Journal for Modern Trends in Science and Technology, 12(07), 39-48. https://doi.org/10.5281/zenodo.20739591

Article Info
Received: 06 June 2026; Revised: 29 June 2026; Accepted: 01 July 2026.

KEYWORDS

Student Performance Prediction, Artificial Intelligence, Machine Learning, Decision Tree (J48), WEKA, Academic Data Analysis, Spring Boot

ABSTRACT

The rapid growth of educational data has created opportunities for improving student performance through intelligent analysis and prediction systems. Traditional methods of evaluating student performance are often manual, time-consuming, and lack the ability to identify hidden patterns in academic data. This creates a need for automated systems that can assist educators in making data-driven decisions. This project presents an AI-based Student Academic Prediction System that predicts the performance level of students using key academic indicators such as attendance, assignment scores, and mid-term marks. The system utilizes the Decision Tree (J48) algorithm implemented in WEKA, which is effective for classification tasks and provides interpretable decision rules. Based on the input parameters, the model classifies students into three categories: At Risk, Good, and Excellent. The system is developed using Spring Boot (Java) for backend processing and MySQL as the database for storing student data. The integration of machine learning with a web-based application enables real-time prediction and efficient data management. The model is trained on historical student data and uses predefined thresholds and learned patterns to generate accurate predictions. Experimental results demonstrate that the system can effectively identify students who are at risk, enabling early intervention and academic support. The proposed system serves as a decision-support tool for educators, helping improve student performance and overall educational outcomes.

1. INTRODUCTION

In recent years, the education sector has witnessed significant growth in the amount of data generated through various academic activities such as attendance tracking, assignment evaluations, and internal assessments. Analyzing this data effectively to improve student performance has become a major challenge for educational institutions. Traditional methods of evaluating student performance are often manual, time-consuming, and lack the ability to identify hidden patterns in academic data.

Academic performance is influenced by multiple factors including attendance, assignment completion, and internal examination scores. Identifying students who are at risk of poor performance at an early stage is crucial for providing timely support and improving overall educational outcomes. However, conventional approaches rely heavily on subjective judgment and are not efficient when dealing with large volumes of student data.

With the advancement of Artificial Intelligence (AI) and Machine Learning (ML), there has been a shift towards data-driven approaches in the education system. Machine learning techniques enable the analysis of large datasets to uncover patterns and relationships that are not easily identifiable through manual methods. These techniques can be used to predict student performance and assist educators in making informed decisions.

This project presents an AI-based Student Academic Prediction System that aims to analyse student data and predict their academic condition. The system considers key academic parameters such as attendance, assignment scores, and mid-term marks to evaluate student performance. Based on these inputs, the system classifies students into three categories: At Risk, Good, and Excellent.

The core of the system is built using the Decision Tree (J48) algorithm implemented in WEKA, which is widely used for classification tasks. The algorithm generates decision rules based on input data, making it easy to interpret and understand. This is particularly useful in educational environments where transparency and clarity are important for decision-making.

The implementation of the system is carried out using Spring Boot (Java) for backend development, which provides a robust and scalable framework for building web applications. The student data is stored and managed using the MySQL database, ensuring efficient data handling and retrieval. The integration of machine learning with a web-based application allows real-time prediction and user interaction.

The proposed system helps educators identify students who require additional attention and support. By predicting student performance in advance, institutions can take proactive measures such as providing extra classes, mentoring, and academic guidance. This not only improves individual student outcomes but also enhances the overall performance of the institution.

In conclusion, the integration of machine learning techniques into academic systems provides an effective solution for student performance prediction. The proposed system demonstrates how AI can be used as a decision-support tool in education, enabling efficient analysis, accurate prediction, and improved academic success.

1.1 OVERVIEW OF STUDENT PERFORMANCE ANALYSIS

Student performance analysis is an essential component of modern educational systems, as it helps institutions evaluate the academic progress and learning outcomes of students. With the increasing number of students and academic data, it has become difficult for educators to manually analyse and track individual performance effectively.

Traditionally, student evaluation is based on periodic examinations, assignments, and attendance records. However, these methods often provide only a limited understanding of a student's overall academic condition. Manual analysis may overlook hidden patterns and relationships between different performance factors, leading to delayed identification of students who require academic support.

In recent years, data-driven approaches have been introduced to enhance the accuracy and efficiency of student performance evaluation. By utilizing historical academic data such as attendance, assignment scores, and internal marks, institutions can gain deeper insights into student behaviour and learning patterns. These insights help in identifying trends and predicting future performance.

