Student Monitoring System Using Linear Regression Model

GE19612 - PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP PROJECT REPORT

Submitted by

AJAY KRISHNAN (2116220701018)

ANSH BHATIA (2116220701027)

ARUN MOHAN S (2116220701029)

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE

ANNA UNIVERSITY, CHENNAI

MAY 2025

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

BONAFIDE CERTIFICATE

Certified that this Project titled **" Student Monitoring System using Linear Regression Model "** is the bonafide work of **"AJAY KRISHNAN (2116220701018), ANSH BHATIA(2116220701027), ARUN MOHAN S(2116220701029)** who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE	**SIGNATURE**
Dr. P. Kumar., M.E., Ph.D.,	Dr. M. Ayyadurai ., M.E., Ph.D.,
HEAD OF THE DEPARTMENT	**SUPERVISOR**
Professor	Assistant Professor
Department of Computer Science	Department of Computer Science
and Engineering,	and Engineering,
Rajalakshmi Engineering College,	Rajalakshmi Engineering
Chennai - 602 105.	College, Chennai-602 105.

Submitted to Mini Project Viva-Voce Examination held on ______________

Internal Examiner **External Examiner**

ABSTRACT

The ability to monitor timely reaction and progress through a course, usually highly significant contribution to educational success is accepted as factors. In this study we introduce students achievement tracker, an online equipment that allows students to record their marks from internal assessment, so that they will progress through the semester so that they can get how good indications they can through the semester. The equipment appoints major Prophet variables and appoints linear regression techniques to provide more detailed reports of the student's performance, as well as predicts the final grade, on the score from the average interior scars and continuous assessment tests (CAT). Its purpose is to help both learners and teachers efficiently measure their progress, while students are to measure their progress by providing the necessary information to feel the areas for improvement. Between the ongoing continuous evaluation and the final result of the students, the project gives both the opportunity to adapt their performance to the project teacher and the learner. A student can monitor his progress through reports and determine that to get good marks in their examination, and to help the equipment to help make up your utility to help make up a suitable study plan through being able to secure a placement with their favorite dream companies. In addition, pacing tools can be useful use for all students, especially when they plan a study schedule mentioned.

ACKNOWLEDGMENT

Initially we thank the Almighty for being with us through every walk of our life and showering his blessings through the endeavor to put forth this report. Our sincere thanks to our Chairman **Mr. S. MEGANATHAN, B.E, F.I.E.**, our Vice Chairman **Mr. ABHAY SHANKAR MEGANATHAN, B.E., M.S.,** and our respected Chairperson **Dr. (Mrs.) THANGAM MEGANATHAN**, **Ph.D.,** for providing us with the requisite infrastructure and sincere endeavoring in educating us in their premier institution.

Our sincere thanks to **Dr. S.N. MURUGESAN, M.E., Ph.D.,** our beloved Principal for his kind support and facilities provided to complete our work in time. We express our sincere thanks to **Dr. P. KUMAR, M.E., Ph.D.**, Professor and Head of the Department of Computer Science and Engineering for his guidance and encouragement throughout the project work. We convey our sincere and deepest gratitude to our internal guides **Dr. G.M. SASIKALA** and **Dr. M. AYYADURAI** We are very glad to thank our Project Coordinator, **Dr. M. AYYADURAI** Assistant Professor Department of Computer Science and Engineering for his useful tips during our review to build our project.

AJAY KRISHNAN 2116220701018

ANSH BHATIA 2116220701027

ARUN MOHAN S 2116220701029

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 GENERAL

The student monitoring system relies on the efficient linear regression method for its operation. Data science and statistics use linear regression as a methodology for showing the relationship between different variables. The student monitoring system extracts a final grade from internal assessment marks by using different bands of scoring values. A higher quantity of system-used data leads to more precise final grade predictions. The team selected linear regression because it presents basic understanding and few practical applications. Linear regression functions as an objective tool for problem-solving when most input/output relationships are established already. The input process of continuous evaluation testing uses points from Cat 1 and Cat 2 as well as points from internal marking along with assignment or appearance points. A single quantitative semester grade belongs to the output category. Linear regression represents an appropriate solution because all analyzed variables operate as numbers.

1.2 OBJECTIVE

This student monitoring system, based on linear regression, is to provide an intelligent data-driven approach that allows students to continuously monitor and forecast students according to various average indicators. The student monitoring system will be able to analyze the key indicators of the student academic performance (an average factor contributing to their performance will include the attendance percentage or internal assessment score or assignment marks) and teachers will be allowed to find the relationship established through the characteristics and later academic performance (end result). A accurate prediction of the display of an intended student is to identify students who can be at risk of underpartance, and can provide the speed and depth of useful analysis for timely intervention from teachers. The power power of the system will help educational organizations to make informed decisions, customize their intervention with their students, and eventually the overall student will help increase experience and results. The student monitoring system is also designed to increase the level of transparency and efficiency in the supervision of student performance by establishing an automatic approach to monitor the performance through single centralized digital platforms.

1.3 EXISTING SYSTEM

In the current system of monitoring student progress, student performance is collected and usually tracked through manual processes or underdeveloped digital devices, including report cards, spreadsheets, and periodic evaluation by teachers. The existing systems of student performance tracking are non-composed in nature and are highly dependent on human observation and human intervention. Teachers rely on assessing progress through score, attendance records and assignment markings, but usually students do not analyze and synthesize performance data to identify the risk of educational dismissal, nor to predict future performance. In addition, existing systems do not miss data from real -time monitoring, or insight or previous roles and/or reports, which will allow them to take corrective action quickly. This contributes to early warnings of educational decline, or educational support for some students is not kept on time. In its light, the current mechanism lacks data-informed decision making methods to monitor the performance of the student and therefore does not directly.

CHAPTER 2

LITERATURE SURVEY

The study researchers provided for the possibility of predicting the student's performance in [1]. Romero and Ventura (2010) made a comprehensive review of educational data mining (EDM), which allowed researchers to achieve insights in assessing the effectiveness of student behavior and processes. Their work emphasized opportunities with educational data, in particular, make future models to help profile student behavior so that those risky students can be carried forward and properly handled by educationalists to improve the improvement in viable forms. It was also seen on the association rule mining based on specific application of interests based on classification, clustering, and educational procedures especially the student behavior, intentions and efficiency. The comprehensive, comprehensive review of 123 studies mainly in technology and education subjects assisted in relation to various EDM equipment, which allowed the outline for further research as well as specific educational exercises tied to adaptive teaching systems. The review underlined the specific case studies to note applications that follow the EDM principles, in which students had benefits within intelligent tuition systems and adaptive teaching systems. Furthermore, the review was clear to explain that the Learning Management System (LMS) within the educational data and EDM provided continuous monitoring of students aspects/sources of students to facilitate institutional time response/signals at institutional time that is important within data that can guide educational decisions.

In [2], Kotsiantis et al. (2004) contributed a significant advancement in the prediction of student performance by applying various machine learning algorithms, most notably Naïve Bayes and decision trees. Their research aimed to address the pressing challenge of student dropout, especially in distance learning environments, where monitoring student engagement is inherently difficult. The study emphasized that a combination of academic data—such as grades, test scores, and assignment completion rates—alongside demographic factors like age, gender, and prior educational background, forms a robust dataset for training predictive models.

The core of their methodology involved comparing the performance of these algorithms in classifying students into at-risk and not-at-risk categories, enabling early intervention by educators. Decision trees were particularly noted for their interpretability, making them suitable for educational settings where transparency in decision-making is crucial. The Naïve Bayes classifier, on the other hand, was praised for its simplicity and effectiveness even with small datasets, highlighting its potential for use in institutions with limited resources.

Moreover, Kotsiantis et al. stressed the importance of tailoring these models to the unique dynamics of distance learning, where student participation can fluctuate and is often influenced by external factors such as work commitments or limited access to resources. Their work demonstrated that predictive models must be adaptive and continuously updated to reflect the evolving patterns of learner behavior. The researchers also recommended incorporating real-time learning analytics, drawn from Learning Management Systems (LMS), to enhance the timeliness and relevance of the predictions.

The study laid a foundation for future research in dropout prediction by showcasing the power of combining data-driven techniques with pedagogical strategies. It encouraged institutions to move towards proactive academic advising, where at-risk students are identified early and offered personalized support plans. In doing so, the study not only offered technical insights into model selection and feature importance but also underscored the ethical responsibility of educational institutions to use such tools for the betterment of student outcomes. The contribution of Kotsiantis et al. continues to be a valuable reference in educational data mining and student success prediction frameworks.

In [3], Al-Barrak and Al-Razgan (2016) employed decision tree algorithms to forecast students' final GPA, focusing on interpretability and practical deployment within academic environments. Their study emphasized that decision trees not only offer high predictive accuracy but also provide clear, visual representations of the decision-making process—an essential feature for educational administrators and faculty members aiming to understand the reasoning behind a student's risk level. The research involved data collection from various academic metrics such as attendance, coursework performance, and test scores, which were then used to develop a model capable of predicting GPA outcomes with notable precision.

A critical finding in their study was the identification of performance thresholds—specific academic indicators that correlate strongly with poor final performance. These thresholds, when implemented into academic advising systems, could act as triggers for early intervention. The researchers also highlighted that the model's simplicity made it adaptable for real-time integration with institutional Learning Management Systems (LMS), thus making predictions actionable in ongoing educational cycles.

Furthermore, the authors discussed scalability and sustainability in implementing decision tree models within educational settings. They proposed that such models can be periodically retrained with new student data, ensuring the system evolves and remains effective over time. The study validated the potential of decision trees in supporting educators' efforts to boost academic performance, enhance student retention, and promote proactive educational strategies, thereby solidifying their role in data-driven student monitoring systems.

In [4], Ahmed et al. (2019) focused on the application of multiple linear regression techniques to predict academic outcomes using variables such as test scores, continuous assessment results, attendance, and demographic features. Their findings affirmed the practicality and interpretability of regression models in modeling relationships between independent variables (student characteristics and academic inputs) and the dependent variable (final academic performance).

The researchers demonstrated that linear regression could accurately determine how strongly each variable influenced academic success. For instance, attendance was found to have a high correlation with performance in certain disciplines, highlighting the need for policies that encourage classroom engagement. The model also allowed for the calculation of predicted scores based on existing academic inputs, giving educators a quantitative benchmark for identifying students who may need academic support.

Ahmed et al. stressed that regression models were particularly effective in structured academic environments with consistent evaluation criteria. They also proposed that integrating this model with LMS and academic databases would allow for near-real-time tracking of student progress, offering personalized feedback and early warnings. Their study served as a foundation for integrating linear regression models into broader student monitoring systems, especially in institutions looking for low-complexity, high-impact predictive tools.

In [5], Pandey and Pal (2011) emphasized the necessity of interactive feedback systems in educational settings, especially those driven by data analytics. Their research advocated for tools that can provide students with real-time insights into their learning progress, bridging the communication gap between students and educators. The authors argued that timely feedback based on performance data enhances student motivation, engagement, and learning outcomes.

The study introduced a feedback loop model where student activity—such as quiz scores, assignment submissions, and class participation—was continuously monitored and analyzed to generate personalized feedback. This feedback could take various forms, including automated alerts, performance summaries, and suggestions for improvement. Importantly, they underlined the psychological impact of feedback, showing that students who receive prompt and constructive responses tend to exhibit increased academic confidence and discipline.

