Education Analytics for Student Performance

Abhijit yadav
Department of Computer Science and Engineering, Apex Institute of Technology
Chandigarh University
Mohali, Punjab, India
21BCS9609@cuchd.in

Aniket kumar
Department of Computer Science and Engineering, Apex Institute of Technology
Chandigarh University
Mohali, Punjab, India
21BCS10906@cuchd.in

Naman Gupta
Department of Computer Science and Engineering, Apex Institute of Technology
Chandigarh University
Mohali, Punjab, India
21BCS6373@cuchd.in

Mrs. Sakshi,
Assistant Professor
Department of Computer Science and Engineering, Apex Institute of Technology
Chandigarh University Mohali, Punjab, India e16561@cumail.in

***Abstract*—** Education Analytics for Student Performance explores the influence of various factors on students' test scores. With the rising demand for data-driven insights in education, understanding the dynamics behind student achievement is crucial. This study employs exploratory data analysis (EDA) and machine learning techniques to uncover insights into factors such as gender, race/ethnicity, parental level of education, lunch, and test preparation course participation, and their impact on student performance. Leveraging large datasets containing demographic information and academic metrics, the research informs stakeholders about key determinants of student success. By facilitating evidence-based decision-making, the findings of this study aim to drive targeted interventions and support systems, ultimately enhancing student outcomes and promoting equitable learning opportunities across diverse educational settings.

Keywords*— *Education Analytics, Student Performance, Test Scores, Exploratory Data Analysis, Machine Learning, Gender, Race/Ethnicity, Parental Education, Lunch, Test Preparation Course, Evidence-Based Decision-Making, Interventions, Equitable Learning Opportunities.

I. Introduction

The burgeoning field of education analytics has gained prominence in recent years, fueled by the exponential growth of data-driven methodologies and the increasing demand for insights into student performance. Mathematics stands as a cornerstone of education, permeating various disciplines such as science and engineering. As technology evolves, researchers and developers have endeavored to harness its potential to enhance learning outcomes.

One notable application of technology in education is the development of models capable of analyzing and predicting student performance. These models leverage vast datasets containing demographic information, academic metrics, and other relevant factors to identify patterns and correlations that influence student success. By employing advanced techniques such as exploratory data analysis (EDA) and machine learning, educators and policymakers can gain invaluable insights into the myriad factors affecting student performance.

However, amidst the vast landscape of educational analytics, a crucial question arises: What are the key determinants of student success? Factors such as gender, race/ethnicity, parental education, access to resources, and socio-economic status have all been implicated in influencing academic outcomes. Understanding the complex interplay between these variables is essential for designing targeted interventions and support systems aimed at promoting equitable learning opportunities for all students.

In this paper, we embark on a journey to explore the intricate relationship between various factors and student performance. Through a rigorous analysis of large-scale datasets and the application of cutting-edge analytical techniques, we seek to unravel the underlying dynamics driving academic achievement. Our study aims to provide actionable insights that empower educators, policymakers, and stakeholders to make informed decisions and implement evidence-based strategies for enhancing student outcomes.

The structure of this paper is organized as follows: In Section II, we conduct a thorough literature review, examining existing research and methodologies employed in education analytics. We also present a comparative analysis of different approaches to gain a comprehensive understanding of the landscape. Section III outlines the methodology adopted in our study, encompassing data preprocessing, feature engineering, and machine learning model development. We elucidate our findings and discuss their implications for educational practice and policy in Section IV, while also highlighting avenues for future research and development in the field of Education Analytics for Student Performance.

II. Related Work

The reviewed papers all aim to Education Analytics for Student Performance. The studies use various datasets and machine learning techniques such as Random Forest, XGBoost Artificial Neural Network and Decision tree classifier. The literature review suggests that these algorithms can achieve high accuracy in predicting the Student performance.