In this project, student performance analysis is carried out using key academic parameters including attendance, assignments, and mid-term marks. These parameters are considered reliable indicators of student engagement and academic consistency. By analysing these inputs, the system is able to assess the overall academic condition of students.

The integration of analytical techniques into education not only improves the evaluation process but also enables institutions to adopt a proactive approach in supporting students. This leads to better academic outcomes and improved student success rates.

1.2 NEED FOR AI IN EDUCATION SYSTEM

The rapid advancement of technology and the availability of large volumes of educational data have created a need for intelligent systems that can analyze and interpret data efficiently. Artificial Intelligence (AI) plays a significant role in transforming traditional education systems into smart and data-driven environments.

In conventional systems, educators rely on manual methods to assess student performance, which can be time-consuming and prone to human error. These methods are not scalable and may fail to identify students who are at risk of poor academic performance at an early stage.

AI-based systems provide the capability to process large datasets and extract meaningful insights automatically. These systems can identify patterns, trends, and relationships in student data, enabling accurate prediction of academic outcomes. By leveraging AI, institutions can move from reactive approaches to proactive strategies.

In the context of this project, AI is used to analyse student data and predict their academic condition. The system utilizes machine learning techniques to classify students into categories such as At Risk, Good, and Excellent. This helps educators take timely actions to support students who require additional attention.

The adoption of AI in education improves decision-making, reduces manual effort, and enhances the overall efficiency of academic management systems. It also supports personalized learning and targeted interventions, leading to better student performance.

1.3 ROLE OF MACHINE LEARNING IN PREDICTION

Machine Learning (ML), a subset of Artificial Intelligence, plays a crucial role in predicting student performance by learning patterns from historical data. ML algorithms are designed to analyse input data, identify relationships between variables, and make predictions based on learned patterns.

In educational data mining, machine learning is widely used for classification and prediction tasks. By training models on past student data, it is possible to predict future academic outcomes with a high degree of accuracy. These predictions help educators identify students who are likely to perform well or face difficulties.

In this project, the machine learning model is developed using the Decision Tree (J48) algorithm in WEKA. The model is trained using input features such as attendance, assignment scores, and mid-term marks. Based on these inputs, the model generates decision rules that classify students into predefined categories.

The use of machine learning provides several advantages:

- Automated analysis of large datasets
- Improved prediction accuracy
- Ability to identify hidden patterns
- Support for data-driven decision making

Unlike traditional methods, machine learning models continuously improve as more data becomes available. This makes them highly effective for real-world applications such as student performance prediction.

1.4 PREDICTIVE ANALYTICS FOR ACADEMIC IMPROVEMENT

Predictive analytics is the process of using historical data, statistical techniques, and machine learning algorithms to forecast future outcomes. In the education sector, predictive analytics is used to analyse student data and predict academic performance.

The primary objective of predictive analytics in education is to identify students who are at risk and provide timely interventions. By analysing factors such as attendance, assignments, and internal marks, institutions can predict the likelihood of a student's success or failure.

In this project, predictive analytics is applied to classify students into three categories:

- At Risk
- Good
- Excellent

These classifications help educators understand the academic condition of students and take appropriate actions. For example, students identified as "At Risk" can be provided with additional support such as extra classes, mentoring, and guidance.

The use of predictive analytics improves the efficiency of academic systems by enabling early detection of performance issues. It also helps institutions allocate resources effectively and enhance overall educational outcomes.

Furthermore, predictive systems promote a proactive approach to education, where decisions are based on data rather than assumptions. This leads to improved student engagement, better academic performance, and increased success rates.

2 LITERATURE SURVEY

The application of machine learning techniques in the education sector has significantly improved the way student performance is analysed and predicted. Educational institutions generate large amounts of academic data such as attendance records, assignment scores, and internal examination marks. Analysing this data manually is inefficient and often fails to reveal meaningful insights.

Recent research focuses on developing intelligent systems that utilize machine learning algorithms to predict student performance. These systems help educators identify students who are at risk and enable timely intervention to improve academic outcomes. This chapter presents a review of existing approaches related to student performance prediction and machine learning techniques used in education.

2.1 STUDENT PERFORMANCE PREDICTION MODELS

Student performance prediction models aim to analyze academic data and forecast future outcomes. Early models were based on statistical methods such as regression analysis, which established relationships between variables like marks and attendance. However, these methods were limited in handling complex and nonlinear data.

