Student Performance Analysis For College Student

Ritik Malviya[1], Ankush Gupta[2], Sohail Ansari[3], Ruchita Kumbhare[4], Ast.Prof. Ravi Asati[5]
[1] Department of Emerging Technologies [AI&ML and AI&DS], SBJITMR, Nagpur, India

Email:ritikm.aiml21@sbjit.edu.in[1],ankushg.aiml21@sbjit.edu.in[2],sohaila.aiml21@sbjit.edu.in[3],ruchitak.aiml21@sbjit..edu.in[4],raviasati@sbjit.edu.in[5]

Abstract—This research paper focuses on conducting a comprehensive analysis of student performance within a specific department of a college. The paper entails collecting historical data on student attendance and examination marks, including Continuous Assessment Examinations (CAE I & II) and Teacher Assessment Examinations (TAE I, II & III). Following data collection, cleaning and preprocessing, a dashboard will be designed to visualize individual student performance metrics. Both teachers and students will have access to the platform, where teachers can input student data and view dashboards, while students can update their profiles, monitor their performance and also access student achievements and certificates. Additionally, a machine learning model will be developed to predict student performance based on current academic data and past exam results and this will be our future scope. The paper aims to facilitate effective communication between students and teachers, enable data-driven decision-making, and enhance student performance through personalized insights and interventions.

Keywords: Novel Student performance analysis, Data collection, Data cleaning, Data preprocessing, Visualization, Machine learning model, Predictive analytics, Performance prediction, Achievement tracking, Intervention strategies.

I. INTRODUCTION

In today's educational landscape, analyzing student performance has become increasingly important for academic institutions. Our project aims to explore student performance within a specific department of our college, leveraging data-driven insights and advanced technologies to provide a comprehensive understanding of student progress and avenues for improvement. [1] The project begins with collecting comprehensive data on student attendance and performance in Continuous Assessment Examinations (CAE I & II) and Teacher Assessment Examinations (TAE I, II & III). This data is essential for understanding students' grasp of course material and their ability to apply knowledge in different contexts. The collected data also includes internal marks derived from CAEs and TAEs, providing a holistic view of students' academic progress.

Once the data is collected, it undergoes rigorous cleaning and preprocessing to ensure accuracy and reliability. Data cleaning involves identifying and rectifying errors, inconsistencies and missing values, while preprocessing tasks such as normalization and feature engineering prepare the data for analysis.

With the cleaned and processed data, we proceed to develop a user-friendly dashboard for visualizing student performance metrics. This dashboard serves as a visual representation of attendance trends, exam scores, and internal marks, allowing teachers and students to monitor progress and identify areas for improvement.

Fig Process Of Dashboard Designing

In addition to dashboard development, we can explore the use of machine learning techniques to predict student's performance. By analyzing past performance data and identifying patterns, we aim to develop predictive models that can forecast how students are likely to perform in upcoming exams. These predictive insights can help students better prepare for exams and enable educators to implement proactive interventions to support student success.[2]

Furthermore, the project includes the development of a web-based platform to facilitate communication and data exchange between students and teachers which means, students can upload their certificate and achievements and teachers can upload student data, and this platform provides secure login credentials for educators to input student data, access performance dashboards, and engage in meaningful dialogue with their students. Similarly, students are empowered to update their profiles, track their progress, and access valuable resources to support their academic endeavors.

In essence, our research paper represents a synergistic blend of data analysis, technology integration, and educational innovation. By harnessing the power of data and technology, we aim to empower students and teachers, foster a culture of academic excellence, and drive positive change within our educational community.[3]

II. LITERATURE SURVEY

The comprehensive analysis of student performance in academic settings has garnered significant attention in recent years due to its implications for educational outcomes and student success. Researchers have explored various aspects of student performance, including attendance, examination scores, teacher assessments, and internal marks, to understand the factors influencing academic achievement. The topic of student performance analysis and educational data analytics provides valuable insights into methodologies and techniques employed in similar projects. In this review, we draw upon existing research to contextualize our approach and highlight key considerations in the analysis of student performance data for college student.[4]

Data Collection and Preprocessing:
Previous studies emphasize the importance of collecting accurate and comprehensive data to facilitate meaningful analysis of student performance. Our project aligns with this recommendation by collecting data on student attendance, CAE I & II marks, TAE I, II & III marks, and internal marks calculated from CAEs and TAEs. The emphasis on data cleaning and then, depending on the results of comparisons, it will provide projections. This is its primary aim, and preprocessing resonates with best practices in educational data analytics, ensuring the accuracy and reliability of our dataset.[9]