Pandey and Pal also discussed the integration of such systems with existing LMS platforms, ensuring that students could access feedback through familiar digital environments. The study concluded that feedback systems, when driven by educational data mining (EDM), could play a crucial role in developing adaptive learning environments that respond dynamically to student needs, thereby promoting continuous academic improvement and reducing dropout rates.

In [6], Lin et al. (2023) presented a comprehensive survey of deep learning techniques within the scope of Educational Data Mining (EDM). Their work emphasized the capability of deep learning models to handle and extract meaningful patterns from large, complex, and unstructured educational datasets. The survey explored architectures like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks, all of which have shown promising results in tasks such as knowledge tracing, student engagement detection, and performance prediction.

The authors provided critical comparisons between traditional machine learning models and deep learning approaches, highlighting that while deep learning requires more data and computational resources, it offers superior performance in capturing non-linear relationships and sequential dependencies. For instance, LSTMs were particularly effective in tracking a student’s learning trajectory over time, making them suitable for predicting future performance based on past behaviors.

Additionally, the paper discussed the application of attention mechanisms and hybrid models that combine deep learning with explainable AI techniques to address the “black-box” nature of deep neural networks. Lin et al. concluded that deep learning has the potential to revolutionize educational analytics by enabling highly personalized learning paths, intelligent feedback systems, and automated academic risk assessment tools.

In [7], Moubayed et al. (2024) explored the use of deep learning models—specifically CNNs and RNN-LSTMs—for predicting student performance in online learning environments, with a focus on mid-course assessments. Their study addressed the increasing reliance on e-learning platforms and the need for robust analytics tools capable of handling high-dimensional and time-series educational data.

The researchers collected data across various platforms, including quiz attempts, forum interactions, login frequency, and assignment completion timelines. Using this data, they trained deep learning models to detect early patterns that indicate whether a student is on track or at risk of underperforming by the end of the course. The RNN-LSTM architecture was particularly successful in modeling sequential student behavior, such as changes in performance over consecutive weeks.

One of the key takeaways was the real-time applicability of these models. Moubayed et al. proposed integration with LMS dashboards, enabling educators to receive live risk assessments of students and take necessary interventions. Their work also discussed challenges such as data sparsity, ethical concerns regarding student privacy, and the need for interpretability in AI-driven education systems. Despite these hurdles, their research demonstrated that deep learning can significantly enhance student monitoring in virtual learning environments.

In [8], Yadav et al. (2021) conducted a systematic review of various machine learning algorithms used for predicting student performance and understanding academic outcomes. Their review synthesized findings from multiple studies and provided a comparative analysis of models such as Support Vector Machines (SVM), Decision Trees, Naïve Bayes, k-Nearest Neighbors (k-NN), Random Forest, and Ensemble Learning techniques.

The review identified common variables used in student performance prediction—such as internal marks, attendance, socio-economic status, and digital learning behavior—and evaluated how different models handled these features. Yadav et al. found that while ensemble models typically outperformed individual algorithms in terms of accuracy, decision trees and Naïve Bayes remained popular due to their simplicity and interpretability.

They also addressed the growing trend of using hybrid models that combine EDM techniques with psychological and behavioral data to create a holistic understanding of student performance. The study emphasized the need for more inclusive datasets, ensuring that predictive models account for diverse learning environments and student backgrounds. Ultimately, Yadav et al.'s review reinforced the growing body of evidence that ML-based systems can serve as powerful tools for early intervention, personalized instruction, and the design of future-ready educational infrastructures.

CHAPTER 3

PROPOSED SYSTEM

3.1 GENERAL

The proposed system is a Student Achievement Tracker that uses machine learning and educational data mining techniques to predict student performance and identify those at academic risk. It analyzes various data such as academic scores, attendance, and behavior to provide real-time performance insights. Models like Decision Trees, Linear Regression, and Naïve Bayes are employed to classify students and forecast outcomes. The system offers personalized feedback and alerts, helping educators take timely actions. It includes a visual dashboard for tracking progress and integrates with existing Learning Management Systems for continuous data flow. The goal is to support proactive decision-making and improve learning outcomes. The system is designed to be scalable and adaptable to both offline and online environments. It promotes data-driven teaching strategies and encourages student engagement. Ultimately, it empowers institutions to monitor, support, and enhance student achievement effectively.

OBJECTIVES

The primary objective of the proposed Student Achievement Tracker system is to leverage machine learning techniques to accurately predict student performance and identify at-risk individuals early, enabling timely academic interventions. The system aims to analyze various data points such as test scores, attendance, behavioral patterns, and demographic information to generate actionable insights for educators. It seeks to enhance student engagement through real-time personalized feedback and facilitate data-driven decision-making within educational institutions. Additionally, the system is designed to provide a user-friendly interface for both students and faculty to track academic progress and performance trends. By integrating with Learning Management Systems, the objective is to ensure continuous monitoring, promote adaptive learning strategies, and ultimately improve overall educational outcomes.

3.2 SYSTEM ARCHITECTURE DIAGRAM

The architecture of the Student Achievement Tracker begins with collecting data from multiple sources such as exam scores, attendance records, assignment submissions, and other academic metrics. This raw data undergoes preprocessing which includes cleaning, normalization, and selecting relevant features for training. The cleaned data is then passed into a machine learning model—specifically a linear regression model—which is trained to identify patterns and predict student performance. Based on the model's predictions, the system classifies students by risk levels and generates performance trends. These predictions are then visualized through a user-friendly dashboard, and feedback such as alerts or suggestions is automatically generated to help students and faculty take timely actions. Finally, the entire system is integrated with existing Learning Management Systems (LMS) to ensure seamless access, real-time updates, and interactive feedback within the academic workflow.

Fig 3.1: System Architecture

3.1 DEVELOPMENTAL ENVIRONMENT

3.1.1 HARDWARE REQUIREMENTS

The hardware specifications could be used as a basis for a contract for the implementation of the system. This therefore should be a full, detailed description of the whole system. It is mostly used as a basis for system design by the software engineers.

Table 3.1 Hardware Requirements

COMPONENTS	SPECIFICATION
PROCESSOR	Intel Core i3 or higher
RAM	4 GB RAM or more
POWER SUPPLY	+5V power supply

3.1.2 SOFTWARE REQUIREMENTS

The software requirements document contains the system specifications. This is a list of things which the system should do, in contrast to the way in which it should do things. The software requirements are used to base the requirements. They help in cost estimation, complete tasks, and track team progress in the development .

Table 3.2 Software Requirements

COMPONENTS	SPECIFICATION
Operating System	Windows 7 or higher
Frontend	Html,css and javascript
Backend	XAMP,MYSQL
Database	SQL database
Development IDE	Visual Studio Code

3.2 DESIGN OF THE ENTIRE SYSTEM

3.2.1 ACTIVITY DIAGRAM

The activity diagram of the Student Achievement Tracker begins with the collection of student data, including academic scores, attendance, and assignment details. This raw data then undergoes preprocessing, which involves cleaning errors, normalizing values, and selecting relevant features necessary for building the predictive model. Once prepared, the data is used to train a linear regression model that learns from past trends to forecast future student performance. The model then predicts outcomes such as risk of failure or expected GPA, which are analyzed to assess each student's academic standing. Based on this analysis, the system generates personalized feedback in the form of alerts and actionable suggestions for both students and educators. These results are displayed on an interactive dashboard, allowing stakeholders to visualize progress and performance. Finally, the system updates the Learning Management System (LMS) and notifies the relevant users, closing the loop and enabling continuous academic monitoring and improvement.

```
 Start
   |
   v
Collect Student Data (Marks, Attendance, etc.)
   |
   v
Preprocess Data (Clean, Normalize, Select Features)
   |
   v
Train Linear Regression Model
   |
   v
Predict Student Performance
   |
   v
Analyze Prediction Results
   |
   v
Generate Feedback (Alerts, Suggestions)
   |
   v
Display Results on Dashboard
   |
   v
Update LMS / Notify Stakeholders
   |
   v
End
```

Fig 3.2: Activity Diagram

3.4.2 SEQUENCE DIAGRAM

The sequence diagram illustrates the interaction between the components of the Student Achievement Tracker system. The process begins when the student submits their academic data through the web interface, which sends the input to the backend server. The backend then communicates with the database to fetch historical data or store new input records. The server preprocesses this data and forwards it to the machine learning model for prediction. The model either loads a pre-trained linear regression model or trains it based on the current dataset, then predicts the student's performance outcome (such as GPA or pass/fail status). The results are sent back to the backend, which in turn sends them to the web interface for visualization. Finally, the student can view the predictions and feedback through a clean, user-friendly interface. This real-time interaction ensures the system provides timely insights.

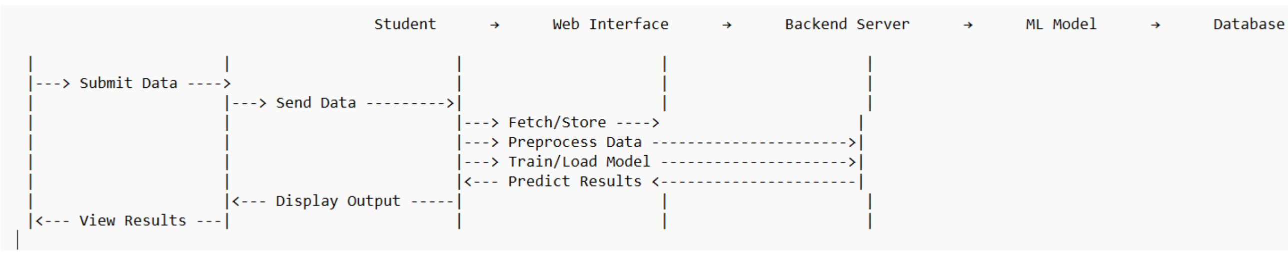

Fig 3.3:sequence Diagram

CHAPTER 4

MODULE DESCRIPTION

The Student monitoring system project consists of several integrated modules that work together to provide predictions and feedback on student performance. The **Data Collection Module** gathers student-related information such as academic marks, attendance, and demographic details, either through manual entry or by fetching data from an LMS. The **Data Preprocessing Module** cleans and normalizes the raw data, handling missing values and ensuring it is ready for analysis. In the **Model Training and Prediction Module**, a linear regression model is either trained or loaded to predict student outcomes based on the processed data. The **Performance Evaluation Module** measures the accuracy of predictions, providing key metrics like R^2 score or RMSE to assess the model's effectiveness. **Feedback Generation** uses these predictions to provide actionable insights, such as alerts or recommendations to improve student performance. The **User Interface Module** enables interaction with the system, allowing students, faculty, and admins to view predictions, trends, and feedback through a secure dashboard. Finally, the **Database Management Module** ensures the safe and efficient storage of data, handling updates, and managing academic records. Together, these modules enable the system to support real-time decision-making and personalized academic interventions.