Modern prediction models use machine learning techniques to improve accuracy and efficiency. These models are trained using historical student data and can identify patterns that are not visible through manual analysis.

Commonly used input parameters in these models include:

- Attendance percentage
- Assignment completion and scores
- Internal examination marks

These parameters are strong indicators of student engagement and academic performance. By analysing these features, prediction systems can classify students into different categories such as low-performing, average, and high-performing.

In this project, a similar approach is followed where student data is used to classify academic condition into At Risk, Good, and Excellent, making the system practical and relevant for real-world educational environments.

2.2 MACHINE LEARNING IN EDUCATION

Machine learning has become a key technology in educational data mining, enabling automated analysis and prediction of student performance. It allows systems to learn from past data and make accurate predictions for new data.

In the education domain, machine learning is used for:

- Predicting student performance
- Identifying at-risk students
- Enhancing personalized learning
- Supporting academic decision-making

Classification algorithms are widely used in this context, as they categorize students based on their academic condition. These algorithms analyze input features and assign them to predefined classes.

Compared to traditional methods, machine learning provides:

- Higher accuracy
- Faster data processing
- Reduced manual effort
- Ability to handle large datasets

In this project, machine learning is used to build a predictive model that classifies students based on their academic indicators, making the system efficient and reliable.

2.3 DECISION TREE (J48) ALGORITHM

The Decision Tree algorithm is one of the most widely used machine learning techniques for classification problems. It creates a tree-like structure where each node represents a decision based on input features, and each branch represents the outcome of that decision.

In this project, the **J48 algorithm in WEKA** is used to build the prediction model. J48 is an extension of the ID3 algorithm and is based on concepts such as entropy and information gain.

The algorithm works as follows:

- It selects the most important attribute that best splits the data

- It creates decision nodes based on input features
- It continues splitting until classification is achieved

In the context of student prediction:

- Attendance, assignment scores, and mid marks are used as input features
- The output is classified into:
    - At Risk
    - Good
    - Excellent

The model generates decision rules such as:

- Low attendance and low marks → At Risk
- Moderate performance → Good
- High performance → Excellent

Advantages of J48

- Easy to understand and interpret
- Suitable for classification tasks
- Works well with both numerical and categorical data
- Provides clear decision rules

Limitations

- Can overfit if not properly controlled
- Depends on quality of input data

Despite these limitations, J48 is highly suitable for this project due to its simplicity, efficiency, and interpretability.

2.4 ANALYSIS OF EXISTING STUDIES

From the review of existing studies, it is observed that machine learning techniques play a crucial role in predicting student performance. Various models have been developed using different algorithms, but they differ in complexity, accuracy, and interpretability.

Many existing systems use complex algorithms such as Random Forest and Neural Networks, which provide high accuracy but lack interpretability.

These models act as black-box systems, making it difficult for educators to understand the reasoning behind predictions.

On the other hand, simpler models like Decision Trees provide clear and interpretable results, making them more suitable for educational environments. They allow educators to understand how different factors influence student performance.

It is also observed that most effective prediction systems use relevant academic parameters such as attendance, assignments, and internal marks. These features provide meaningful insights into student behaviour and academic progress.

Based on these observations, the proposed system adopts the Decision Tree (J48) algorithm, which provides a balance between accuracy and interpretability. The system focuses on practical input features and generates clear classification results, making it suitable for real-world implementation.

3 PROPOSED METHODOLOGY

The proposed method is an AI-based student academic prediction system that uses student data such as **attendance, assignment marks, and mid-term marks** to predict performance levels.

The system applies the **Decision Tree J48 algorithm in WEKA** to classify students into **At Risk, Good, and Excellent** categories. It is built using **Spring Boot Java** for backend processing and **MySQL** for data storage.

Working steps

1. Student or faculty enters academic details.
2. The system validates and preprocesses the data.
3. The J48 model analyzes the records.
4. The prediction result is generated.
5. The result is stored in the database and shown in the dashboard.

3.1 SYSTEM ARCHITECTURE

The system uses a **4-layer architecture**: User Interface, Application Layer, Machine Learning Layer, and Database Layer.

Architecture flow

- **User Interface**: Accepts student details and shows prediction results.
- **Application Layer**: Spring Boot processes requests and controls the workflow.
- **Machine Learning Layer**: WEKA J48 predicts student performance.
- **Database Layer**: MySQL stores student records and prediction history.

