STUDENT PERFORMANCE TRACKING AND VISUALISING

Keerthana C (Roll no: 240701253), Keerthana G (Roll no: 240701254), Kiruthika S (Roll no: 240701265), Krithika L (240701277)

Department of Computer Science and Engineering

Rajalakshmi Engineering College, Chennai, India

keerthana.c.2024.cse@rajalakshmi.edu.in, keerthana.g.2024.cse@rajalakshmi.edu.in, kiruthika.s.2024.cse@rajalakshmi.edu.in, krithika.l.2024.cse@rajalakshmi.edu.in

***Abstract*— Monitoring one's own academic performance is a crucial part. With traditional methods, it is cumbersome and time-consuming to keep track of all factors such as marks, attendance, and performance in all courses. To improve the process of tracking one's own performance, it is necessary to have a system which takes various inputs and generates the desired output in an interactive way where the marks, attendance can be visually presented. The proposed system "Student Performance Tracking and Visualizing" helps students to trace their academic journey in a visual format, and it also provides them the ability to predict future GPA and CGPA based on their past performance. The system was built using Python and data science techniques, utilizing the vast libraries of the Python ecosystem. Its main target is to provide an intuitive dashboard and use simple models like linear regression, logistics regression, polynomial regression to predict the future outcome and also enables students to decide on what if scenarios.**

Keywords—Student performance analysis, Academic progress tracking, Data Visualization, Educational data mining, Predictive analytics

I. INTRODUCTION

In today's society, focusing on academic performance has become an essential part. It's a critical indicator of a student's performance and it's vital for predicting their future outcomes. Tracking academic performance involves considering multiple factors such as internal assignment marks, internal examinations, external examinations, projects, attendance, participation in extra other activities, and the consistency of performance across the semesters. To do this, a traditional method needs manual records where the data is spread across various spreadsheets that are fragmented. This process involves lots of effort and is time-consuming. The method is also most likely to be error prone, and manual efforts give limited insights. The main focus is to find the hidden pattern in the provided records and produce them in an interactive way which attracts students and also reduces their time in analysing.

There is a strong demand for grade analysing tools as they save time, identify the learning gaps, improvised assessment and engagement, and also allows personalized mentoring. This demand is proven by the statistical evidence segmented into financial valuation of educational market, time-savings experienced by teachers and the proven impact of analytical techniques. Table I shows the substantial market commitment and institutional adoption rates of technologies, while Table II provides specific quantitative metrics depicting the reduction in administrative workload, at last Table III highlights the positive correlation between the grade analysis and improvements in core student outcomes.

TABLE I. Market Demand and Sector Growth for Educational Analytics

Metric	**Value**	**Source Type**
Global EdTech Market Size (2024)	USD 171.2 Billion [1]	Baseline valuation of the industry supporting analytics platform
Projected EdTech Market CAGR (2024-2033)	14.1% [1]	High CAGR indicates strong, sustained investment in digital learning and data solutions.
K-12 Sector Revenue Share (2024)	39.40% [2]	It is the largest market segment, demonstrating susbstantial demand for technology based grading and tracking systems.
Projected Global Education ERP Market Size (2030)	USD 61.23 Billion [3]	Enterprise Resource Planning (ERP) systems in education integrate student information and grading analytics
Institutional Adoption of LMS	75% of K-12 institutions globally expected to adopt Learning Management Systems(LMS) by 2025 [1]	LMS platforms are the primary host for grade analysis and performance reporting features.

TABLE II. Educator efficiency and time savings automation

Metric	Value	Source Type
Reduction in Weekly Grading Time	35% [4]	Empirical evidence showing the average reduction in teacher grading hours
Automated Grading Coverage	Median coverage of 92.2% of submitted student work [5]	A high median rate of automated score generation was reported, demonstrating the system's capacity to streamline evaluation at scale.

TABLE III. Pedagogical Impact and student performance metrics

Metric	Value	Source Type
Student performance prediction accuracy	88.3% [6]	Demonstrated accuracy of a random forest classification model in predicting successful students using educational data mining techniques
Student assignment completion rate increase	Increase from 70% to 88% [4]	Increased rate directly correlates with the provision of timely and consistent automated feedback from analysis systems [4].
Focus of Learning Analytics Research	Approximately 70% of research investigated student performance prediction using grades/ GPAs [7]	Existing grade data is the predominant and most valuable feature for predictive modelling

II. Literature survey

There are many works and plenty of papers about the grade analyzing and predicting the performance of students. But it's found that most of the studies did not use the visualization approaches to analyze the academic data.[8]

A. The Importance and Evolution of Digital Grading Systems

As already stated there is a strong demand and need for these grade analysing tools. Early research has provided the tools to enhance the learning process by predicting and content visualization.