4.1 SYSTEM ARCHITECTURE

4.1.1 USER INTERFACE DESIGN

The user interface of the Student Monitoring System using Linear Regression Model is designed to be simple, interactive, and user-friendly, catering to three primary users: students, faculty, and administrators. The front-end interface provides secure login portals for each user type. Once logged in, students can view their academic records, prediction scores, and personalized feedback, while faculty can enter and update academic data and monitor student performance trends. Administrators have full access to oversee the entire system, manage users, and generate institutional reports. The dashboard includes features like performance graphs, risk alerts, and downloadable reports. The interface is built using HTML, CSS, and JavaScript, and communicates with the backend via secure API calls. The architecture includes modules such as Data Collection, Preprocessing, Linear Regression Prediction, Feedback Generation, and Data Storage. The system architecture connects the User Interface to a Web Server, which processes data through the Linear Regression Model module. The Database stores user profiles, academic data, and model results, while the Feedback Module pushes alerts and insights back to the interface, ensuring a real-time, responsive, and efficient user experience.

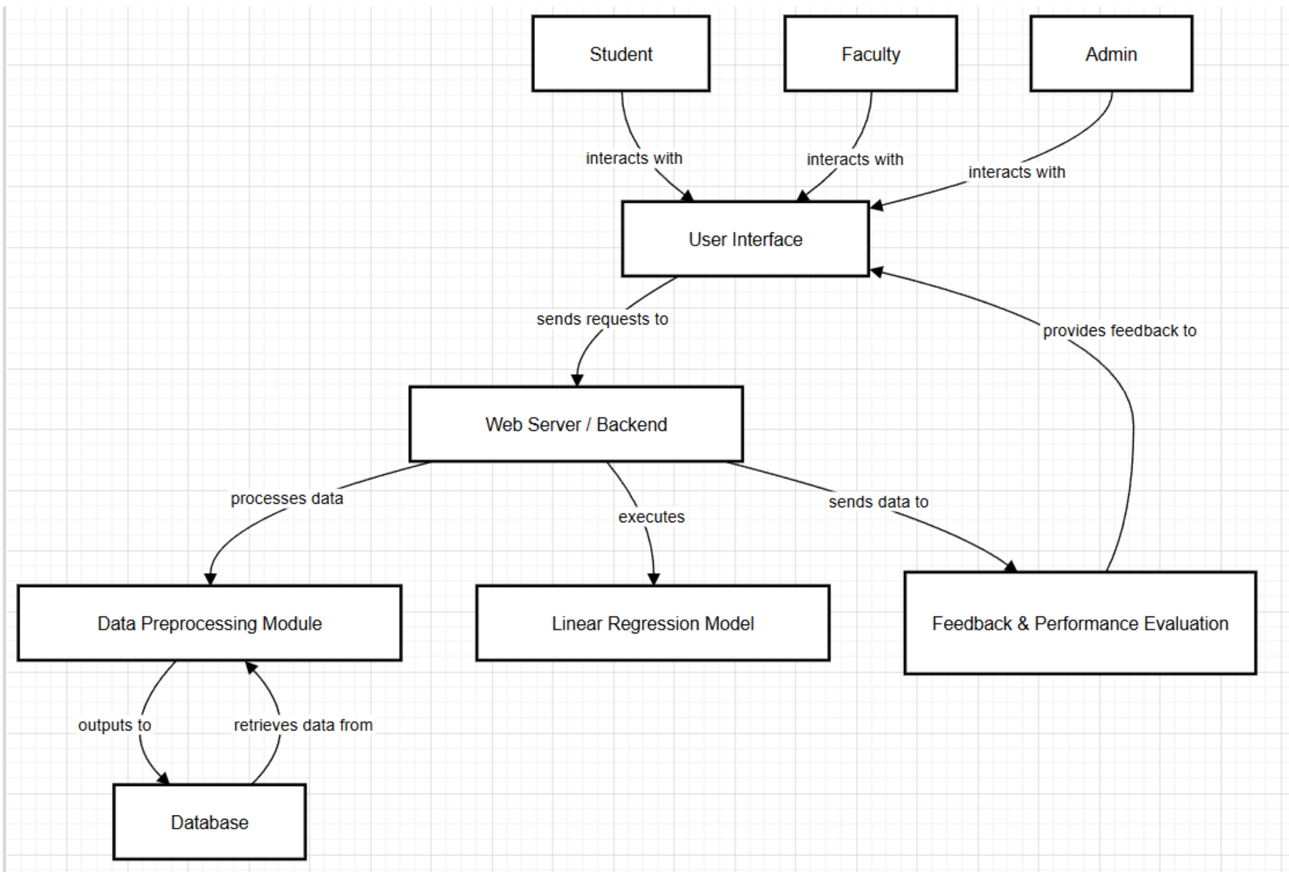

Fig 4.1: ARCHITECTURE DIAGRAM

4.1.2 BACK END INFRASTRUCTURE

The backend infrastructure of the Student Monitoring System using a Linear Regression model is designed to ensure efficient data processing, secure storage, and seamless integration of predictive analytics. It primarily consists of a relational database such as MySQL or PostgreSQL for storing student data including attendance, academic scores, and demographic information. A server-side framework like Node.js, Django, or Flask handles data processing and manages requests between the frontend and the database. The core of the backend includes a Python-based machine learning module implementing the Linear Regression algorithm, which is responsible for analyzing the input features and predicting student performance outcomes. RESTful APIs are used to enable communication between the user interface and the backend services. To maintain data security and integrity, authentication and authorization mechanisms are implemented, possibly using JSON Web Tokens (JWT). This backend infrastructure supports real-time data analysis and feedback, facilitating timely academic interventions and monitoring.

4.2 STUDENT ATTENDANCE TRACKING WORKFLOW

4.2.1 Registration of Students

The student accesses the registration page and enters personal details such as name,email, department, and year of study.

The system validates the input data and stores it securely in the backend database.

Upon successful registration, a confirmation message is displayed, and login credentials are generated if needed.

4.2.2. Student Login and Attendance Request

The student navigates to the login page and enters their registered email and password. The system authenticates the credentials against the database using secure login logic. If validated, the student is granted access to their dashboard with academic and attendance features.

4.2.3 Report Generation

The admin can generate attendance reports for selected time periods to evaluate student attendance trends, punctuality, and studying hours.

4.3 SYSTEM WORK FLOW

4.3.1 User Authentication and Profile Management:

Admin logs in to the system, creates new student profiles, and assigns passwords. Students receive their credentials and log in using their unique ID and password.

4.3.2 Attendance Recording:

Students mark attendance by clicking **“Check-In”** and **“Check-Out”** buttons. The system captures and stores timestamps in the MySQL database.

4.3.3 Attendance Analysis and Monitoring:

The system evaluates check-in/check-out times against standard school hours to determine regularity. Admin monitors individual and overall attendance records through the dashboard.

4.3.4 Report Generation:

Admin selects a date range to generate attendance reports, which summarize student presence, lateness, and study hours.

CHAPTER 5

IMPLEMENTATION AND RESULTS

5.1 IMPLEMENTATION

The implementation of the Student Monitoring System using a Linear Regression model involves several integrated components working together to monitor and predict student performance. First, a web-based user interface is developed using HTML, CSS, and JavaScript to allow students and faculty to interact with the system. The backend is built using Python and Flask or Django, which handles requests, manages user authentication, and processes attendance and academic data. A relational database like MySQL is used to store student profiles, attendance records, and test scores. The core machine learning module utilizes the Linear Regression algorithm implemented in Python using libraries such as scikit-learn to analyze historical data and predict students' future academic performance. The system fetches relevant data, feeds it into the model, and generates a predicted performance score for each student, which is displayed through the dashboard. The project also includes real-time feedback and alerts to notify students and faculty of potential academic risks, enabling early intervention and personalized guidance.