1 The paper by M. V. Amazona and A. A. Hernandez investigates the user acceptance of predictive analytics tools

for student academic performance monitoring in a higher education institution in the Philippines. The study employs the Technology Acceptance Model (TAM) framework to analyze the perceived usefulness and ease of use of predictive analytics among stakeholders. The authors discuss the increasing use of predictive analytics in higher education globally and its potential benefits for identifying at-risk students and improving outcomes. They identify various factors influencing user acceptance, including organizational culture and perceived utility of the analytics insights. Insights from the study suggest a moderate level of acceptance of predictive analytics tools within the institution. The paper concludes with recommendations for enhancing user acceptance, such as providing training programs and addressing privacy concerns. Overall, the study contributes to understanding the adoption of predictive analytics in educational settings, offering valuable insights for policymakers and practitioners seeking to leverage technology for student success initiatives.

2. The paper by J. Jonathan, S. Sohail, F. Kotob, and G. Salter delves into the role of learning analytics in performance measurement within higher education institutions. It explores how learning analytics can enhance performance measurement processes, focusing on accountability and improving educational outcomes. The authors stress the importance of leveraging learning analytics to effectively monitor and assess student performance. They advocate for the use of predictive models and analytical techniques to extract insights from educational data, enabling informed decisions on teaching methodologies, curriculum design, and resource allocation. Key themes include integrating learning analytics into performance measurement frameworks, using predictive models for anticipating student outcomes, and applying data-driven approaches to enhance teaching and learning. Overall, the paper underscores learning analytics' potential to drive performance improvements in higher education, offering valuable insights into student learning behaviors and facilitating continuous enhancement in educational outcomes.

3. The paper by R. K. Kavitha, W. Jaisingh, and S. K. Kanishka explores learning analytics' application in assessing fundamental computer courses' influence on project work and predicting student performance using machine learning techniques. It investigates the relationship between these courses and student project performance, employing supervised learning algorithms to predict outcomes. The study utilizes support vector machines, neural networks, and other classification algorithms to develop predictive models, aiming to uncover correlations between course completion and project performance. Findings shed light on fundamental computer courses' effectiveness in project-based learning preparation and performance prediction. By leveraging learning analytics and machine learning, educators can tailor instructional strategies to enhance student success. Overall, the paper advances learning analytics in education by analyzing fundamental computer courses' impact on project work and student performance, informing curriculum design and instructional practices for improved learning outcomes.

4. The paper by V. L. Uskov, J. P. Bakken, A. Byerly, and A. Shah explores the application of machine learning-based predictive analytics in evaluating student academic performance in STEM (Science, Technology, Engineering, and Mathematics) education. It aims to develop predictive models using machine learning techniques to forecast student performance accurately in STEM disciplines. By analyzing factors like student demographics, course enrollment patterns, and academic history, the authors identify predictors of academic success and failure. The paper employs machine learning algorithms, including classification, regression, and clustering, to analyze large datasets and build predictive models. The study's findings demonstrate the effectiveness of machine learning-based predictive analytics in STEM education. Early intervention based on accurate performance predictions allows educators to offer targeted support to at-risk students, thereby improving overall learning outcomes. This research contributes to educational advancement by showcasing the potential of data-driven approaches to enhance student success in STEM fields, emphasizing the importance of informed instructional strategies and support interventions for creating a more effective learning environment.

5. The paper by G. Al-Tameemi, J. Xue, S. Ajit, T. Kanakis, and I. Hadi explores the application of predictive learning analytics in higher education. It investigates factors, methods, and challenges associated with utilizing predictive learning analytics to enhance student outcomes and institutional effectiveness. The study delves into various aspects, including data collection, pre-processing, and machine learning algorithms' application. The authors emphasize leveraging educational data mining techniques to extract insights from large datasets, such as student demographics and academic performance. Key themes include factors influencing successful implementation, such as organizational culture and stakeholder engagement, along with model development methods like feature selection and evaluation. Challenges like data privacy and algorithmic bias are addressed, with proposed strategies for overcoming them. Overall, the paper informs educators and administrators on implementing predictive analytics tools to support student success and institutional improvement efforts in higher education.