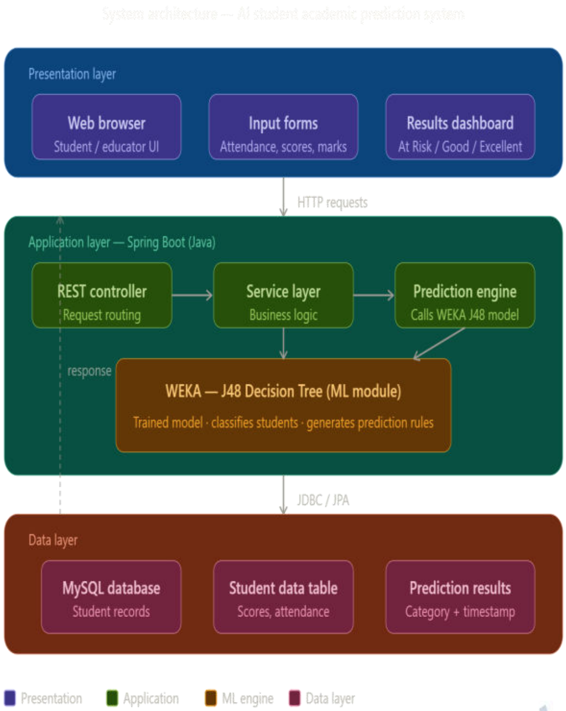

Fig 3.1 Proposed System Architecture

The system architecture of the Student Academic Prediction System follows a layered approach for better organization and performance. The first layer is the **presentation layer**, where students and educators interact with the system through a web browser. Users enter academic details such as attendance, assignment scores, and mid-term marks using input forms. The interface collects and forwards this data to the backend for processing.

The second layer is the **application layer**, developed using Spring Boot. It contains the REST controller, service layer, and prediction engine. The controller handles user requests, the service layer processes the data, and the prediction engine uses the WEKA J48 Decision Tree model to classify students as At Risk, Good, or Excellent.

The final layer is the **data layer**, which uses MySQL to store student records and prediction results. The backend communicates with the database to save and retrieve data efficiently. This layered structure ensures smooth data flow, accurate predictions, and effective data management.

3.2 USE CASE DIAGRAM

The main actors are **Student**, **Faculty**, and **Admin**.

- **Use cases**
- **Student**
- Login
- View prediction
- Check performance category
- **Faculty**
- Login
- Enter student data
- Upload bulk data
- Generate predictions
- View reports
- **Admin**
- Manage users
- Manage student records
- View logs
- Monitor system

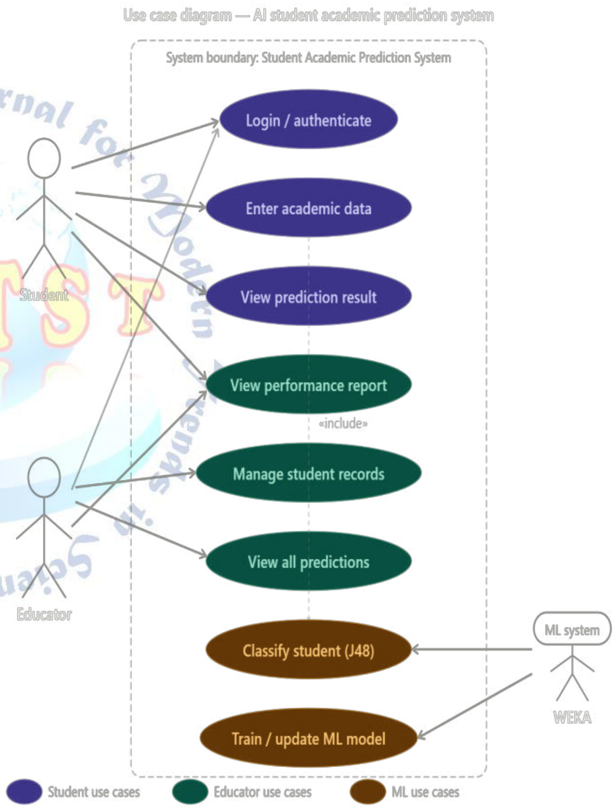

Fig 3.2 Use Case Diagram

The use case diagram illustrates how different users interact with the Student Academic Prediction System. The primary actors in the system are the Student and the Educator. The Student can log in to the system, enter academic details such as attendance and marks, view the predicted performance category, and access a performance report. These actions allow students to monitor their academic standing and understand areas that need improvement.

The Educator actor has additional privileges within the system. Educators can manage student records, review submitted academic data, view all prediction results, and generate performance reports. This helps them track overall class performance and identify students who require early academic support or intervention.