- Addresses the Educator workload
- Enhances fairness and consistency

Grade analysis is the backbone for the implementation of Learning Analytics and Educational Data Mining. The analysis converts the raw data into useful insights which makes understanding the pattern and the progress easier.

B. The Role of Predictive Modelling

There are several studies which have applied machine learning algorithms to predict student performance. Albreiki et al. [9] conducted literature review on student performance prediction using machine learning techniques, emphasizing the importance of the models in educational data mining. Similarly, Martinez and Miller [10] used logistic regression for determining how likely a student is to graduate after four years in college with a high degree of accuracy. In the same way, Fernandes et al. [11] performed descriptive statistical analysis, and Gutierrez et al [12] used random forest.

C. Data Preprocessing and Feature Selection

For building robust predictive models, data preprocessing and feature selection are most important. Jenhani et al.[13] proposed a remedial action classification-based approach which highlights the role of data preprocessing in model performance.

TABLE IV. List of Research Papers referenced

Paper Title and Year Published	Algorithm / Technology used	Result obtained	Matrix used to analyse the result	Future enhancement
Academic Performance Analysis Supported by a Web-Based Visual Analytics Tool (2020)	Random Forest, Linear Regression, Decision Tree, MLP, D3.js, Scikit-learn	Achieved effective visualization of academic factors	Confusion matrix, accuracy, F1-score	Include more diverse datasets and behavioral features for higher prediction accuracy
Explainable Student Performance Prediction Models: A Systematic Review (2021)	XAI, Decision Trees, Logistic Regression	Identified best-performing interpretable models for academic prediction	Accuracy, Precision	Combine explainable models with visualization platforms
A Comparative Analysis for GPA Prediction of Undergraduate Students Using Machine and Deep Learning (2024)	SVM, Random Forest, ANN	Early-semester data effectively predicted final GPA	Confusion Matrix, Accuracy	Apply deep learning for continuous GPA prediction
Early Prediction of Students's Performance Using Machine Learning Techniques (2014)	Decision Trees, Naïve Bayes, KNN	Early identification of low-performing students	Confusion Matrix	Extend to real-time tracking and visualization

Course Learning Outcome Performance Improvement: A Remedial Action Classification-Based Approach (2016)	Classification Algorithms	Enhanced outcome prediction through remedial action plans	Confusion Matrix	Apply method to semester-level academic analytics
Student-Facing Learning Analytics Dashboard for Remote Labs (2022)	Web Dashboard, Python, SQL	Enabled real-time monitoring of lab performance	Usability and Accuracy Matrices	Integrate AI-based personalized insights
Data Mining Techniques for Academic Performance Prediction (2018)	Decision Tree, SVM, KNN	Achieved 86% accuracy on academic dataset	Confusion Matrix	Apply deep learning models on larger datasets
Predictive Analytics in Academic Performance Monitoring (2020)	Linear Regression, Random Forest	Predicted semester GPA trends effectively	R^2, RMSE	Expand model to include attendance and behavior features
Educational Dashboards for Monitoring Student Engagement (2019)	Dashboard Frameworks, Data Analytics	Improved engagement through visual tracking	Usability and Feedback Metrics	Combine with predictive performance indicators
Machine learning analysis and inference of student performance and visualization of data results based on small data set of student information (2021)	Pandas, Numpy, Regression algorithms	One hot encoding result did have an effect compared to all elements	Mean Absolute Error and Root Mean Square Error (RMSE)	Larger student information datasets, to add videos or pictures
Student Performance patterns in Engineering at the university of Johannesburg: An Exploratory Data Analysis (2023)	Pandas, Seaborn,Matplotlib, Missingo, EDA technique	Gender disparity, Dropout reason	Correlation Coefficient / Heat Map	Include non academic factors, Machine learning algorithms
Attendance Tracking Using Smartphone GPS and NFC Technologies	GPS, NFC, Android Framework	Increased attendance accuracy through mobile tracking	Precision, Recall	Combine with academic data for performance correlation
Visual Analytics for Student Assessment and Feedback	D3.js, Tableau	Improved data-driven decision-making in education	Usability Metrics	Incorporate ML predictions into dashboards
Data-Driven Student Performance Tracker with Alerts	Python, Pandas, Threshold Algorithms	Implemented automatic alerts for low performance	Confusion Matrix	Add adaptive learning recommendation engine
Student Attendance Management Systems: A Systematic Literature Review and New Taxonomy	RFID, Biometric Systems	Enhanced automation and reliability of attendance systems	Not applicable (review paper)	Integrate attendance data with performance analytics
Academic Progress Tracking Dashboard for Students	Python, Streamlit, Plotly	Developed interactive dashboards for student progress visualization	User Testing Matrix (Usability Evaluation)	Add predictive analytics module for CGPA goals
Predicting students final academic performance using feature selection approaches	Classification (machine learning)	Feature selection improves accuracy	F1-score	Generalisation and scalability
Predictive Model of Undergraduate Student Grading Using Machine Learning for Learning Analytics (2022)	Bayes Network, Naive Bayes, Simple logistics, JRip, OneR,SMO and Random forest	High accuracy with JRip and random forest	Accuracy and confusion matrix	Expanding scope and curriculum impact
FICAvis:Data Visualization to prevent University Dropout (2020)	Power BI, user-centred design	Developed an interactive dashboard for stakeholders	Analytical and empirical evaluation methods	Integration into university services and security layer