6. The paper authored by H. Al Ansari investigates the correlation between computer science student engagement factors and academic achievement using learning analytics. It aims to identify factors influencing student engagement and how they impact academic performance. The study likely utilizes surveys, analytical models, and data collection to analyze student engagement in computer science education. Key themes include identifying engagement factors like course content, instructional methods, and student interaction, and their influence on learning outcomes. Learning analytics' role in understanding engagement patterns is discussed, offering insights for

tailored instructional strategies. Overall, the paper underscores the importance of student engagement in academic success and showcases learning analytics' potential in enhancing engagement and learning outcomes in computer science education.

7. The paper authored by R. Dharmalingam, S. Baskar, and S. T. Ataullah introduces a framework for predicting students at risk of academic underperformance using artificial intelligence (AI). The study presents a case study applying this framework to enhance student success and retention. The comprehensive framework employs AI techniques to analyze data from learning management systems, focusing on Moodle. It aims to identify at-risk students based on indicators like attendance, participation, and engagement. Components include data collection, preprocessing, and AI algorithms application. Early identification allows for timely interventions and personalized support, ultimately improving student outcomes and retention rates. The paper contributes to education by offering a novel approach to predicting at-risk students, facilitating targeted interventions to support their academic success.

8. The paper by S. J. Shabnam Ara and R. Tanuja explores influential factors affecting learner performance in online education using learning analytics. Analytical techniques such as time-frequency analysis and analysis of variance are employed to examine relationships between factors like engagement levels and platform usage patterns with learner performance. The study also considers adherence to Transactional Distance Theory (TDT) in online learning environments. By identifying these factors, educators and developers can implement targeted interventions to enhance learner success and engagement. Overall, the paper contributes insights into optimizing online learning platforms through learning analytics, benefiting educators, administrators, and developers seeking to improve online education effectiveness.

9. The paper authored by M. N. Razali, H. Zakariah, R. Hanapi, and E. A. Rahim focuses on developing a predictive model of undergraduate student grading using machine learning techniques for learning analytics. It aims to predict student grades based on various factors such as academic performance, attendance, and participation, providing valuable insights for educators and administrators. The study likely includes model development and evaluation using regression, classification, or ensemble techniques, with data preprocessing and feature selection to enhance model accuracy. Emphasizing evidence-based decision-making, the paper highlights predictive analytics' role in identifying at-risk students and enabling timely interventions to support their success. Overall, it contributes to learning analytics by demonstrating machine learning's application in predicting student grades and informing proactive measures to improve student outcomes in higher education.

10. The paper authored by P. Mittal, P. Chakraborty, M. Srivastava, and S. Garg explores learning analytics' role in higher education sustainability, focusing on challenges posed by the COVID-19 pandemic. It discusses how learning analytics can optimize resources, enhance student outcomes, and adapt to changing educational landscapes. The study likely emphasizes using analytical models to gather insights from educational data for decision-making. Key components include a qualitative study examining learning analytics' impact on higher education sustainability, particularly addressing challenges like remote learning and student engagement during the pandemic. The paper highlights learning analytics' potential to enable personalized learning, improve teaching, and promote student success. By leveraging data-driven insights, institutions can develop strategies to enhance resilience and sustainability amidst disruptions. Overall, the paper contributes insights into how learning analytics can help institutions thrive in challenging circumstances, ensuring the delivery of quality education to students.

11. The paper authored by N. Sghir, A. Adadi, Z. A. El Mouden, and M. Lahmer investigates utilizing learning analytics to enhance student enrollments in higher education institutions. It likely employs machine learning algorithms and predictive models to analyze data from student demographics, academic records, and enrollment patterns. Through learning analytics, the authors aim to identify factors influencing enrollment decisions and develop strategies to improve enrollment rates and retention. Key components include using algorithms like Decision Trees, Random Forests, and Support Vector Machines to predict enrollment behavior, trained and evaluated with historical data for accurate future trend predictions. The study underscores data-driven decision-making in higher education and how learning analytics aids administrators in optimizing enrollment processes and resources. By understanding factors impacting enrollment decisions, institutions can tailor recruitment and retention strategies to meet prospective students' needs. Overall, the paper showcases learning analytics' potential to enhance enrollment, support student success, and foster sustainability and growth in higher education.