The ML system represents the internal machine learning component responsible for prediction tasks. It classifies students using the J48 Decision Tree algorithm and supports training or updating the model with historical data. The use case diagram clearly defines user roles and system functionalities, showing how each actor interacts with the system to achieve specific goals.

3.3 CLASS DIAGRAM

The system consists of several main classes, each with a specific role. The **Student** class stores student information such as studentId, name, attendance, assignmentMarks, and midMarks. It includes methods like getDetails() and updateDetails() to retrieve and modify academic data.

The **DataProcessor** class is responsible for validating and preprocessing the input data using methods such as validateData() and preprocessData(). The **PredictionModel** class handles the machine learning tasks, where buildModel() is used to train the model and predictPerformance() is used to classify students as At Risk, Good, or Excellent.

The **DatabaseManager** class manages data storage with methods like saveData(), getData(), and updateData(). The **Faculty** class can enter student data and generate predictions, while the **Admin** class manages users and views system logs. Together, these classes ensure organized data handling and accurate performance prediction.

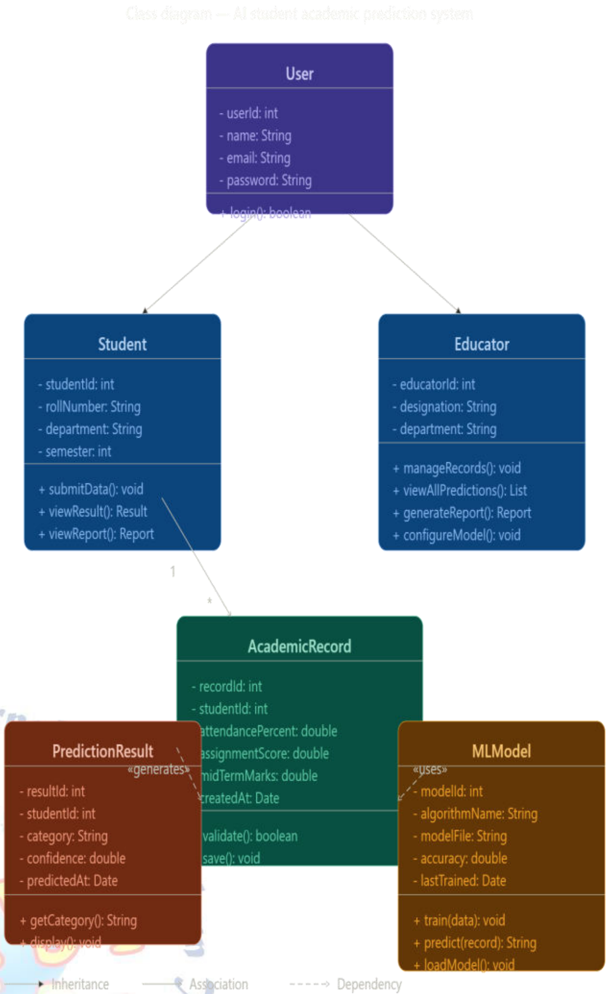

Fig 3.3 Sequence Diagram

The class diagram represents the structural design of the Student Academic Prediction System and shows the relationships between different classes. The User class acts as the parent class, providing common attributes such as userId, name, email, and login credentials. The Student and Educator classes inherit from the User class, demonstrating the concept of inheritance. This allows both classes to share common properties while having their own specific functionalities.

The Student class stores academic and personal details such as roll number, department, and semester. It includes methods like submitData(), viewResult(), and viewReport(), which allow students to enter their academic information and access prediction outcomes. The Educator class has additional responsibilities, including managing student records, viewing all prediction results, generating performance reports, and configuring or updating the prediction model settings.

The AcademicRecord class is responsible for storing academic indicators such as attendance percentage, assignment scores, and mid-term marks. These records are used as input for the prediction process. The MLModel class handles machine learning operations, including training the model, loading the trained model, and predicting student performance. After prediction, the PredictionResult class stores the final performance category (At Risk, Good, or Excellent), along with confidence level and prediction date. Together, these classes form a well-structured and organized system design.