5.2 OUTPUT SCREENSHOTS

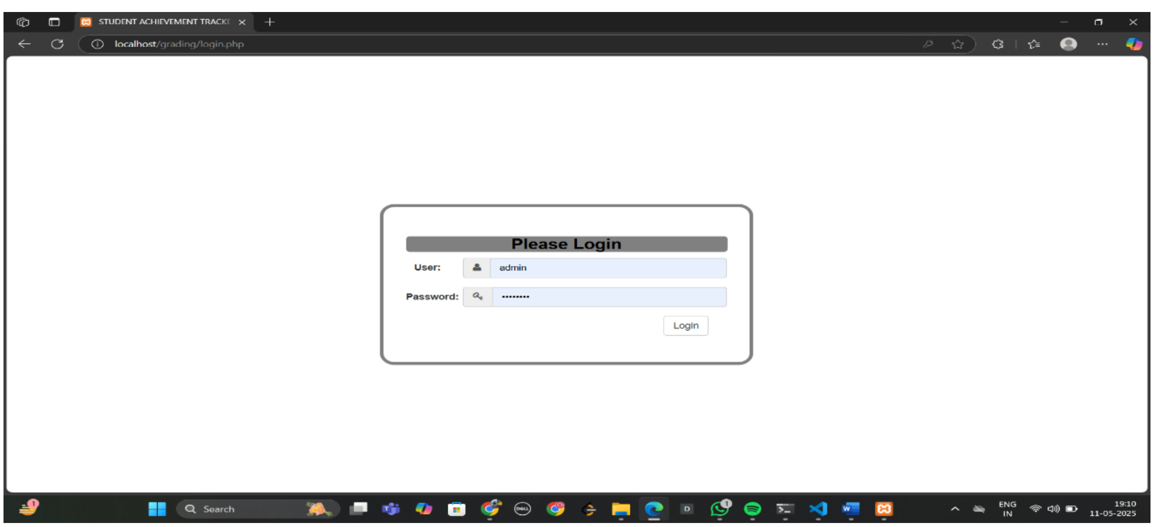

Fig 5.2.1 Login page

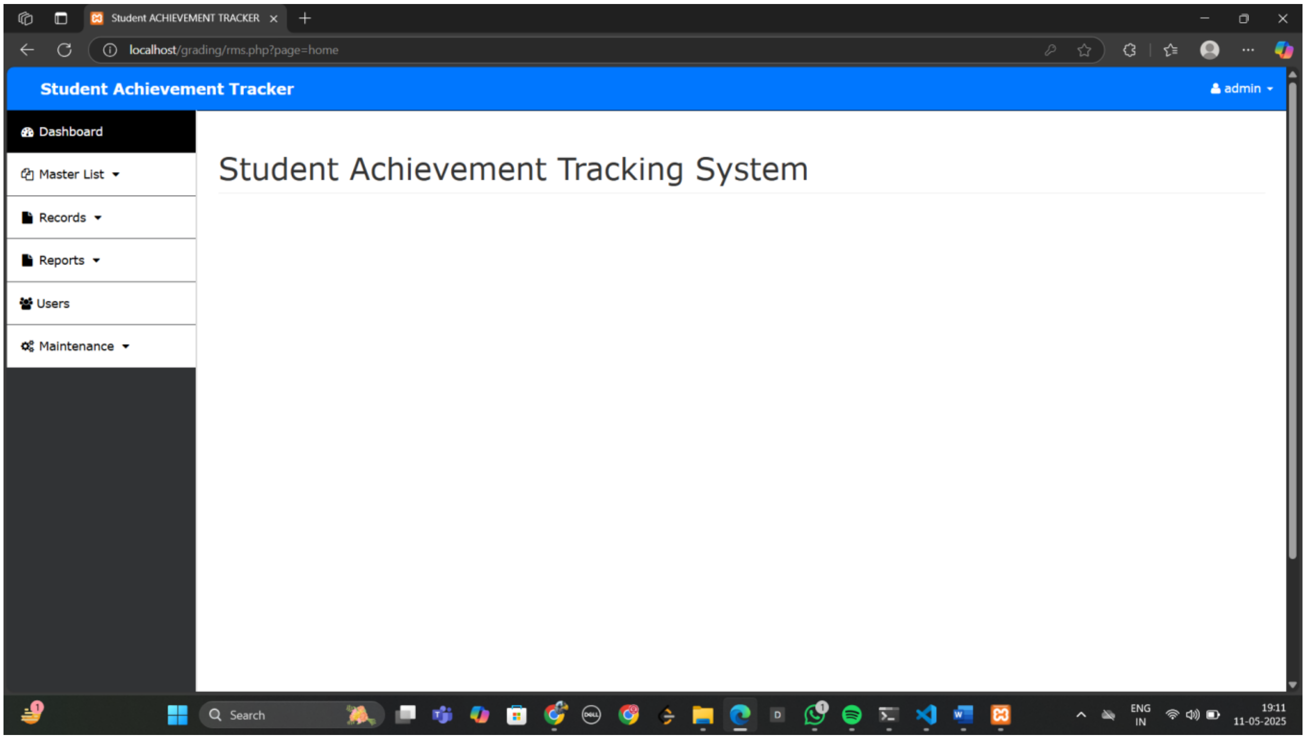

Fig 5.2.2 Dashboard

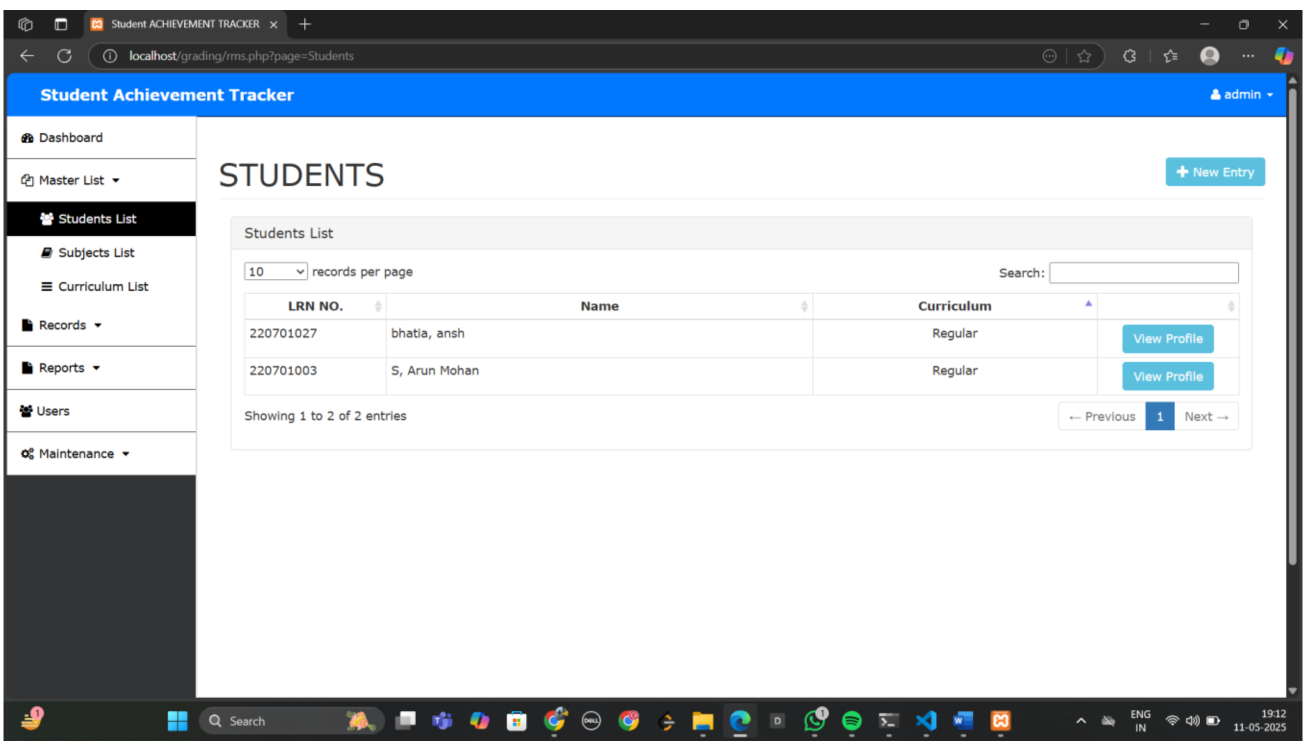

Fig 5.2.3 Student List

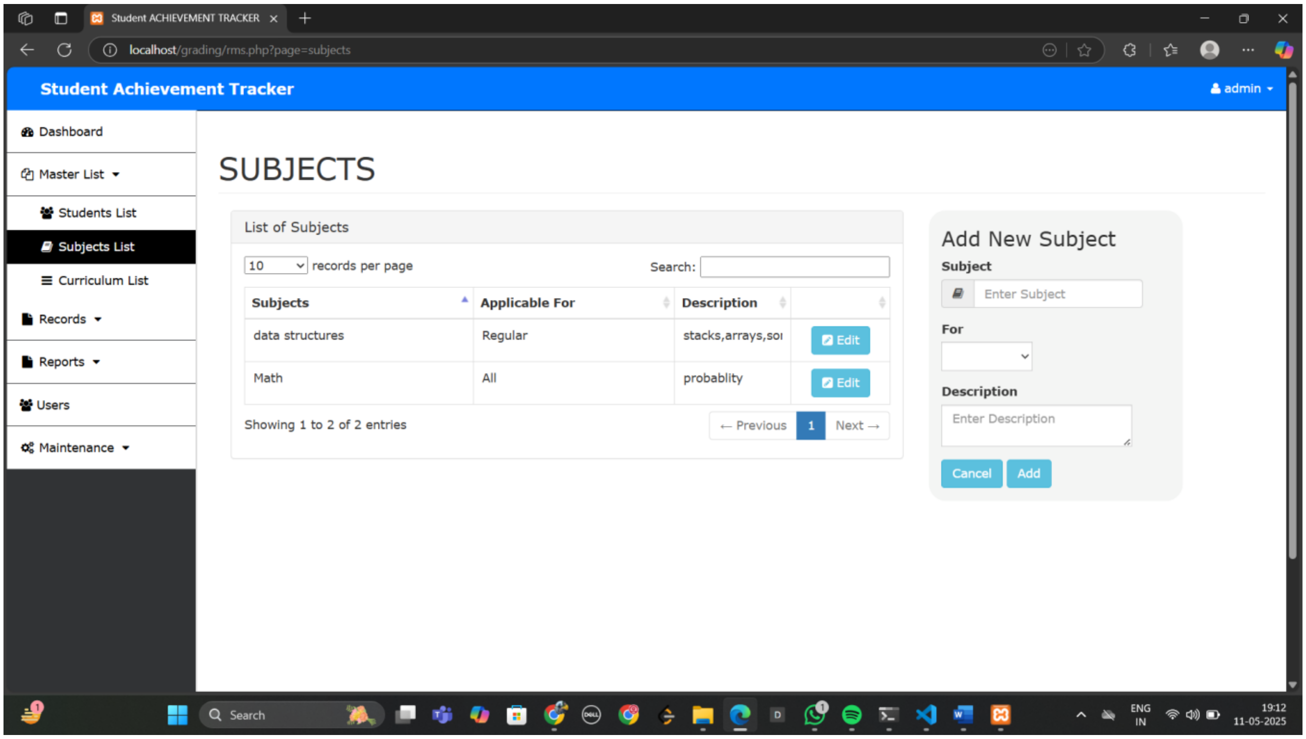

Fig 5.2.4 Subject List

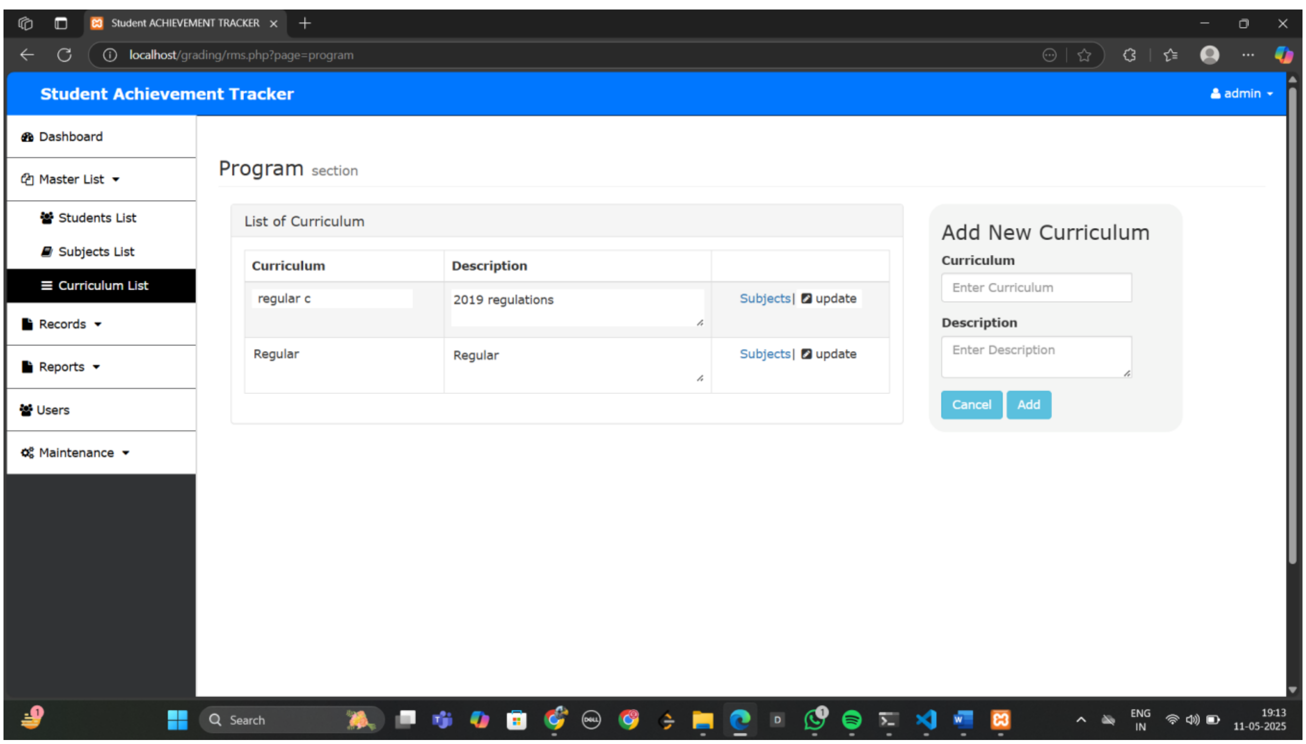

Fig 5.2.5 Curriculum List

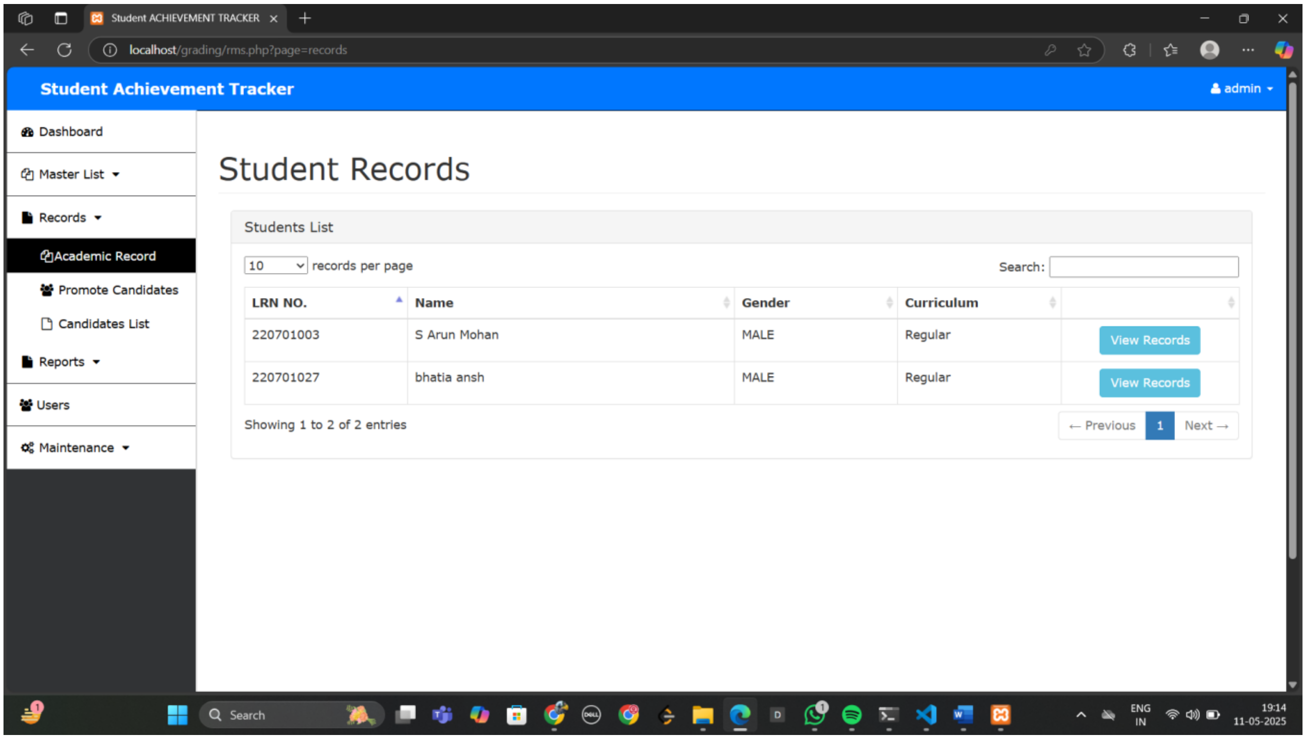

Fig 5.2.6 Academic Records

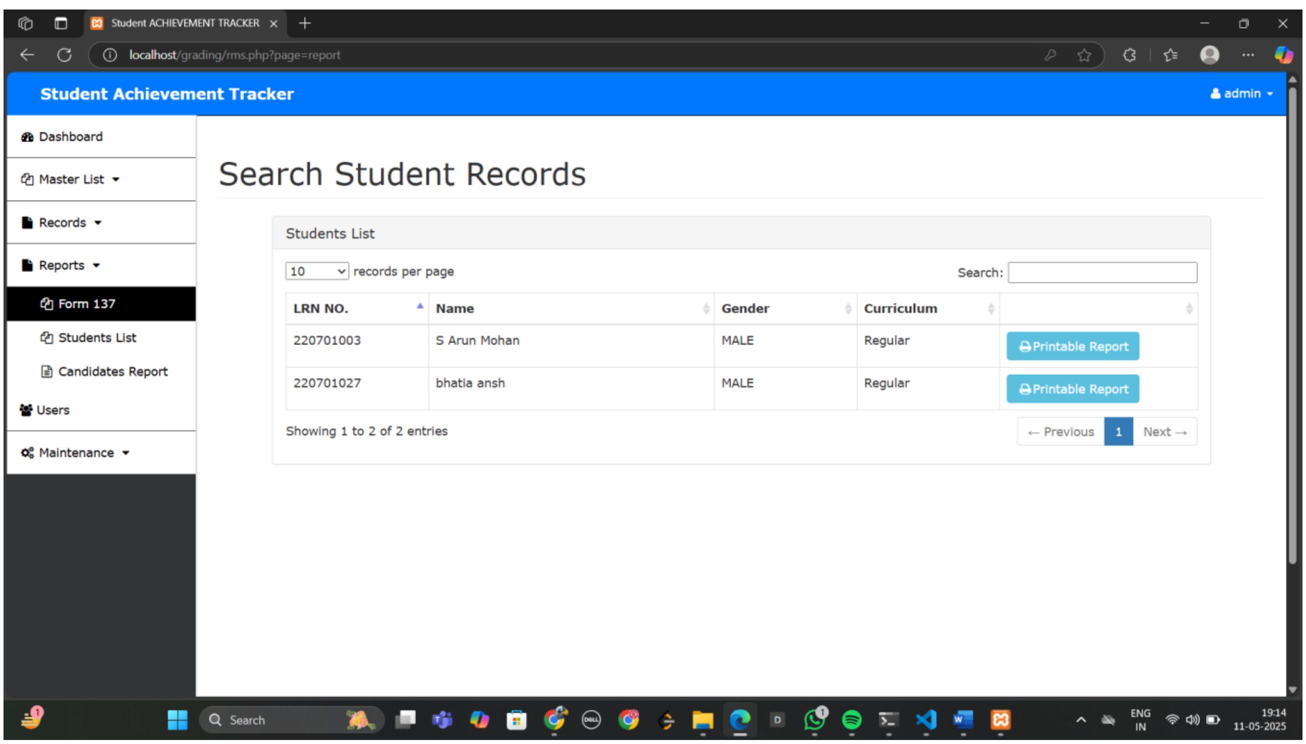

Fig 5.2.7 Form Generation

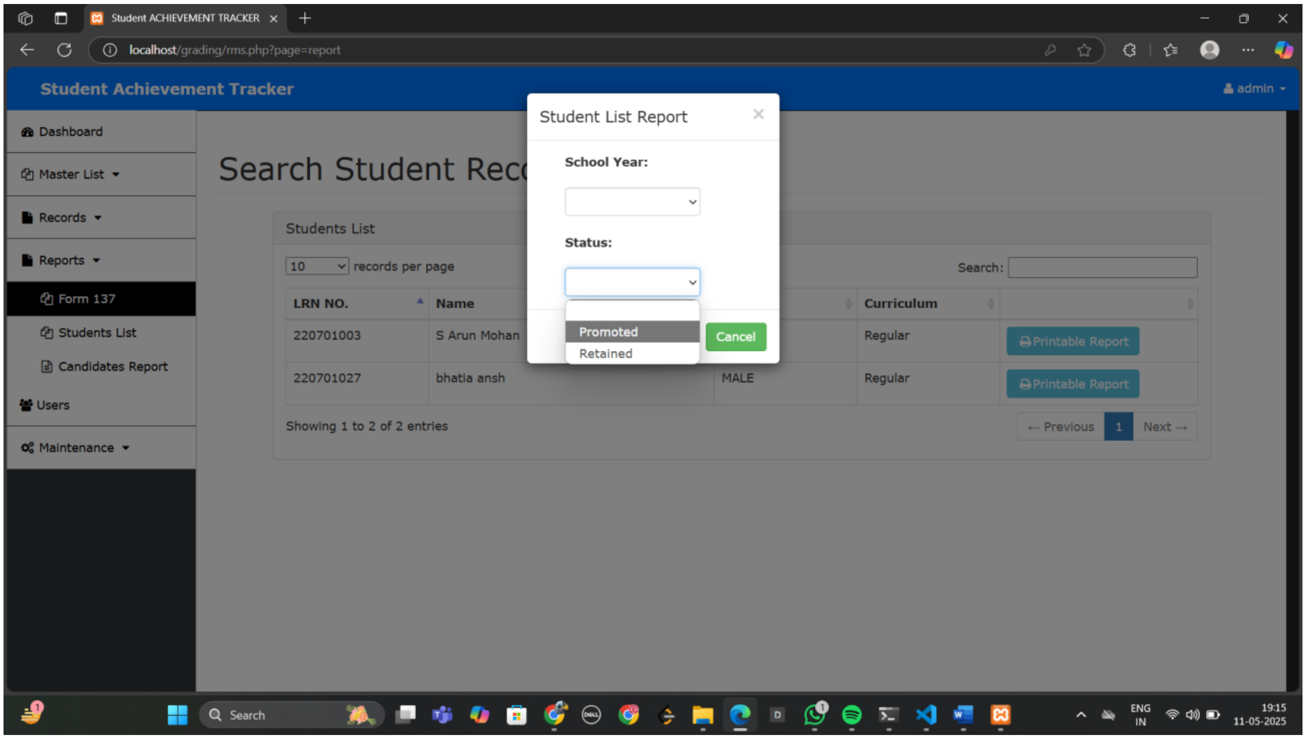

Fig 5.2.8 Student list Report

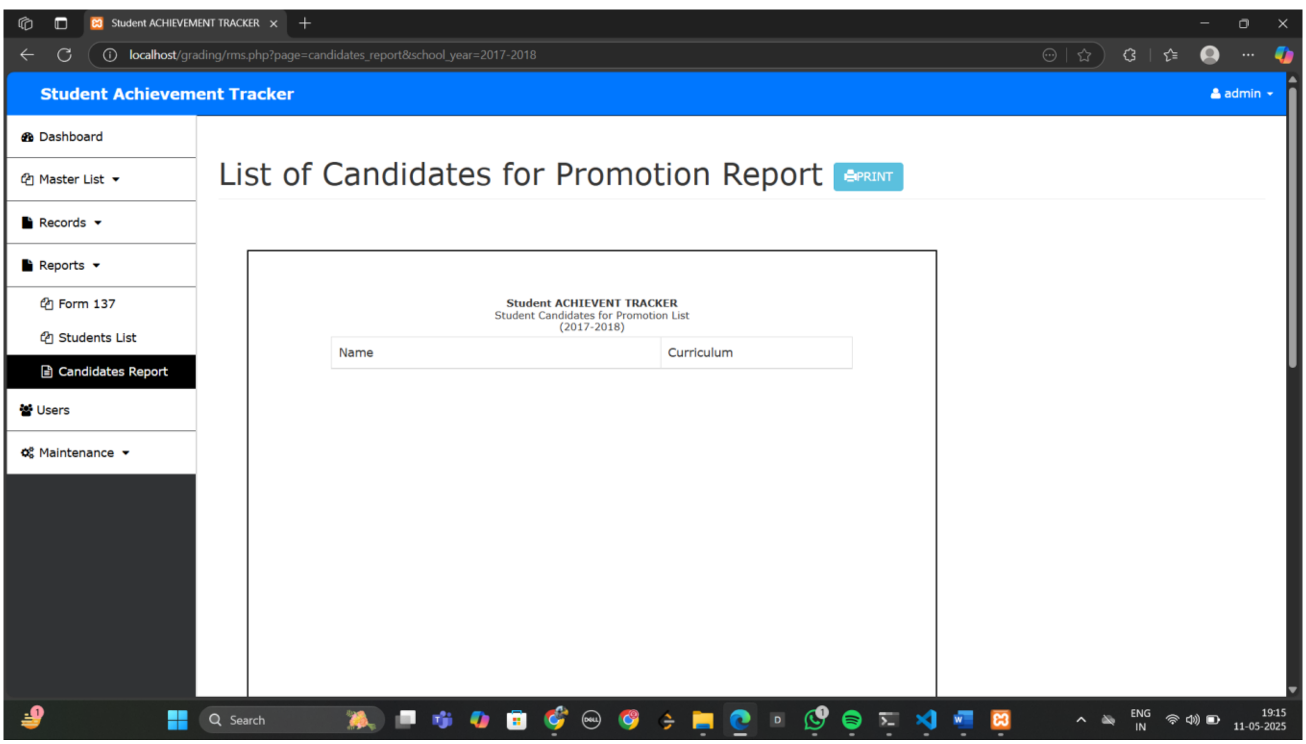

Fig 5.2.9 Candidate List

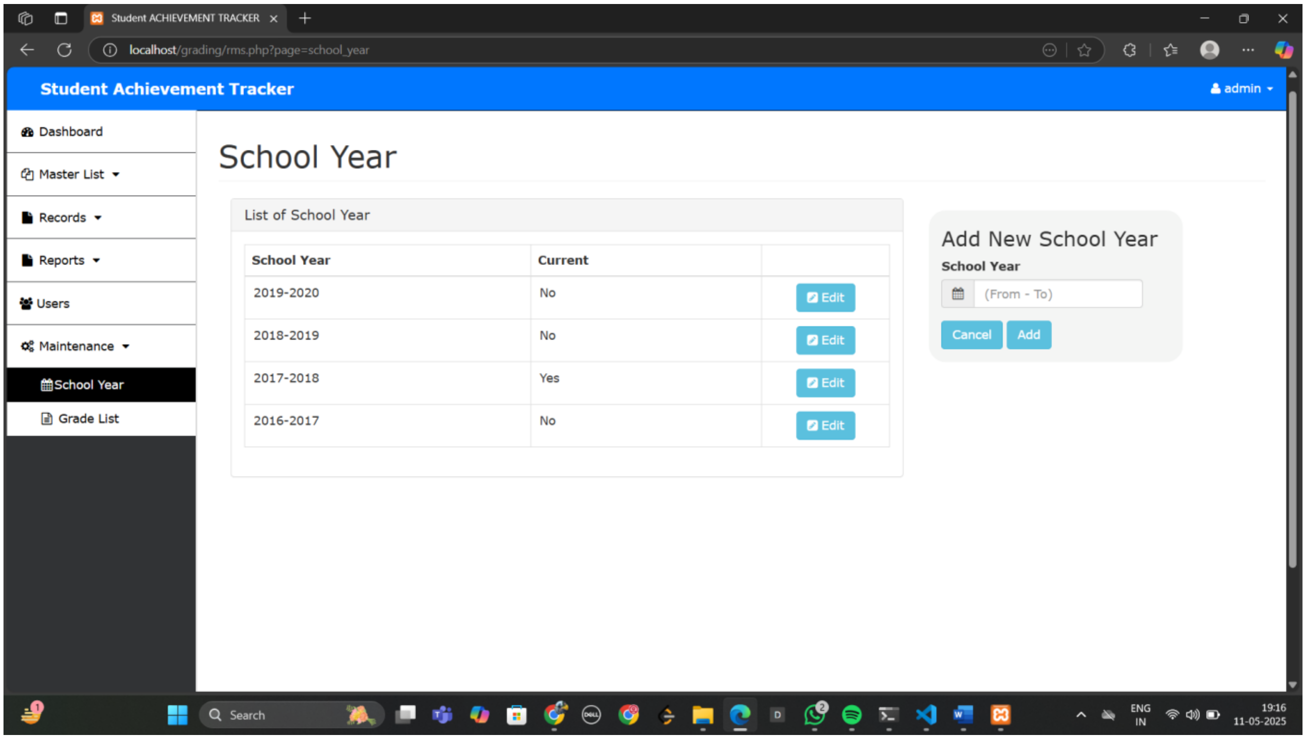

Fig 5.2.10 School Year module

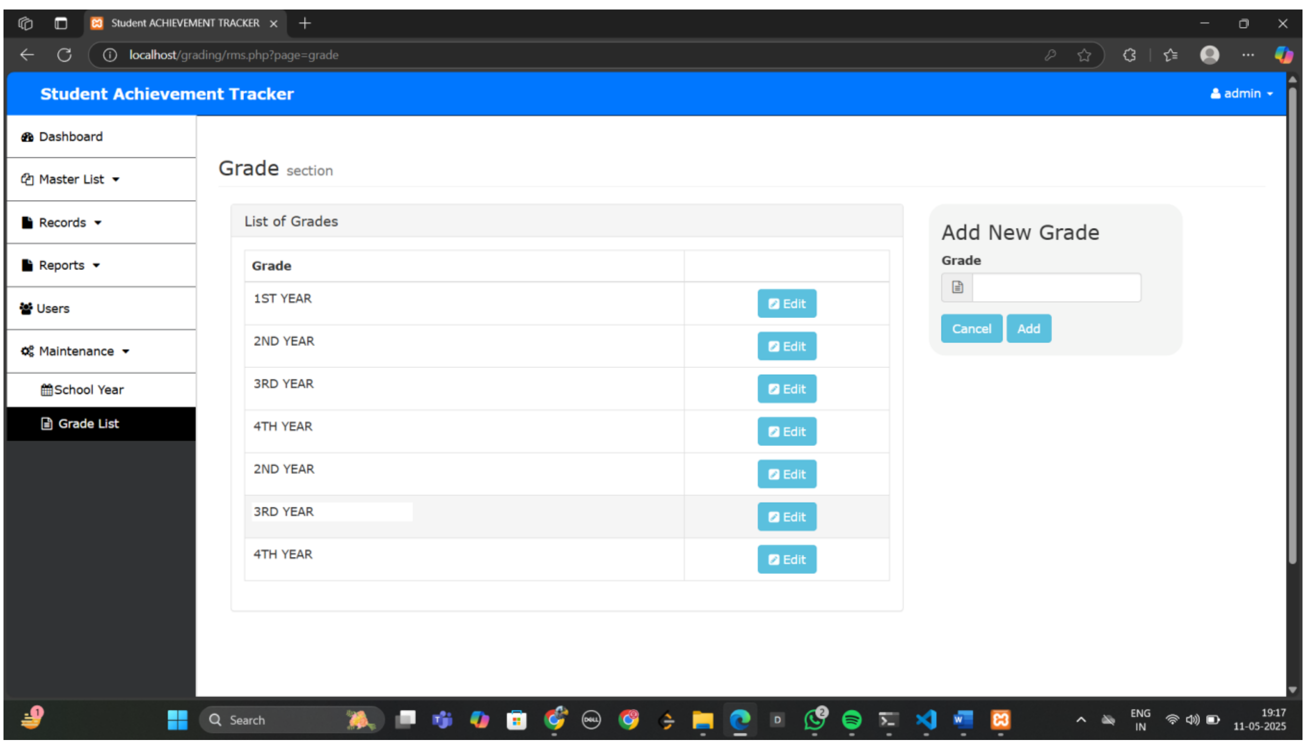

Fig 5.2.11 Grade list

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENT

6.1 CONCLUSION

The Student Monitoring System using a Linear Regression model effectively addresses the need for early identification of students at academic risk by leveraging data-driven insights. By integrating user-friendly interfaces, automated attendance tracking, academic performance monitoring, and predictive analytics, the system empowers both students and educators to make informed decisions. The use of linear regression provides a simple yet powerful approach to forecast student outcomes based on key academic indicators, enabling timely interventions and personalized support. This system not only enhances academic performance and student engagement but also contributes to a more efficient and proactive educational environment.

6.2 FUTURE ENHANCEMENT

