Predicting Student Performance Using Educational Data

[1]Mr. T. Sesha Sai,[2]Gavini Hari Vikas,[3]Charan Kakumanu,[4]Gollapudi Kesava Naga Gupta

[1]Assistant Professor, Dept of Computer Science and Engineering, St. Ann's College of Engineering and Technology, Chirala-523187, India.

[2,3,4]B. Tech Student, Dept of Computer Science and Engineering, St. Ann's College of Engineering and Technology, Chirala-523187, India.

ABSTRACT

Predicting student academic performance is an important task for educational institutions to enhance learning outcomes and provide timely support to students. Traditional evaluation methods mainly rely on examinations and manual observation, which often fail to identify students at risk in the early stages. This paper presents a machine learning-based system that analyzes educational data such as attendance, study hours, previous exam marks, and assignment scores to predict student performance. The system applies data preprocessing and classification techniques to generate accurate predictions. A web-based interface is developed using Flask to allow users to input student data and obtain results in real time. The proposed model helps educators identify students who may require additional academic support. By utilizing data-driven insights, the system improves decision-making and enhances academic planning. Overall, the approach contributes to efficient performance monitoring and supports the development of intelligent educational systems.

KEYWORDS: ***Student Performance Prediction, Machine Learning, Educational Data Mining, Flask Framework, Data Analysis, Academic Analytics.***

INTRODUCTION

Education plays a vital role in the development of individuals and society, yet many institutions face challenges in identifying students who may struggle academically. Traditional methods of evaluating student performance rely on examinations, assignments, and teacher observations, which often provide results only after assessments are completed. This delay makes it difficult to take early corrective actions for students who require additional academic support. As a result, many students fail to achieve their full potential due to the lack of timely intervention. With the rapid growth of data science and machine learning, educational institutions can now analyze large volumes

of student- related data to improve learning outcomes. Educational Data Mining (EDM) enables the extraction of meaningful patterns from data such as attendance, study habits, previous academic performance, and participation. These insights help educators understand student behavior and predict future performance more effectively. The proposed system aims to develop a machine learning-based model that predicts student performance using educational data. A web-based interface is implemented using Flask, allowing users to input student details and receive predictions in real time. The system supports efficient data processing and accurate prediction generation. By providing early identification of at-risk students, the system helps educators take timely actions and improve academic outcomes. This approach enhances decision-making and promotes data-driven education.

LITERATURE SURVEY

Student performance prediction has become an important research area in educational data mining, where machine learning techniques are used to analyze academic data and improve learning outcomes. Various studies highlight that traditional evaluation methods are limited in identifying at-risk students at an early stage. As a result, researchers have explored the use of predictive models to analyze student behavior and academic patterns. Several approaches have been proposed using machine learning algorithms such as decision trees, logistic regression, support vector machines, and random forest models. These techniques help in analyzing large datasets and extracting meaningful insights related to student performance. Educational data such as attendance, study hours, and previous academic records are commonly used to train prediction models. Recent research also emphasizes the importance of real-time data analysis and user-friendly systems that assist educators in decision-making. Web-based platforms and learning management systems have improved accessibility and usability. Additionally, studies focus on data security and privacy while handling student information. Overall, the literature supports the use of machine learning for predicting student performance and highlights the need for efficient, accurate, and accessible prediction systems in education.

RELATED WORK

Recent studies in student performance prediction have focused on the application of machine learning techniques to analyze educational data and improve academic outcomes. Various models such as decision trees, random forest, and support vector

machines have been widely used to predict student success based on academic and behavioral factors. These approaches demonstrate that predictive analytics can effectively identify students at risk of poor performance. Research also highlights the importance of feature selection, where attributes such as attendance, study hours, previous grades, and participation significantly influence prediction accuracy. Data preprocessing techniques, including normalization and handling missing values, are essential for improving model performance. Additionally, ensemble methods have been explored to enhance prediction reliability by combining multiple algorithms. Web-based applications integrated with machine learning models have improved system accessibility and usability for educators. However, many existing systems lack real-time prediction capabilities and user-friendly interfaces. These studies emphasize the need for efficient, scalable, and interpretable models to support academic decision-making and improve student performance.

EXISTING SYSTEM

In the existing system, student performance is evaluated using traditional methods such as examinations, assignments, and teacher observations. These approaches rely heavily on manual analysis and provide results only after assessments are completed, which limits the ability to identify struggling students at an early stage. Educational institutions often collect large amounts of student data, but this data remains underutilized due to the lack of analytical tools. Manual evaluation methods are time- consuming and prone to errors, especially when dealing with large datasets. Additionally, these systems do not provide predictive insights or real-time feedback, making it difficult for educators to take timely corrective actions. The absence of automated analysis also reduces the efficiency of academic monitoring. Overall, the existing system lacks accuracy, scalability, and early prediction capabilities. This creates challenges in identifying at-risk students and prevents institutions from implementing proactive measures to improve academic performance.

PROPOSED SYSTEM

The proposed system introduces a machine learning-based approach for predicting student performance using educational data. It is designed as a web-based application that allows users to input student-related information and obtain prediction results efficiently. The system analyzes key academic parameters such as

attendance, study hours, previous exam marks, and assignment scores to determine the expected performance of a student. The system utilizes data preprocessing techniques to clean and format the input data before passing it to the machine learning model. Classification algorithms are used to analyze patterns in historical data and generate accurate predictions. The model identifies whether a student is likely to perform well or may require additional academic support. This helps educators take proactive measures to improve student outcomes. A user-friendly interface is developed using HTML, CSS, and JavaScript, while Flask is used for backend processing and handling user requests. The system ensures smooth interaction between the user interface and the prediction model. Secure data handling mechanisms are implemented to protect user information.