Dashboard Design and Visualization:
The design of dashboards for visualizing student performance metrics is a prominent term in educational data analytics literature. Our project aligns with this trend by developing a user-friendly dashboard that provides teachers and students with intuitive access to performance metrics. By visualizing attendance records, exam scores, and internal marks, our dashboard facilitates data-driven decision-making and promotes student engagement in their academic journey.[10]

Machine Learning-Based Predictive Modelling:
The incorporation of machine learning techniques for predictive modelling is a common approach in educational data analytics research. Our project follows suit by exploring the use of predictive models to forecast future student performance based on past data. By leveraging techniques such as linear regression and correlation analysis, we aim to identify predictors of academic success and provide personalized recommendations for academic improvement but at present we are not focusing on machine learning models. This will be a part of our future scope.[6]

Web Development and Platform Integration:
The integration of web development technologies into educational platforms is another emerging trend in educational data analytics. Our project leverages web development tools to create a platform that seamlessly connects students and teachers, facilitating data exchange and communication. By providing teachers with login credentials to input student data and access performance dashboards, and students with the ability to update profiles and view achievements, our platform enhances collaboration and engagement in the educational process.[11]

Overall, our paper draws upon the literature on student performance analysis and educational data analytics to inform our approach and methodology. By integrating best practices in data collection, preprocessing, visualization, predictive modelling, and platform development, we aim to contribute to the growing body of knowledge in this field and provide practical insights for improving student outcomes within a college department. [4]

III. METHODOLOGY

Our research methodology is meticulously crafted to harness the potential of the provided data and unearth invaluable insights into student performance while laying the groundwork for the development of predictive models. This methodology unfolds through a series of well-planned steps to address specific research objectives and contribute to a complete understanding of factors influencing student success. The initial phase revolves around data collection and preparation, wherein we gather historical records on student attendance, CAE I & II marks, TAE I, II & III marks, and internal marks derived from CAEs and TAEs. This is followed by a comprehensive data cleaning and preprocessing phase

aimed at ensuring the accuracy and consistency of the dataset by addressing issues such as missing values and outliers.

Subsequently, we delve into feature selection and engineering, identifying relevant features from the dataset that may influence student performance. Through feature engineering, we aim to create new variables or transform existing ones to enhance predictive performance, laying the groundwork for robust predictive modelling. Our focus is on visualization of insights gleaned from our analysis, designing a dynamic dashboard using Power BI to visualize student performance metrics and predictive insights. This dashboard will incorporate interactive features to empower users to explore performance trends, set goals, and access personalized recommendations. It includes visualizations of monthly attendance and overall attendance of a particular semester for a particular subject, a complete analysis of CAE I and II which includes total marks scored in subjective paper, total marks scored in object paper, and total marks scored in that particular subject.

Furthermore, we have represented score vs question comparison in our dashboard. Additionally, we have analysed CAE I and II papers in which we have mapped teachers' teaching plan vs attendance vs questions asked in the CAE exams. This helps students to analyse their performance and also helps them to understand their topic clarity and understanding. This will definitely result in the improvement of performance in an individual student.

In our project, we utilize various visualization techniques to present insights and findings derived from our research methodology. These visualizations serve to enhance understanding and facilitate decision-making processes.[15]

Here's an explanation of the charts we employ:

1. Pie Charts: Pie charts are used to represent the overall attendance of students. This visualization provides a quick and intuitive overview of attendance distribution among students, allowing stakeholders to identify patterns and trends at a glance.[12]

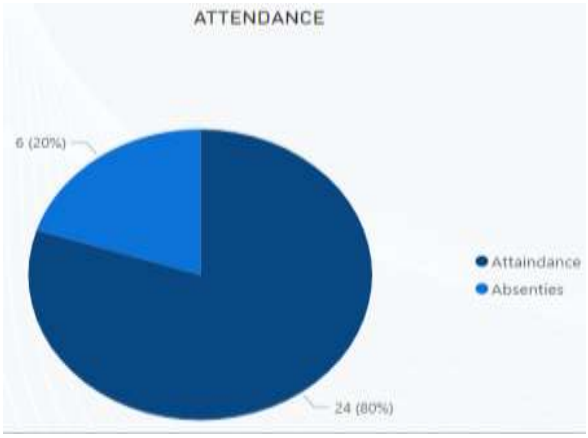

Fig 2: Pie Chart (For Overall Attendance)