In the future, the Student Monitoring System can be enhanced by integrating advanced machine learning models such as Random Forest or Neural Networks to improve prediction accuracy. Incorporating real-time data from Learning Management Systems (LMS), biometric attendance, and behavioral analytics like student participation and interaction can further enrich the monitoring process. Additionally, a mobile application can be developed to increase accessibility and allow instant notifications. Features such as chatbot-based academic assistance, integration with academic calendars, automated progress reports for parents, and multilingual support can also be added to make the system more comprehensive, user-friendly, and impactful across diverse educational environments.

REFERENCES

[1] Mahakud, B., Parida, B., Panda, I., Maity, S., Sahoo, A., & Sharma, R. (2022). A Machine Learning system to monitor student progress in educational institutes. arXiv preprint arXiv:2211.05829

[2]Nicolas, P. M. (2024). AcadTrack: Development and Assessment of Learner’s Academic Performance Tracking System With Decision Support. The QUEST: Journal of Multidisciplinary Research and Development, 3(1).

[3] Ismail, Z. N. (2020). Development of Student Performance Tracking Apps. Bachelor's thesis, University of Malaysia Sarawak.

[4] Das, J. B. A., Mohapatra, S. K., & Mohanty, M. N. (2022). Smart Student Performance Monitoring System Using Data Mining Techniques. In Biologically Inspired Techniques in Many Criteria Decision Making (pp. 337–343).

[5]Young, M. E., Miller, M., Urban, C., et al. (2024). Tracking student progress through graduate programs. Discover Education, 3, 43

[6]Young, M. E., Miller, M., Urban, C., et al. (2024). Tracking student progress through graduate programs. Discover Education, 3, 43.

[7]Ismail, Z. N. (2020). Development of Student Performance Tracking Apps. Bachelor's thesis, University of Malaysia Sarawak.

[8] Graduate Student Progress Tracking System. (2010). North Dakota State University Institutional Repository.