4 RESULTS

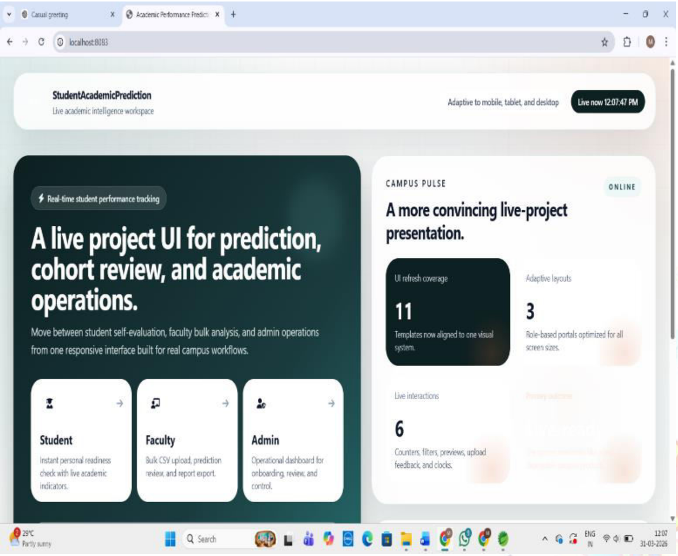

Fig 4.1: Academic Performance Prediction Home Page

This figure illustrates the main landing page of the Academic Performance Prediction system. It provides an overview of the platform, including different user roles such as Student, Faculty, and Admin. The interface emphasizes real-time performance tracking, adaptive design, and system capabilities. It acts as a navigation hub for accessing various modules and showcases the project as a live, interactive application.

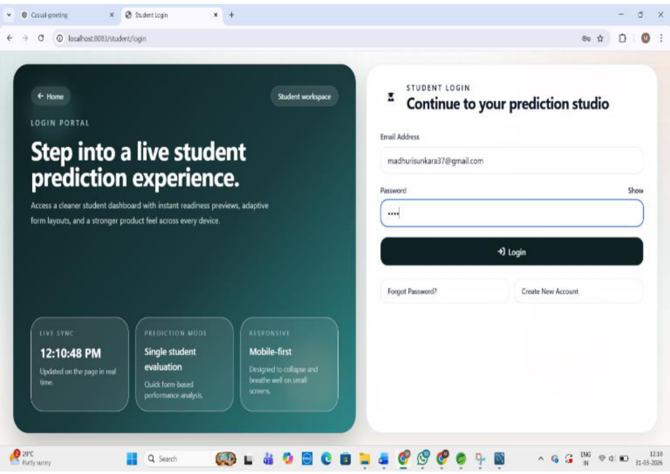

Fig 4.2: Student Login Interface

This figure shows the student login portal of the system. The interface allows students to securely access their prediction workspace by entering their email and password. The design includes a responsive layout with a clear call-to-action login button, along with options for password recovery and new account registration. The left panel highlights key features such as real-time synchronization, single-student evaluation, and mobile responsiveness, providing an overview of the system capabilities.

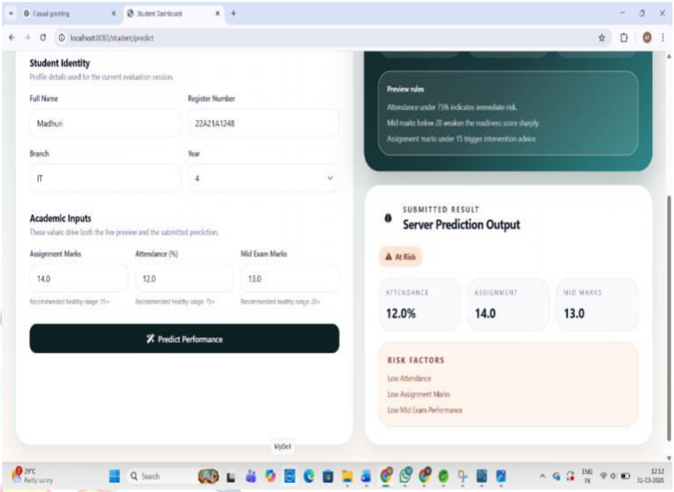

Fig 4.3: Student Prediction Input and Output Dashboard

This figure represents the student prediction dashboard where users input academic details such as assignment marks, attendance percentage, and mid-exam scores. After submission, the system generates a prediction result indicating the student's academic risk level. The output section displays performance metrics and highlights risk factors such as low attendance or poor marks, helping students understand areas needing improvement.