III. Proposed methodology

From the existing systems, it is found that several significant gaps are present in the current state of

educational data mining and learning analytics. There are limitations in methodology, challenges in generalization and deployment, and limited data.

The proposed system works in the following manner.

A. Addressing data and feature gaps

The application addresses the data limitation by taking input of various factors per course that goes beyond the Standard Cumulative GPA.

- Course Name, Credits, Marks obtained, Attendance percentage are collected from students for past semesters. This provides the rich data set required for the analysis.
- The internal calculation has a predefined dictionary that maps the categorical inputs to numerical values, effectively performing the data transformation.

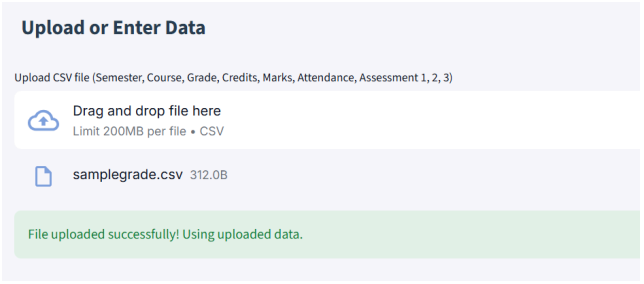

Fig I. Uploading input data

B. Predictive Modelling and Evaluation

The application uses linear regression,logistics regression, k-nearest neighbours to analyse the gpa trend across the completed semesters.

- This allows the student to predict the gpa of next semester at an early stage and prepare accordingly.
- It also visualizes the predicted gpa trend, making it easily understandable by the students.

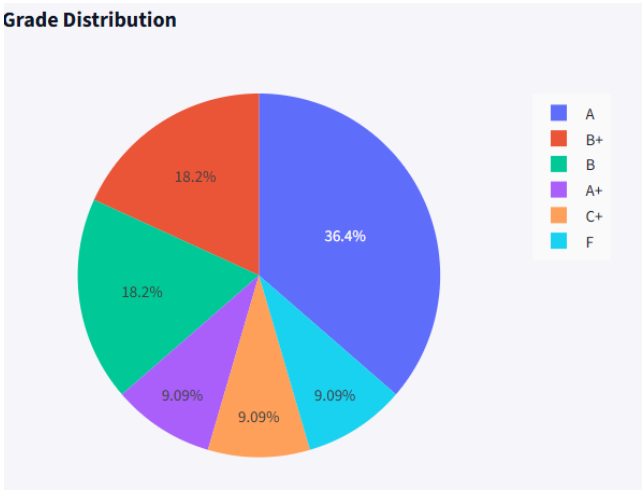

Fig II. Pie chart showing grade distribution

Fig III. Correlation heat map

Fig IV. GPA prediction comparison

C. Target Simulation

- Students can provide the target cgpa and the system calculates the minimum average grade point needed.
- Scenario simulation, it allows students to immediately test the impact on their overall cgpa by providing any specific grades of future courses.

Fig V Target setting

D. Visualization and Insights

- The system replaces the static tables with interactive plots using plotly. The gpa trend is depicted through line charts, marks and attendance are shown through bar charts, radar

charts and heatmaps are used for highlighting the students' performance in each subject.
- Early alert system is also implemented that flags the courses with below specified thresholds.

E. User-Centric Design

The application is built using streamlit which provides an intuitive user interface which does not require specialized software.

- Easily accessible through a web browser
- The application is very simple and clear that separates the input, calculation and visualization and prediction into labeled sections.