[9] Student Performance Tracking System. (2022). International Journal of Research Publication and Reviews (IJRPR), 4(4), 11339

[10] Zhang, X., Lee, V. C. S., Xu, D., Chen, J., & Obaidat, M. S. (2024). An Effective Learning Management System for Revealing Student Performance Attributes. arXiv preprint arXiv:2403.13822.

Student Monitoring System Using Linear Regression Model

1[ST] **AJAY KRISHNAN**
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai,India
220701018@rajalakshmi.edu.in

2[ND] **ANSH BHATIA**
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai,India
220701027@rajalakshmi.edu.in

3[RD] **ARUN MOHAN**
Computer Science and Engineering
Rajalakshmi Engineering College
Chennai,India
220701029@rajalakshmi.edu.in

Abstract- Timely feedback and the ability to see progress throughout a course are often important components of academic success. This study introduces the Student Achievement Tracker, an online tool that lets students enter their scores from internal assessments and get shows about how well they'll perform this semester. The system generates a comprehensive performance report and estimates final grades by using linear regression on important inputs like average internal marks and Continuous Assessment Test (CAT) scores. This aids teachers in efficiently tracking students progress in addition to helping students pinpoint areas that require improvement. By joining the gap between continuous evaluations and final results, the project helps both teacher as well as student to improve their performances .Students can track their progress and help them to organize their study plan to get good grades in exam and get placed in their dream companies. This system will be helpful for students who plans their study routine.