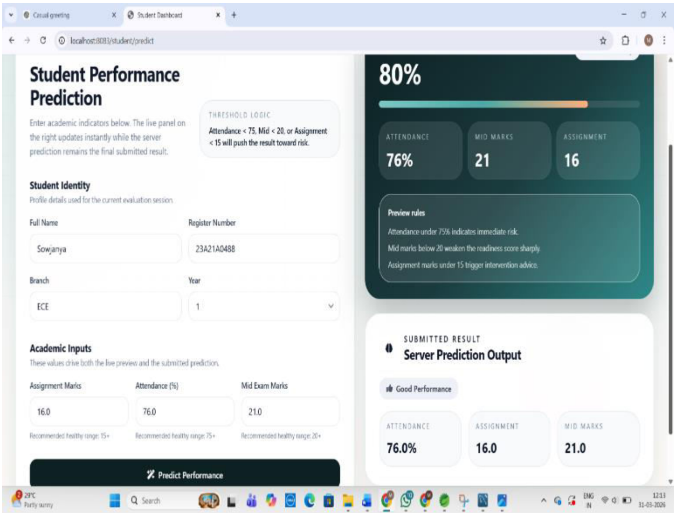

Fig 4.4: Student Prediction Form with Moderate Performance Input

This figure illustrates the student prediction interface populated with sample academic data representing moderate performance. The system displays a live readiness score of 80%, based on inputs such as attendance (76%), mid marks (21), and assignment marks (16). The threshold logic and preview rules highlight how values close to minimum requirements influence the prediction outcome, providing immediate feedback before final submission.

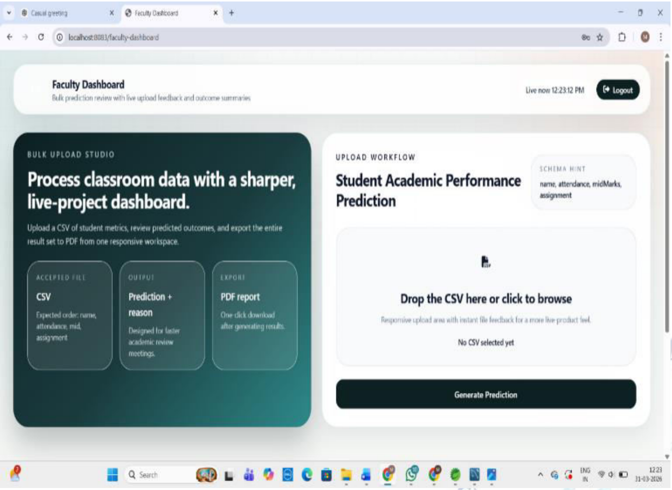

Fig 4.5: Faculty Dashboard for Bulk Student Performance Prediction

This figure shows the faculty dashboard designed for bulk processing of student data. It enables instructors to upload a CSV file containing multiple student records and generate predictions simultaneously. The interface includes features such as file upload, schema guidance, prediction output with reasoning, and an option to export results as a PDF report. This module enhances scalability and supports academic performance analysis at the classroom level.

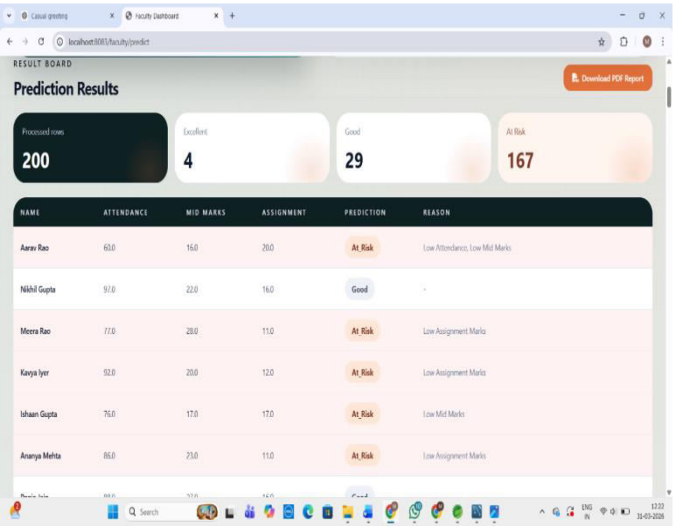

Fig 4.6: Prediction Results Dashboard

This figure presents the Prediction Results page, displaying the outcomes of the student performance analysis. It includes summary metrics such as total processed records, number of excellent, good, and at-risk students. A detailed table lists each student's attendance, marks, predicted category, and reason for the prediction. The interface also provides an option to download the results as a PDF report for further analysis and documentation.