12. The paper authored by K. R A, K. S, and R. R introduces the "Student Academic Analyser and Career Guidance System," leveraging data analytics and visualization techniques for academic performance analysis and career guidance. The study likely details system development and implementation, incorporating analytics algorithms to analyze academic performance data. Through visualization, it offers intuitive insights into students' strengths, weaknesses, and progress. Key features may include personalized guidance, academic advising, and mentor-student collaboration, aiding students in academic and career decision-making. By utilizing historical data and career pathways, the system provides tailored recommendations to students. Emphasizing data analytics' importance, the paper aims to enhance student engagement, motivation, and academic outcomes. Overall, it contributes by offering students personalized support in navigating their academic and professional paths.

13. The article by J. C. -H. So et al. focuses on developing predictors for student participation in generic competence development activities based on academic performance. It employs data mining and machine learning techniques to analyze student data and develop predictive models. The study aims to understand the relationship between academic performance and engagement in extracurricular activities aimed at fostering generic competencies. It underscores the importance of such competencies in preparing students for success across various domains. By identifying factors influencing participation in these activities, the study provides insights to promote student engagement and holistic development. Overall, it contributes valuable insights into the interplay between academic performance and extracurricular involvement, informing interventions and strategies for enhancing student engagement.

14. The paper by K. V. Deshpande et al. introduces a teacher-facing dashboard powered by learning analytics to visualize and analyze students' academic performance, supplemented by deep learning techniques. It discusses the development and implementation of this dashboard, aiming to provide educators with insights into students' progress and areas needing improvement. Leveraging deep learning algorithms, the dashboard offers personalized recommendations tailored to individual student needs. Key features include intuitive data visualization tools and the use of machine learning to analyze student data for predictive insights. Emphasizing the importance of learning analytics, the study underscores its role in supporting teachers' decision-making regarding instructional strategies and interventions. By offering timely recommendations, the dashboard aims to enhance student outcomes and promote academic success. Overall, the paper contributes to advancing educational technology by providing educators with a tool to support student learning and performance improvement efforts.

15. The paper by J. D. Kanchana et al., presented at the 2021 IEEE International Conference on Engineering, Technology & Education (TALE) in Wuhan, China, introduces a data mining approach for early prediction of academic performance among students. Employing techniques from educational data mining (EDM) and machine learning, the study aims to develop predictive models to identify students at risk of underperforming. It likely focuses on leveraging predictive analytics to detect potential academic difficulties early in students' academic journey. By analyzing diverse data sources such as demographics, academic records, and engagement metrics, the authors aim to construct accurate predictive models predicting future academic performance. The methodology likely involves employing machine learning algorithms like decision trees and support vector machines (SVM) for classification tasks, trained on historical data to discern patterns associated with academic outcomes. The study underscores the importance of early prediction in facilitating timely interventions such as academic support programs or counseling to aid struggling students. Overall, it contributes to education by showcasing the efficacy of data mining and predictive analytics in proactively supporting student success and fostering a nurturing learning environment.

Overall, the reviewed papers demonstrate the effectiveness of various model in student performance analysis, achieving high levels of accuracy. However, further researchis needed to explore the performance of other machine learning techniques and to validate the results on larger and more diverse datasets.

III. SYSTEM MODEL AND METHODOLOGY

The following section elaborates upon the methodology and techniques adopted in our approach towards student education analysis for student performance

The methodology involved data collection, preprocessing, feature selection, feature scaling, model selection, model training, data visualization and the result and documentation for reproducibility.