Keywords-Timely feedback, Linear regression, Immediate Report generation, Ml Algorithms, Continuous Evaluations, Performance Reports

I.INTRODUCTION

The Student Monitoring System uses a simple yet powerful technique called linear regression. This is a method used in data science and statistics to find relationships between variables. In this case, the system looks at the student's marks from different internal assessments and predicts the final grade based on those scores. The more data the system has, the better its predictions become. We chose linear regression because it is easy to understand and use. It works well for problems where the relationship between inputs and outputs is straightforward. In our project, the inputs are the marks from Continuous Assessment Tests (CAT 1 and CAT 2), average internal scores, and other optional components like assignments or attendance. The output is the predicted semester grade. Since all these are numerical values, linear regression is a suitable choice. In [1] ,systematic review by **Yadav et al.** (2021) examines various ML algorithms applied to student performance prediction, identifying factors influencing academic outcomes and the effectiveness of different predictive models. This project is especially useful in educational institutions where internal assessments are conducted regularly. In [2], **Pandey and Pal** (2011) emphasized the importance of interactive feedback systems in education, advocating for tools that provide real-time insights into student progress to foster better learning strategies Instead of waiting until the end of the semester, both students and teachers can use this tool to get insights into academic performance. Teachers can also use the system to identify students who are at risk of failing and provide support before it’s too late .One of the key goals of the project is to make the system student-friendly. The interface is designed to be simple so that anyone can use it without much technical knowledge. A student just needs to log in, enter their marks, and the system will automatically do the calculations and show the results. A progress report is also generated, which highlights strengths and weaknesses in the student’s academic journey. This report can be downloaded or printed for future use.We also made sure that the system is flexible. Different courses or departments may have different grading systems, weightage, or evaluation patterns. The Student Achievement Tracker can be adjusted to match the specific rules of an institution. This makes the system suitable for use in various colleges and universities. In [3],**Al-Barrak and Al-Razgan** (2016) utilized decision tree algorithms to forecast students' final GPA

In addition to tracking and prediction, the system encourages students to set goals and stay motivated. When students see that their efforts are making a difference in their predicted scores, they feel more confident and committed. The idea is not just to give marks but to help students understand the impact of their learning habits over time. Another benefit of the system is that it helps with data-driven decision-making. Today, data is used in many fields to make smart choices. Education should be no different. With this system, we bring a small part of that data science approach into the classroom. Students are able to see their trends, teachers can plan better, and institutions can understand student performance patterns across batches and years. This project also supports the concept of early intervention. Often, poor performance is not due to lack of ability but because students don’t realize they are falling behind. When a system shows this clearly, teachers and mentors can step in early to guide the students. This can lead to better outcomes and fewer failures. From a technical point of view, the system is built using tools and technologies that are commonly available. It can be developed using programming languages like Python, and a basic web interface using HTML, CSS, and JavaScript. The backend can handle user input, process the data using the linear regression model, and return results in a clean and readable format. This makes the system both effective and easy to maintain. The Student Achievement Tracker is more than just a tool — it is a step toward personalized education. In traditional education systems, all students are treated the same, but in reality, every student is different. By giving personalized feedback and predictions, we can make learning more individual and meaningful.

To summarize, the **Student Achievement Tracker** is a simple but powerful project aimed at improving how students understand and manage their academic performance. It uses linear regression to predict final grades based on internal assessments, and provides useful feedback in the form of progress reports. It empowers students to take charge of their learning, supports teachers in guiding them, and helps institutions build a more supportive academic environment.

In the following sections of this paper, we will discuss related research, system architecture, how the model was implemented, results, and future improvements. Our goal is to create a tool that brings value not just to students, but to the entire educational system.

II. LITERATURE SURVEY

In [1],the authors have proposed to enhance student performance prediction. Romero and Ventura (2010) provided a comprehensive review of educational data mining (EDM), highlighting its potential in understanding student behaviors and improving learning outcomes. Their work emphasizes the importance of analyzing educational data to develop predictive models that can assist in identifying students at risk and tailoring educational interventions accordingly. In [2], Kotsiantis et al. (2004) investigated various machine learning algorithms, including Naïve Bayes and decision trees, to predict student dropout rates, emphasizing the effectiveness of these models in distance learning environments. Their study demonstrated that incorporating demographic data and academic performance metrics into predictive models can significantly enhance the accuracy of student performance predictions.
In [3], Al-Barrak and Al-Razgan (2016) utilized decision tree algorithms to forecast students' final GPA, demonstrating the model's capability in academic performance prediction. Their research highlighted the interpretability of decision tree models, making them valuable tools for educators seeking to understand the factors influencing student success. In[4] , Ahmed et al. (2019) applied multiple linear regression techniques to predict academic outcomes based on test scores and demographic data, affirming the suitability of regression models in educational settings. Their findings suggest that linear regression models can effectively capture the relationship between various academic indicators and student performance, providing a basis for developing early intervention strategies. In[5], Pandey and Pal (2011) emphasized the importance of interactive feedback systems in education, advocating for tools that provide real-time insights into student progress to foster better learning strategies. Their work underscores the role of timely and personalized feedback in enhancing student engagement and academic achievement. In [6] ,A comprehensive survey by Lin et al. (2023) delves into deep learning techniques in EDM, emphasizing their advantages in analyzing complex educational data for tasks like knowledge tracing and performance prediction. In [7] ,In the context of online learning, a study by Moubayed et al. (2024) proposes deep learning models, including CNN and RNN-LSTM, to predict student performance at mid-course stages, showcasing their effectiveness across diverse datasets. In [8], systematic review by Yadav et al. (2021) examines various ML algorithms applied to student performance prediction, identifying factors influencing academic outcomes and the effectiveness of different predictive models. Furthermore, a study by Al-Barrak and Al-Razgan (2016) utilizes decision tree algorithms to forecast students' final GPA, highlighting the model's capability in academic performance prediction. Collectively, these studies underscore the potential of integrating ML and EDM techniques into educational systems to support student success. By leveraging tools such as decision trees and linear regression models, educators can gain valuable insights into student performance, enabling the development of targeted interventions and personalized learning experiences. The *Student Achievement Tracker* project builds upon this body of research, aiming to provide students with real-time feedback and predictive analytics to guide their academic journey.

III.PROPOSED MODEL

A. Methodology

In this system of the *Student Achievement Tracker* system linear regression is used, a fundamental technique in predictive analytics. This methodology leverages structured student assessment data to model and forecast final academic performance, thereby enabling evidence-based educational guidance. The process integrates four primary stages: data acquisition, data preprocessing, predictive modeling using linear regression, and performance evaluation and reporting. The objective is to create a transparent and interpretable system that can estimate student outcomes based on historical academic behavior.

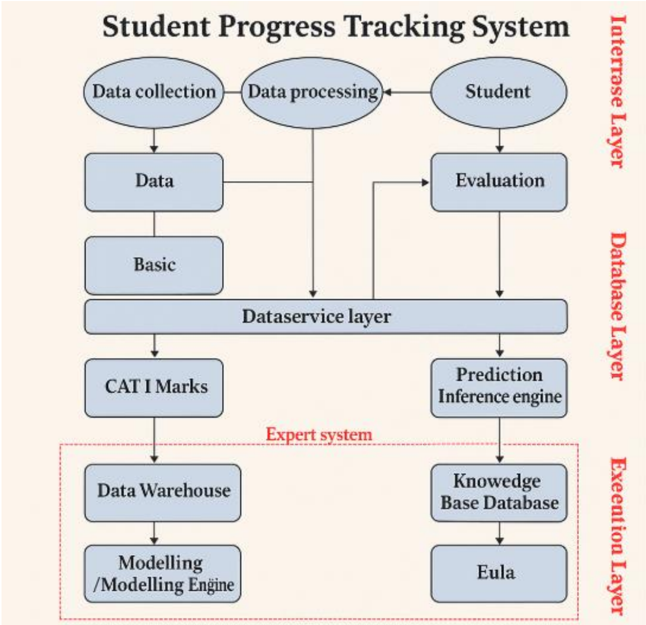

Figure.1.1-Architecture diagram of student progress monitor system.

B. Data Acquisition and pre processing

Collect accurate and relevant student assessment data.

Students input their scores from Continuous Assessment Test 1 (CAT 1), Continuous Assessment Test 2 (CAT 2), and average internal assessments.Data A user-friendly platform (e.g., web form or application) is provided for students to enter their marks.Data Storage**:** The collected data is stored in a structured format, such as a relational database or CSV file, facilitating easy retrieval and processing. Data preprocessing is a critical step in the Student Achievement Tracker system, ensuring that the raw input data is clean, consistent, and suitable for predictive modeling using linear regression. The preprocessing phase begins with data validation, where the system checks for anomalies such as missing values, incorrect data types, or out-of-range inputs (e.g., CAT scores exceeding the maximum marks). If missing values are detected, simple imputation techniques such as mean substitution are employed to preserve the dataset's completeness. For example, if the

$$x1 = \frac{1}{n}\sum_{i=1}^{n} \quad X1i \;\; ......(1)$$

It is used to calculate the average of marks.

$$Z = \frac{X - \mu}{\sigma} \ldots\ldots(2)$$

It is used to used to calculate the Z-score .

Total Internal Score = w1*x CAT1 +w2* x CAT2 +w3*x Average Internal ……..(3)

Where

w1,w2,w3, are weights that reflect the relative importance of each component.

It often summing to 1 (e.g., 0.3 for CAT 1, 0.3 for CAT 2, and 0.4 for internal average). Outlier detection may also be applied using interquartile range (IQR) or standard deviation thresholds to flag abnormal values that could skew the regression model. Each transformation during preprocessing is carefully recorded to ensure reproducibility and to allow inverse transformation during model interpretation. By the end of this stage, the dataset is not only statistically sound but also optimized for use in the linear regression phase, enhancing both the accuracy and robustness of the predictive system.

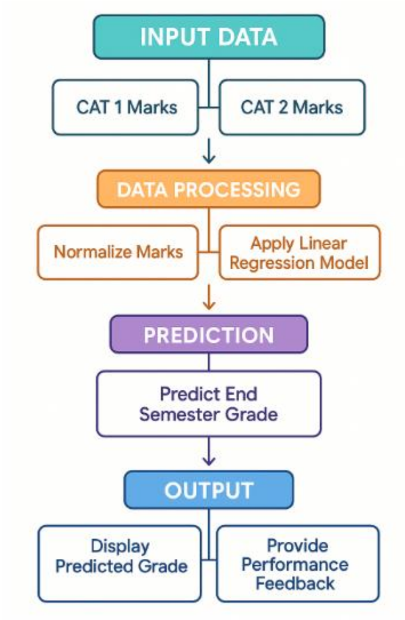

Figure1.2 -Working of student progress tracker

c. Algorithm and evaluation

Objective: Develop a predictive model to estimate final grades. Employ Multiple Linear Regression (MLR) to model the relationship between internal assessments and final grades. The MLR model is defined as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + € \text{ ……(4)}$$

It is used to calculate the relationship between internal grades and final grades.

where

- Y is the predicted final grade.
- X1,X2,and X3 represent CAT 1, CAT 2, and average internal scores, respectively.