2. Column Charts: Column charts are utilized to depict monthly attendance and CAE I & CAE II marks. By displaying attendance and marks data over time, stakeholders can observe fluctuations and identify any recurring patterns or anomalies throughout the session.[13]

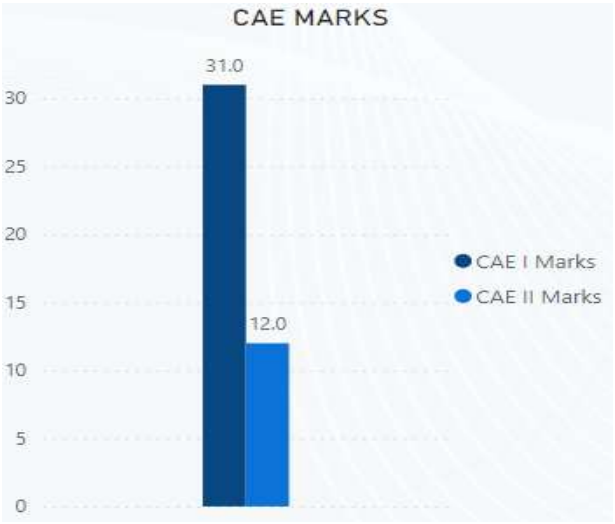

Fig 3: Column Chart (For CAE Marks and Attendance)

3. Stacked Column Charts: Stacked column charts are employed to visualize CAE marks. This type of chart allows for the comparison of marks across different TAEs, providing insights into variations in marks scored in all three TAEs.[14]

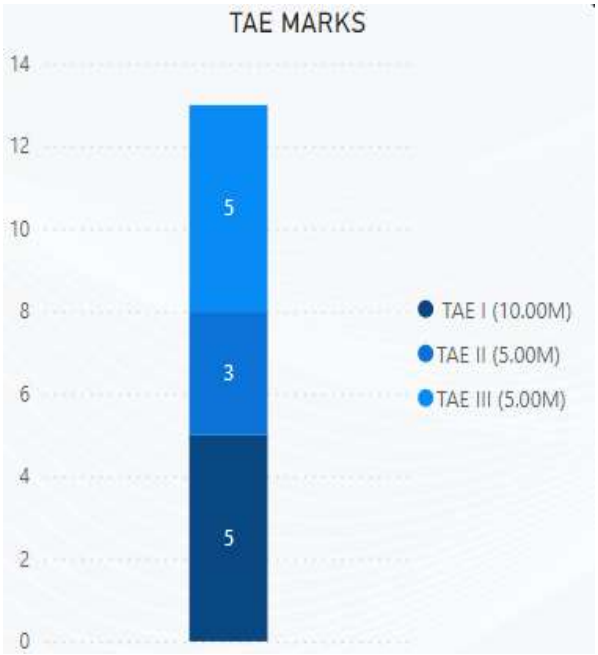

Fig 4: Stack Column Chart (For TAE Marks)

4. Meter Gauge Charts: Meter gauge charts are used to showcase students' total internal marks with respect to target values, indicating their performance levels and also telling whether the student has secured sufficient marks in internal marks or not so that a student must be able to gain good marks in his end semester exam. This visual

representation allows stakeholders to assess student performance relative to predefined benchmarks or goals, aiding in the identification of students who may require additional support or intervention.[16]

Fig 5: Meter Gauge Chart (For Total Internal Marks)

Overall, our project involves meticulously collecting and preparing data on various factors influencing student performance, such as attendance and exam scores. We then employ advanced techniques, including feature selection, correlation analysis, and predictive modelling, to uncover insights and develop robust predictive models. Additionally, we create dynamic dashboards using Power BI to visualize student performance metrics and predictive insights, allowing stakeholders to explore trends, set goals, and access personalized recommendations. By integrating dashboard visualizations into a web-based platform, we aim to deliver a solution that drives positive change in student performance and learning outcomes within the department.

IV. RESULTS

Our research yielded insightful findings across various facets of student performance analysis within the college department. Through meticulous data collection, preprocessing, and analysis, we were able to derive meaningful insights that shed light on key determinants of academic success and avenues for intervention.