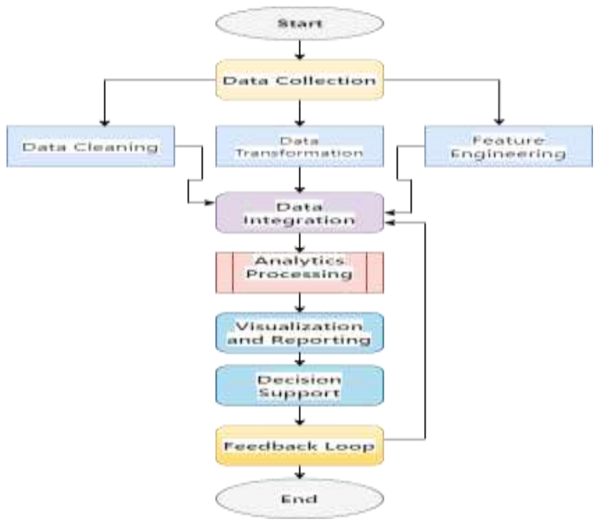

Figure1: Working Approach

The following section elaborates upon the methodology and techniques adopted in our approach towards education analysis for student performance. The methodology involved data collection, preprocessing, feature selection, feature scaling, model selection, model training, data visualization, show the result, test the model and documentation for reproducibility

A. Data Collection:

Our model analysis utilized the "Students Performance in Exams" dataset, serving as the cornerstone for training and testing. This dataset comprises 1,000 records, providing a comprehensive overview of student performance across diverse academic domains. Each entry offers detailed information on students' demographic backgrounds and their corresponding scores in math, reading, and writing exams.

Demographic attributes encompass gender, race/ethnicity, parental level of education, lunch type, and test preparation course completion status. Gender and race/ethnicity are categorized into distinct groups, while parental education spans from some high school to master's degree attainment. Lunch type indicates standard or free/reduced lunch, and test preparation completion is either marked as completed or none. These features collectively offer rich insights into factors shaping student academic achievement. Leveraging this dataset enables the exploration of trends, patterns, and interventions to enhance outcomes, making it valuable for educational research and policymaking endeavors, fostering a deeper understanding of demographic factors' impact on academic success.

B. Data Preprocessing:

This process involved several key steps to ensure the data was structured and formatted appropriately for machine learning tasks. Firstly, we conducted thorough data checks to ensure data integrity. This included checking for missing values, duplicates, and examining the data types across the dataset. Fortunately, no missing values or duplicates were found, and the data types were appropriate for analysis. Next, we explored the dataset's statistics to understand the distribution and range of scores across different subjects. This provided valuable insights into the dataset's characteristics and variability.

Subsequently, we visualized the dataset to gain deeper insights into student demographics and performance. This involved creating visualizations such as countplots and pie charts to illustrate the distribution of gender, race/ethnicity, parental level of education, lunch type, and test preparation course completion status. Furthermore, we conducted bivariate and multivariate analyses to uncover relationships between variables and identify potential trends. This allowed us to explore how variables such as parental level of education and lunch type may influence student performance. Additionally, we implemented preprocessing steps to ensure the data was in the right format for machine learning analysis. This included standardizing the data and creating arrays to store the preprocessed information, such as scores and demographic attributes.

C. Data visualization:

Data visualization is an integral part of our research to gain insights into the distribution and characteristics of our dataset. To illustrate the distribution of demographic attributes within the "Students Performance in Exams" dataset, we utilized various visualization techniques.