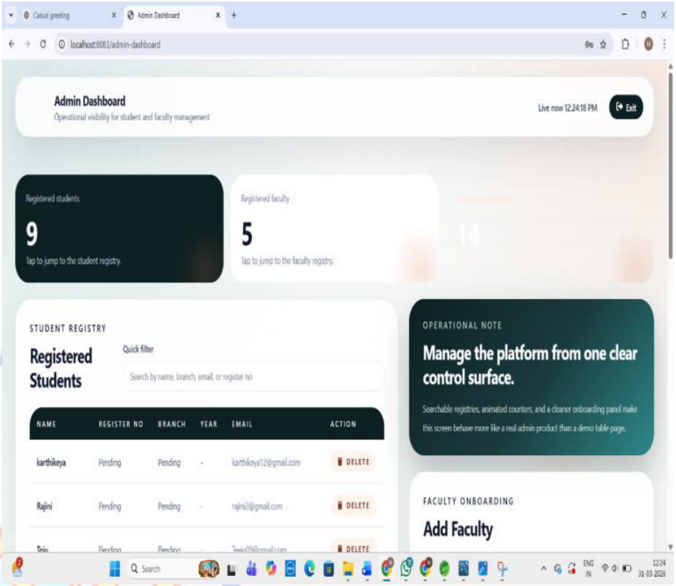

Fig 4.7: Admin Dashboard – Student and Faculty Management

This figure presents the Prediction Results page, displaying the outcomes of the student performance analysis. It includes summary metrics such as total processed records, number of excellent, good, and at-risk students.

A detailed table lists each student's attendance, marks, predicted category, and reason for the prediction. The interface also provides an option to download the results as a PDF report for further analysis and documentation.

5 CONCLUSION

The Student Academic Prediction System developed in this project successfully demonstrates the application of machine learning techniques in the education domain. The system analyzes key academic parameters such as attendance, assignment marks, and mid-term marks to predict the performance of students.

The implementation of the **Decision Tree (J48) algorithm using WEKA** provides accurate and interpretable results. The system classifies students into categories such as *At Risk, Good, and Excellent*, which helps educators understand the academic condition of students effectively.

The integration of **Spring Boot, MySQL, and WEKA** ensures smooth functioning of the system. The web-based interface allows users to easily input data and obtain prediction results in real time.

The proposed system overcomes the limitations of traditional manual evaluation methods by providing:

- Automated prediction of student performance
- Early identification of at-risk students
- Improved accuracy and efficiency
- Support for data-driven decision making

Overall, the system proves to be an effective tool for academic performance analysis and can assist institutions in improving student outcomes.

6 FUTURE ENHANCEMENT

Although the system performs effectively, there are several enhancements that can be implemented to improve its capabilities in the future.

- **Incorporating More Parameters:**

Additional factors such as extracurricular activities, behaviour, and previous academic history can be included for better prediction accuracy.

- **Using Advanced Algorithms:**

More advanced machine learning algorithms like Random Forest or Neural Networks can be implemented to improve performance.

- **Real-Time Data Integration:**

The system can be connected to live academic databases for real-time data analysis.

- **User Role Management:**

Different user roles such as admin, teacher, and student can be added for better access control.

- **Graphical Representation:**

Visualization of results using graphs and charts can improve understanding.

- **Mobile Application Development:**

A mobile version of the system can be developed for easier accessibility.

These enhancements can make the system more powerful, flexible, and suitable for large-scale implementation in educational institutions.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Han, J., Kamber, M., & Pei, J., Data Mining: Concepts and Techniques, 3rd Edition, Morgan Kaufmann, 2011.

[2] Witten, I. H., Frank, E., & Hall, M. A., Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2016.

[3] Quinlan, J. R., "C4.5: Programs for Machine Learning," Morgan Kaufmann Publishers, 1993. (Basis for J48 Decision Tree Algorithm)

[4] Kotsiantis, S. B., "Use of Machine Learning Techniques for Educational Data Mining: A Survey," International Journal of Computer Science and Applications, 2012.

[5] Romero, C., & Ventura, S., "Educational Data Mining: A Review of the State of the Art," IEEE Transactions on Systems, Man, and Cybernetics, 2010.

[6] WEKA Documentation, University of WaikatoAvailable: https://www.cs.waikato.ac.nz/ml/weka/

[7] Oracle Corporation, Java DocumentationAvailable: https://docs.oracle.com/javase/

[8] Spring Boot DocumentationAvailable: https://spring.io/projects/spring-boot

[9] MySQL DocumentationAvailable: https://dev.mysql.com/doc/

[10] Aggarwal, C. C., Machine Learning for Text, Springer, 2018.