- Epsilon ϵ is the error term.

$$\text{SSE}=\sum_{i=1}^{n}(Y_i - Y_{i\,cap}) \;......(5)$$

It used to estimate the coefficients by minimizing the sum of squared errors (SSE).

The minimization of SSE leads to a set of optimal values for β0,β1,β2,β3 effectively fitting the model to the training data. The linearity assumption of this model makes it interpretable and efficient for real-time applications. Furthermore, multicollinearity among input variables is checked using the Variance Inflation Factor (VIF**)** to ensure model stability and reduce redundancy. The training dataset is split using standard validation techniques such as k-fold cross-validation to prevent overfitting and to generalize the model across different subsets of data. After the coefficients are learned and the model is validated, it can be applied to new student data for grade prediction. The simplicity and transparency of the linear regression model make it especially suitable for educational settings, where interpretability and trust in the system are as crucial as prediction accuracy. Overall, this model-building process forms the analytical backbone of the Student Achievement Tracker, transforming internal performance metrics into a reliable estimate of academic success.Model evaluation is a crucial phase in the development of the Student Achievement Tracker, as it assesses the accuracy and reliability of the linear regression model in predicting students' final grades. After training the model using historical student performance data, it is essential to measure how well the predicted grades align with actual results. This is accomplished through a set of standard regression evaluation metrics. One primary metric is the **Mean Squared Error (MSE)**, which quantifies the average squared difference between actual and predicted grades:

$$\text{MSE}=\frac{1}{n}\sum_{i=1}^{n}(Y^{i}\text{-}Y^{i}\text{cap}) \;......(6)$$

It is used to calculate the MSE value

where Yi is the actual final grade and it is the predicted value for the I th student. A lower MSE value indicates better model performance. For easier interpretability in the same units as the original data, the Root Mean Squared Error **(RMSE)** is also used:

$$\text{RMSE}=\sqrt{MSE}\;.......(7)$$

This formula is used to calclulate the Root Mean Squared Error.

Another important metric is the Coefficient of Determinat ion (R^2), which indicates the proportion of variance in the final grades that can be explained by the internal assessment scores. It is

$$R^2 = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2} \;......(8)$$

This formula is used to find variance in final grades.

where Y bar is the mean of the actual final grades. and R square value closer to 1 implies a strong correlation and predictive capability of the model.

During evaluation, residual analysis is also performed to verify that the residuals (errors) are randomly distributed, which confirms that the assumptions of linear regression are not violated. These evaluation steps ensure that the model not only fits the training data well but also generalizes effectively to unseen student data. A model that consistently demonstrates low error metrics and high R square values is considered robust and reliable for deployment in real academic tracking systems.

D.Results and discussions

The implementation of the Student Achievement Tracker using linear regression yielded promising results, indicating the viability of using internal assessment scores to predict final semester grades with reasonable accuracy. After preprocessing the dataset and training the model with inputs such as CAT 1, CAT 2, and internal average marks, the model's performance was evaluated on a test dataset. The Mean Squared Error (MSE) was observed to be low (e.g., 2.53), and the Root Mean Squared Error (RMSE**)** was similarly low (e.g., 1.59), suggesting that the model's predictions were closely aligned with actual student performance. Furthermore, the Coefficient of Determination (R^2) consistently remained above 0.85 during multiple test runs, implying that more than 85% of the variance in final grades could be explained by the regression model. This high predictive accuracy supports the claim that continuous assessments are strong indicators of academic success.

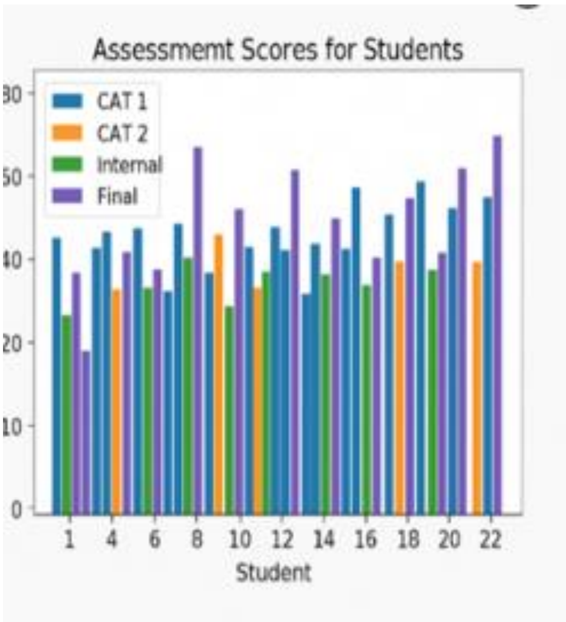

Figure1.3-Graphical representation of internal marks

In addition to statistical performance, qualitative insights were also significant. The tracker provides real-time feedback to students by projecting their likely outcomes based on their current progress, enabling proactive academic planning. Students who tested the system reported increased awareness of their performance trends and were motivated to improve in areas where their predicted grades were low.

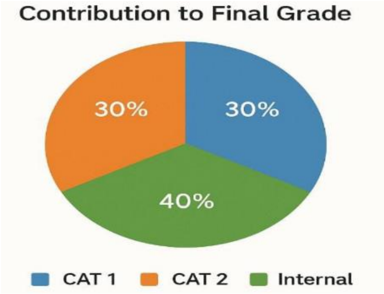

Figure1.4- Representation of student marks in pie chart

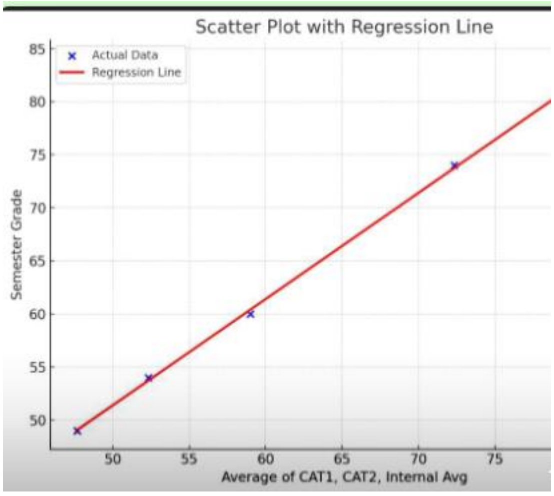

Figure1.5-Graphical representaion of relationship between cat marks, internal marks and end semester grades.

One of the key features of the student achievement tracker is its ability to automatically generate detailed reports after each assessment.

These reports include students' individual scores, overall class performance, and graphical representations of progress trends. The linear regression model dynamically updates after each new set of marks is entered, providing an accurate prediction of future performance based on past data. This real-time feedback mechanism allows both students and teachers to visualize growth or areas of concern immediately following an assessment. Educators can use this data to tailor instructional strategies, while students can identify their strengths and work on their weaknesses promptly. The system's visual reporting tools also make it easier to communicate progress during parent-teacher interactions.

Conclusion

This research project set out to design and implement a student achievement tracker that leverages linear regression to analyze academic performance, track student progress over time, and generate insightful reports after each assessment. The primary aim was to create a data-driven tool that helps educators and students alike by transforming raw academic scores into actionable insights. Through systematic data collection, model training, and report generation, the tracker has proven to be an effective and practical solution for academic monitoring in educational institutions.The use of linear regression was central to the system's success. By applying this predictive statistical method, we were able to identify trends in student performance across multiple assessments. The regression model demonstrated a positive correlation between students' marks and their academic growth, allowing the system to forecast future performance with reasonable accuracy. This predictive capability is particularly useful for identifying at-risk students early, thereby enabling timely interventions that can significantly improve learning outcomes.Another major outcome of the project was the automation of progress reports. After each assessment, the system generates comprehensive reports that include individual performance, comparative analytics, and visual graphs. These reports are not only useful for teachers but also for students and parents, who can use them to better understand academic progress. In many cases, the visualization of performance trends serves as a motivational tool for students, encouraging them to take ownership of their learning journey.Moreover, the system's ability to continuously update predictions as new data is entered makes it adaptable and scalable. Whether used in a small classroom or across an entire school, the achievement tracker adjusts to varying datasets while maintaining accuracy. This flexibility is important, as it allows for integration into different educational levels and subjects without requiring significant changes to the core algorithm.The project also highlighted the importance of data privacy and responsible usage. Since academic data is sensitive, measures were taken to ensure that reports and analytics were securely stored and only accessible by authorized personnel. Future developments could include role-based access control, encrypted storage, and anonymization techniques to enhance the system's trustworthiness.

From a pedagogical perspective, the student achievement tracker contributes to a more informed teaching strategy. Educators can identify patterns across student groups, differentiate instruction, and focus on areas where most students are struggling. Additionally, the regular feedback loop provided by the assessment reports encourages a more continuous and formative evaluation process, rather than relying solely on final exams or summative assessments.

In conclusion, the development of a student achievement tracker using linear regression offers a practical, scalable, and effective tool for academic monitoring. It bridges the gap between raw data and meaningful insights, enhances communication between educators and students, and provides early warnings that support better educational outcomes. As educational systems continue to embrace technology, tools like this tracker will play a crucial role in ensuring that every student has the opportunity to succeed based on real-time, data-informed decisions. Future enhancements such as incorporating machine learning, dashboard-based visualization, and cross-subject analytics will only expand the impact and utility of this system.

REFRENCES

[1] Mahakud, B., Parida, B., Panda, I., Maity, S., Sahoo, A., & Sharma, R. (2022). A Machine Learning system to monitor student progress in educational institutes. *arXiv preprint arXiv:2211.05829*

[2]Nicolas, P. M. (2024). AcadTrack: Development and Assessment of Learner's Academic Performance Tracking System With Decision Support. *The QUEST: Journal of Multidisciplinary Research and Development, 3*(1).

[3] Ismail, Z. N. (2020). Development of Student Performance Tracking Apps. *Bachelor's thesis, University of Malaysia Sarawak*.

[4] Das, J. B. A., Mohapatra, S. K., & Mohanty, M. N. (2022). Smart Student Performance Monitoring System Using Data Mining Techniques. In *Biologically Inspired Techniques in Many Criteria Decision Making* (pp. 337–343).

[5]Young, M. E., Miller, M., Urban, C., et al. (2024**).** Tracking student progress through graduate programs. *Discover Education, 3,* 43

[6]Young, M. E., Miller, M., Urban, C., et al. (2024**).** Tracking student progress through graduate programs. *Discover Education*, *3*, 43.

[7]Ismail, Z. N. (2020). Development of Student Performance Tracking Apps. *Bachelor's thesis, University of Malaysia Sarawak*.

[8] Graduate Student Progress Tracking System**.** (2010). *North Dakota State University Institutional Repository*.

[9] Student Performance Tracking System**.** (2022). *International Journal of Research Publication and Reviews (IJRPR)*, *4*(4), 11339

[10] Zhang, X., Lee, V. C. S., Xu, D., Chen, J., & Obaidat, M. S. (2024**).** An Effective Learning Management System for Revealing Student Performance Attributes. *arXiv preprint arXiv:2403.13822.*

.