F. Architecture Diagram

- Input data: it gets a dataset which can be either cleaned or with missing values.
- Preprocessing: These are then preprocessed and produce the cleaned data for further analysis.
- Analysis: The system uses various machine learning algorithms like linear regression,logistics regression,polynomial regression, k-nearest neighbours and performs eda analysis that gives in depth analysis of the student performance.
- Visualization: the analysis is later represented using line charts,bar charts, pie charts, etc that makes the understanding better.
- The whole system uses a very simple methodology which is depicted in the below diagram.

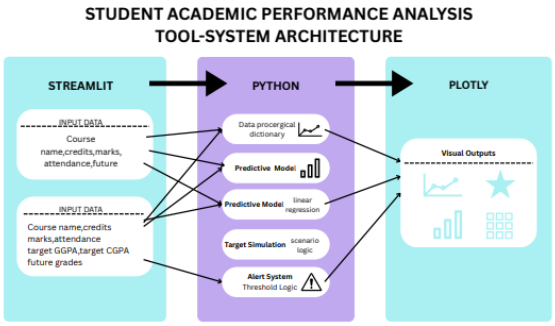

Fig VI. Architecture Diagram

IV. Conclusion

The proposed system Student Performance tracking and Visualising implements the core features required after listing the limitations of existing methods. The application uses user-centric design which is interactive to use, it incorporates the data like marks, credit, grades and attendance across completed semesters. It also uses linear regression, logistics regression,k-nearest neighbours,t test,z test and f test which acts as a basic predictive gpa trend, the target cgpa simulation enables immediate response to the scenario provided.

V. Future enhancements

The proposed system can be expanded by adding a variety of features, first converting it to a mobile application, then connecting with a database for real time analysis in college portal systems, and implementing authentication systems. If possible integration of chatbot would make the whole portal interactive and complete academic system.

References

1] DMR, "Global EdTech Market to Reach USD 561.8 bn by 2033," Dimension Market Research, [Online]. Available: [Insert relevant URL, e.g., market research firm website]. (Accessed: October 2025).

2] Grand View Research, "Education Technology Market Size, Share & Trends Analysis Report," [Online]. Available: [Insert relevant URL, e.g., market research firm website]. (Accessed: October 2025).

3] Verified Market Research, "Education Enterprise Resource Planning (ERP) Market Size," [Online]. Available: [Insert relevant URL, e.g., market research firm website]. (Accessed: October 2025).

4] S. H. Al-Mekhlafi, "Automated Grading and Feedback Systems: Reducing Teacher Workload and Improving Student Performance," Journal of Asian Development Studies, vol. 13, no. 4, pp. 248-261, 2024.

5] G. L. Yang and A. E. K. M. D. E. M. S. S., "Implementation Considerations for Automated Al Grading of Student Work," arXiv preprint arXiv:2506.07955, Jun. 2025.

6] N. S. Patel and S. R. Varma, "Al and Teacher Productivity: A Quantitative Analysis of Time- Saving and Workload Reduction in Education," ResearchGate, [Online]. Available: [Insert relevant URL, e.g., research repository link]. (Accessed: October 2025).

7] B. Y. M. S. V. T. B., "Predicting Student Performance in Higher Educational Institutions Using Video Learning Analytics and Data Mining Techniques," Applied Sciences, vol. 10.

8] R. Etemadpour, Y. Zhu, Q. Zhao, Y. Hu, B. Chen, M. A. Sharier, S. Zheng, and J. G. S. Paiva, "Academic Performance Analysis Supported by a Web-Based Visual Analytics Tool," Proc. IEEE Conf. on Visualization and Computer Graphics, City College of New York, USA, 2020

9] B. Albreiki, N. Zaki, and H. Alashwal, "A Systematic Literature Review of Student' Performance Prediction Using Machine Learning Techniques," Education Sciences, vol. 11, no. 9, p. 552, 2021.

10] J. Martinez and A. Miller. Analysis of student success. Mission College Undergraduate Research Journal (MURJ), 2015

11] E. Fernandes, M. Holanda, M. Victorino, V. Borges, R. Carvalho, and G. Van Erven. Educational data mining: Predictive analysis of academic performance of public school students in the capital of brazil. Journal of Business Research, 94:335–343, 2019

12]. L. Gutiérrez, V. Flores, B. Keith, and A. Quelopana. Using the belbin method and models for predicting the academic performance of engineering students. Computer Applications in Engineering Education

13] I. Jenhani, G. Brahim, and A. Elhassan, "Course Learning Outcome Performance Improvement: A Remedial Action Classification-Based Approach," in Proc. 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 2016, pp. 408-413