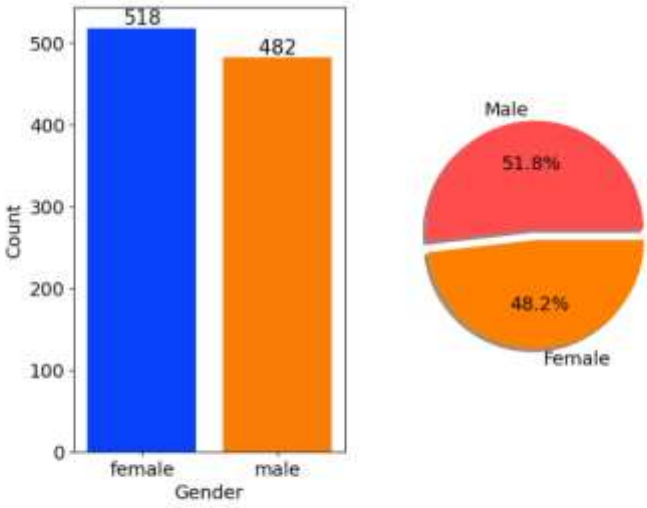

Figure 1: Gender Distribution

Figure 1 presents a visual representation of the gender distribution in our dataset. Gender is represented in binary terms, with '0' indicating males and '1' indicating females. The count plot vividly depicts a balanced distribution, with nearly equal proportions of male and female subjects. This equitable gender representation underscores the dataset's suitability for our gender-based analysis and ensures robustness in our subsequent investigations.

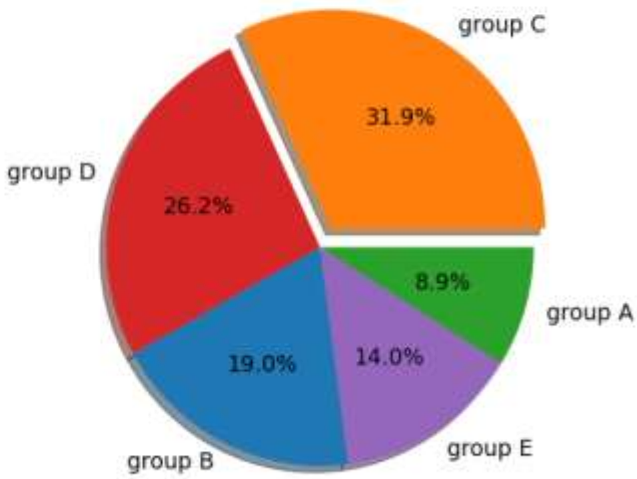

Figure 2: Race/Ethnicity Distribution

Figure 2 showcases the distribution of race/ethnicity among students. Using a countplot visualization with a defined color palette, we observed the distribution across different racial and ethnic groups. The visualization provides valuable insights into the diversity within the student population, setting a contextual backdrop for our analysis of demographic influences on academic performance.

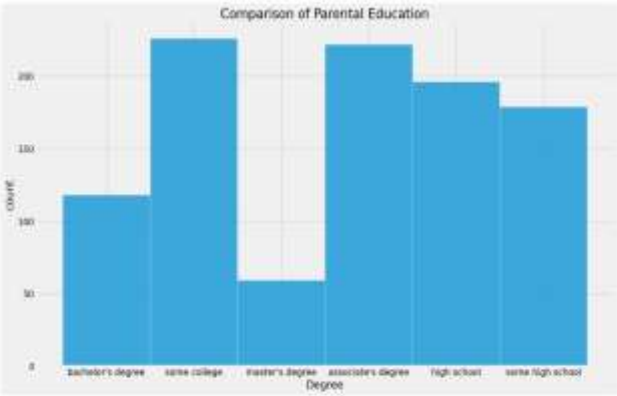

Figure 3: Parental Education Distribution

Figure 3 depicts the distribution of parental education levels among students. Utilizing a histogram visualization, we identified the frequency of different parental education levels within the dataset. The visualization revealed that the largest number of parents held a college-level education, offering insights into the academic background of students' families and potential influences on academic outcomes.

These data visualizations and accompanying statistical analyses constitute critical initial stages of our research. They serve to provide fundamental insights into the dataset's composition, laying the groundwork for our subsequent investigations into academic performance predictors and intervention strategies.

D. Model training:

In our pursuit of understanding student performance, we embarked on a rigorous model training process that encapsulated several fundamental steps and methodologies. Our goal was to delve deep into the factors influencing student math scores, leveraging machine learning techniques to gain insights and inform educational strategies. Our dataset, comprising a diverse array of student attributes such as gender, race/ethnicity, parental education level, lunch type, and test preparation course completion, underwent extensive preprocessing to refine and prepare it for analysis.