SYSTEM ARCHITECTURE

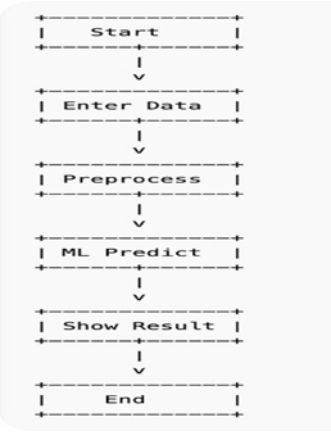


Fig 1: Computer Vision-Based Food Calorie Estimation System: Intelligent Dietary Monitoring

The Student Performance Prediction System is designed using a three-tier architecture to provide modularity, scalability, and efficient processing of the data. The three-tier architecture of the system consists of the Presentation Layer, Application Layer, and Model Layer. The Presentation Layer is the interface through which the user interacts with the system. The user is provided with the interface to insert student information such as attendance, study hours, previous marks, and assignment scores. The prediction result is also shown to the user in a simple and easy-to-understand format. The presentation layer is developed using web technologies such as HTML, CSS, and JavaScript to provide a user-friendly interface.

The Application Layer is the main processing unit of the system. The user interface and machine learning model interaction occur in this layer of the system. The user input is cleaned and processed for analysis in this layer. This layer is developed using Python programming language with frameworks such as Flask that handles server-side operations.

METHODOLOGY DESCRIPTION

The methodology of the Student Performance Prediction System follows a systematic approach to collect educational data and process it using machine learning techniques for analysis. The process starts with collecting the data, where student-related information is collected through the user interface. This includes student information related to their attendance, study hours, previous exam marks, and assignment marks. The next step is to process the collected data using machine learning techniques. Here, the collected data is cleaned and converted to a suitable format for analysis. The third step is to extract the most prominent features that affect student performance using feature extraction techniques. The collected features are fed into the machine learning model. In the model training step, machine learning algorithms such as Decision Tree, Random Forest, or Logistic Regression are applied to the collected educational data to train the model. The model is designed to learn from the historical educational data collected from various sources.

RESULTS AND DISCUSSION

The proposed Student Performance Prediction System was successfully developed and tested using both real-time and sample educational datasets. The system provides a user-friendly web interface where users can input student details such as attendance, study hours, previous exam marks, and assignment scores. The home page acts as the entry point, allowing users to navigate easily and access prediction features.

Home Page:

The home page serves as the entry point of the Student Performance Prediction System. It provides a simple and user-friendly interface that allows users to navigate through the application easily.



Fig 2: Home Page

Prediction Page:

The prediction page allows users to input detailed student information required for performance analysis. Users can enter parameters such as attendance, study hours, previous exam marks, and assignment scores. Additional attributes like gender, section, and parental education level are selected using dropdown menus.



Fig 3: Prediction Page

Result Page:

The result page displays the predicted student performance in a clear and organized manner. It presents subject-wise predicted scores such as Mathematics, Reading, and Writing. The system ensures that the results are easy to understand, enabling users to interpret the output effectively.This helps in identifying students who may require additional academic support.



Fig 4: Result Page

CONCLUSION

The Student Performance Prediction System using Educational Data provides an effective solution for analyzing and predicting student academic outcomes. By utilizing machine learning techniques, the system processes important educational parameters such as attendance, study hours, previous marks, and assignment scores to generate accurate predictions. The web-based interface ensures ease of use, allowing users to input data and obtain results in real time. The system reduces manual effort and enhances the efficiency of academic monitoring. It helps educators identify students who may require additional support and enables timely intervention. Overall, the proposed system demonstrates the practical application of data-driven approaches in education and contributes to improving student performance and decision-making in educational institutions.

FUTURE SCOPE

The Student Performance Prediction System can be further enhanced by incorporating advanced machine learning techniques such as deep learning and ensemble models to improve prediction accuracy. Future improvements may include integration with Learning Management Systems to automatically collect student data such as attendance and assignment performance. Real-time monitoring of student activities and

behavior can provide more dynamic and continuous predictions. The system can also be extended to generate personalized learning recommendations based on prediction results. Developing a mobile application would increase accessibility for both students and educators. Additionally, implementing advanced data visualization tools and stronger security mechanisms can improve usability and data protection. These enhancements can transform the system into a comprehensive academic analytics platform.

REFERENCES

1.Kedari, M., Harini, P., & Narayana, N. L. (2024). ANALYZE AND PREDICT OF HUMAN CYBER ATTACKERS USING ARTIFICIAL NEURAL NETWORK. *JOURNAL OF BASIC SCIENCE AND ENGINEERING*, *21*(1), 1529-1536.

2. Python Software Foundation, Python Official Documentation. Available: https://docs.python.org/3

3. Flask, Flask Web Framework Documentation. Available: https://flask.palletsprojects.com

4. Pandas, Data Analysis Library Documentation. Available: https://pandas.pydata.org

5. NumPy, Scientific Computing Library Documentation. Available: https://numpy.org

6. GeeksforGeeks, Machine Learning and Data Science Tutorials. Available: https://www.geeksforgeeks.org

7. W3Schools, HTML, CSS, and JavaScript Tutorials. Available: https://www.w3schools.com

8. J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2012.

9. R. Baker and K. Yacef, “The State of Educational Data Mining in 2009,” Journal of Educational Technology & Society, 2009.

10. C. Romero and S. Ventura, “Educational Data Mining: A Survey,” IEEE Transactions on Systems, Man, and Cybernetics, 2010.