Here, we present a summary of the key results obtained from our study:

Attendance Trends: Analysis of attendance records revealed notable trends in student attendance patterns throughout the academic term. By visualizing attendance data over time, we observed fluctuations in attendance rates and identified periods of heightened engagement and disengagement among students. These insights provide educators with valuable information to tailor intervention strategies aimed at improving student attendance and overall academic engagement.

Examination Scores: Examination score analysis provided valuable insights into student performance on Continuous Assessment Examinations (CAE) and Teacher Assessment Examinations (TAE). Through the examination of trends and distributions of examination scores, we identified areas of strength and weakness among students. Additionally, correlation analysis enabled us to discern relationships between examination scores and other performance metrics, offering insights into factors influencing academic achievement

.

Internal Marks: The computation and analysis of internal marks derived from CAEs and TAEs provided a comprehensive view of students' academic progress. By examining trends in internal marks over time, we were able to assess students' mastery of course material and their ability to apply knowledge in different contexts. Furthermore, correlation analysis facilitated the identification of correlations between internal marks and other performance indicators, guiding educators in targeted intervention strategies.

Web Platform Deployment: The deployment of the web-based platform provided educators and students with a centralized hub for communication, data exchange, and collaboration. Feedback from users highlighted the platform's usability and accessibility, with educators praising its intuitive interface for inputting student data and accessing performance dashboards. Students appreciated the platform's functionality for updating profiles, tracking progress, and accessing resources to support their academic journey. [5]

Overall, the results of our research underscore the importance of leveraging data-driven insights and advanced technologies to enhance student performance analysis within educational institutions. By providing educators and students with actionable insights and intervention strategies, our study contributes to fostering a culture of academic excellence and student success within the college department.

V. DISCUSSION

The comprehensive analysis of student performance within educational institutions is a paramount undertaking with profound implications for both educators and students alike. Our project endeavors to leverage data-driven insights and advanced technologies to deepen our understanding of student performance within a specific college department. Through meticulous collection and analysis of historical data encompassing various performance metrics, including attendance, examination scores, and internal marks, we aimed to construct a robust framework facilitating informed decision-making and proactive intervention strategies to elevate student outcomes.

A primary focal point of our research was the development of a user-friendly dashboard tailored to provide educators and students with intuitive access to performance metrics. This dashboard serves as a visual representation of attendance trends, examination scores, and academic

achievements, empowering users to pinpoint areas for improvement and monitor progress longitudinally. Leveraging interactive features and visualization techniques, such as Power BI, we sought to enhance user engagement and foster a culture of data-driven decision-making.

The integration of machine learning techniques for predictive modelling constituted another pivotal aspect of our research endeavour. By discerning patterns within past performance data, our aim was to construct predictive models capable of forecasting future student performance. These models offer educators invaluable insights into student progress, enabling them to proactively implement interventions in support of academic success. Moreover, correlation analysis facilitated the identification of key predictors of student achievement, enriching our predictive modelling efforts and providing actionable insights for educators.

The development of a web-based platform emerged as a requisite element for facilitating seamless communication and data exchange between educators and students. This platform affords educators secure login credentials to input student data, access performance dashboards, and engage in meaningful dialogue with their students. Simultaneously, students are empowered to update their profiles, monitor academic progress, and access pertinent resources to fortify their educational journey. By fostering collaboration and engagement, the platform cultivates an environment conducive to student success.
Throughout our research journey, we encountered various challenges and considerations that merit deliberation. The processes of data collection and preprocessing proved to be intricate and resource-intensive endeavors, necessitating meticulous attention to detail to ensure the accuracy and fidelity of the dataset. Additionally, the development of predictive models presented challenges in terms of model selection, evaluation, and interpretation of results. Future research endeavors could explore alternative machine learning methodologies and evaluation metrics to further refine and augment the predictive accuracy of our models. Furthermore, the deployment and integration of the dashboard and predictive models into the web-based platform mandated careful consideration of usability and accessibility parameters. While conscientious efforts were directed towards designing an intuitive interface, ongoing user testing and feedback mechanisms are indispensable to pinpoint areas for enhancement and refinement. Moreover, safeguarding data privacy and security within the platform remains an imperative to safeguard sensitive student information and uphold ethical standards.

In conclusion, our research endeavors represent a significant stride towards advancing student performance analysis within educational contexts. By harnessing the synergy of data-driven insights and technological innovations, we aspire to empower educators and students alike to make informed decisions and undertake proactive measures to elevate academic outcomes. Despite the encountered challenges, our commitment to continuous improvement and refinement underscores our dedication to fostering a culture of academic excellence within the college department.