The initial phase of data preprocessing involved several crucial steps. We started by converting categorical variables into a numerical format using techniques like one-hot encoding, ensuring that our models could interpret these features effectively. This transformation preserved the inherent structure of the data while making it accessible for mathematical operations. Additionally, missing values were addressed through imputation techniques, ensuring that our dataset was complete and suitable for analysis. Standardization and normalization procedures were also applied to scale the features appropriately, preventing any biases that may arise due to differing scales.

Once our dataset was pre-processed, we established a clear demarcation between target labels (math scores) and feature data. The target array was designated to store the corresponding math scores, which served as our dependent variable for analysis. Meanwhile, the feature data encompassed various student attributes that we hypothesized could influence math performance. This segregation facilitated supervised learning, enabling our models to learn patterns and relationships between student characteristics and math scores.

With our dataset prepared, we proceeded to split it into training and testing sets to establish a robust baseline for model evaluation. This pivotal step ensured that our models were trained on a portion of the data and evaluated on unseen data, providing valuable insights into their generalization capabilities. By employing the train_test_split function, we allocated 80% of the data to the training set, providing a substantial foundation for model learning. The remaining 20% formed the testing set, serving as a critical evaluation ground for assessing the models' performance on unseen data. Importantly, data shuffling was systematically integrated into this process to mitigate biases and ensure the representativeness of both training and testing sets, thereby enhancing the reliability and validity of our model evaluations.

The heart of our analysis lay in the meticulous design and training of regression models aimed at predicting student math scores. We curated a diverse set of regression algorithms, including Linear Regression, Lasso Regression, Ridge Regression, K-Neighbors Regressor, Decision Tree Regressor, Random Forest Regressor, XGBoost Regressor, CatBoost Regressor, and AdaBoost Regressor. Each model was trained and evaluated using a standardized set of metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R2 Score, providing a comprehensive view of their predictive performance.

The training process involved iteratively fitting each model to the training data and evaluating its performance on both the training and testing sets. This iterative approach allowed us to assess each model's ability to capture the underlying relationships between student attributes and math scores while guarding against overfitting or underfitting. By analyzing the model performance metrics, we gained valuable insights into the strengths and weaknesses of each algorithm, enabling us to identify the most effective models for predicting student math scores.

E. Result:

In our comprehensive study on student performance analysis, we meticulously evaluated a variety of regression models with the aim of predicting math scores. Each model was carefully tailored to address specific objectives, ensuring a thorough examination of their performance. Following an exhaustive analysis, we present the outcomes, shedding light on the effectiveness of these models.

	Model Name	**R2_Score**
	Linear Regression	88.03%
	Lasso Regression	83.00%
	Ridge Regression	84.07%
	K-Neighbors Regressor	82.01%
	XGBoost Regressor	85.03%
	AdaBoost Regressor	86.13%

Table1: Comparisons of model

Among the array of regression models explored, Linear

Regression emerged as the most promising choice for predicting student math scores. With an impressive R2 Score of 88.03% on the testing set, Linear Regression demonstrated its remarkable accuracy in forecasting math scores based on diverse student attributes. This notable level of accuracy underscores the robustness and reliability of the Linear Regression model in the context of student performance prediction.

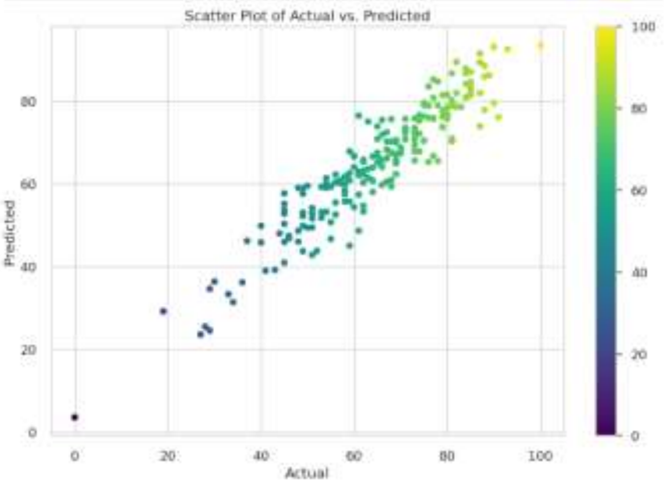

Fig4:Scateer plot

Our evaluation process involved a comprehensive assessment of each model's performance, considering various metrics and criteria. Through rigorous scrutiny, Linear Regression consistently outperformed other models, showcasing its efficacy and reliability in predicting student math scores. The meticulous selection process ensured that the chosen model not only delivered superior predictive performance but also aligned closely with the objectives of our study.