VI. CONCLUSION

In conclusion, our paper aims to revolutionize student performance analysis within the college department by leveraging data-driven insights and predictive modelling techniques. Through the collection and analysis of historical data on attendance, exam scores, and teacher assessments, we have laid the foundation for a comprehensive understanding of student performance factors. By developing a user-friendly dashboard and integrating predictive models, we empower teachers and students to make informed decisions and take proactive steps to improve academic outcomes. The utilization of machine learning algorithms and correlation analysis allows us to identify key predictors of student success and provide personalized recommendations for academic improvement. Furthermore, the deployment of a web-based platform facilitates seamless communication and data exchange between teachers and students, fostering a collaborative learning environment. With continuous feedback and iteration, we strive to enhance the effectiveness and usability of our platform and models, ultimately driving positive change in student performance and learning outcomes within the college department.

VII. REFERENCES

[1] Bum, S. 1, Iorliam, I. B. 2, Okube, E. O. 1, and Iorliam, *Prediction of Student's Academic Performance Using Linear Regression*. Benue State University, Makurdi, Nigeria: Department of Mathematics & Computer Science.

[2] Pallavi Asthana, Sumita Mishra, Nishu Gupta, Mohammad Derawi And Anil Kumar ,Prediction of Student's Performance With Learning Coefficients Using Regression Based Machine Learning Models ,Gjøvik, Norway: Amity School of Engineering and Technology, Amity University, Lucknow Campus, Uttar Pradesh 226028, India & Department of Electronic Systems, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology.

[3] Institute for Intelligent Systems, University of Johannesburg, Johannesburg 2006, South Africa 2School of Electronic and Electrical Engineering, University of Leeds, LS2 9JT Leeds, U. K., "Student Performance Patterns in Engineering at the University of Johannesburg: An Exploratory Data Analysis."

[4] *Comparison of Linear Regression and Logistic Regression Algorithms for Ground Water Level Detection with Improved Accuracy*. C. Gnaneshwar Rajuudha and Sajiv G.

DOI: 10.1109/ICONSTEM56934.2023.10142495

[5] Student Performance Patterns in Engineering at the University of Johannesburg: An Exploratory Data Analysis Institute for Intelligent Systems, University of Johannesburg, Johannesburg 2006, South Africa 2School of Electronic and Electrical Engineering, University of Leeds, LS2 9JT Leeds, U. K., *Comparison of Linear Regression and Logistic Regression Algorithms Ground Water Level*.

[6] Oyerinde O. D. University of Jos and Jos, Nigeria. Chia, Predicting Students' Academic Performances – A Learning Analytics Approach using Multiple Linear Regression. Bauchi Road, Jos, Nigeria: International Journal of Computer Applications, 2017.

[7] Andrea Janes, Alberto Sillitti, and Giancarlo Succi, Effective Dashboard Design. ©2013 Cutter Information LLCCUTTER IT JOURNAL, 2013.

[8] Smt Sri Swathi, Preshitha Puppala, Divya Vasara, Sowmya Chippa, Prasanna Kumari Bhukya, Student Performance Analysis. India: IJSDR, 2022.

[9] Sumian Peng, "Research on Data Preprocessing Process in the Web Log Mining," IEEE Conference. Information Science and Engineering (ICISE), 2009 1st International Conference, pp. 942 - 945, 26-28 Dec. 2009

[10] M. Alhamadi. Challenges, strategies and adaptations on interactive dashboards. In Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization, pp. 368–371, 2020.

[11] Han, Hao, Tokuda, Takehiro. 2008/08/14 SN - 978-0-7695-3261-5. A Method for Integration of Web Applications Based on Information Extraction DO- 10.1109/ICWE.2008.29

[12] CROXTON F. E., STRYKER R. E.: Bar Charts Versus Circle Diagrams. Journal of the American Statistical Association 22, 160 (1927), 437–482

[13] SIMKIN D., HASTIE R.: An Information-Processing Analysis of Graph Perception. Journal of the American Statistical Association 82, 398 (1987), 454–465.

[14] GeeksforGeeks, https://www.geeksforgeeks.org/power-bi-create-a-stacked-column-chart/

[15] HEER J., BOSTOCK M.: Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. In Proceedings CHI (2010), pp. 203–212.

[16] Springer,Singapore, DOI:https://doi.org/10.1007/978-981-15-0772-4_24 Print ISBN978-981-15-0771-7