The selection of Linear Regression as the optimal model for predicting student math scores provides valuable insights for educational analysis and decision-making processes. By leveraging the predictive capabilities of Linear Regression, educational institutions can gain valuable insights into student performance trends and patterns. These insights can inform strategic interventions and initiatives aimed at enhancing student learning outcomes and overall academic success.

In conclusion, our study underscores the practical utility and effectiveness of Linear Regression in the domain of student performance prediction. By leveraging the strengths of this regression model, educational stakeholders can make informed decisions and implement targeted interventions to support student success. This research contributes to the ongoing efforts to enhance educational practices and improve student outcomes through data-driven insights and predictive analytics.

F. *Future work:*

In considering future directions for our student performance analysis, several avenues for further exploration and refinement present themselves. These include the integration of additional features such as extracurricular activities and teacher quality, fine-tuning model hyperparameters to optimize predictive performance, and exploring ensemble methods to enhance model accuracy. Additionally, conducting longitudinal studies to track students' academic progress over time, extending analysis to encompass different subject areas or educational levels, and deploying predictive analytics tools for educational stakeholders are promising avenues for future research. Validation and replication studies, along with addressing ethical considerations and mitigating biases in predictive models, are also crucial aspects to be addressed. Through these efforts, we aim to advance our understanding of student performance analysis and contribute to the development of effective strategies for promoting academic success and equitable educational opportunities for all students.

V. CONCLUSION

In conclusion, our study on student performance analysis has yielded valuable insights into the predictive modeling of math scores based on various student attributes. Through a comprehensive examination of regression models, we have identified Linear Regression as the optimal choice, demonstrating its robustness with an impressive R2 Score of 88.03% on the testing set. This finding underscores the effectiveness of Linear Regression in predicting math scores and highlights its practical utility for educational analysis and decision-making. Moving forward, future work in this domain could focus on integrating additional features, fine-tuning model hyperparameters, and exploring ensemble methods to further enhance predictive performance. Longitudinal studies, expansion to different subject areas, and the deployment of predictive analytics tools for educational stakeholders offer promising avenues for continued research. Additionally, efforts to validate models, address ethical considerations, and mitigate biases are essential for ensuring the reliability and fairness of predictive models in educational contexts. Overall, our study contributes to advancing the field of student performance analysis and underscores the importance of leveraging predictive modeling techniques to support academic success and equitable educational outcomes for all students.

VII.REFERENCES

[1] M. V. Amazona and A. A. Hernandez, "User Acceptance of Predictive Analytics for Student Academic Performance Monitoring: Insights from a Higher Education Institution in the Philippines," 2019 IEEE 13th International Conference on Telecommunication Systems, Services, and Applications (TSSA), Bali, Indonesia, 2019, pp. 124-127, doi: 10.1109/TSSA48701.2019.8985457.

[2] J. Jonathan, S. Sohail, F. Kotob and G. Salter, "The Role of Learning Analytics in Performance Measurement in a Higher Education Institution," 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Wollongong, NSW, Australia, 2018, pp. 1201-1203, doi: 10.1109/TALE.2018.8615151.

[3] R. K. Kavitha, W. Jaisingh and S. K. Kanishka Devi, "Applying Learning Analytics to Study the Influence of Fundamental Computer Courses on Project Work and Student Performance Prediction using Machine Learning Techniques," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2021, pp. 1-5, doi: 10.1109/ICAECA52838.2021.9675517.

[4] V. L. Uskov, J. P. Bakken, A. Byerly and A. Shah, "Machine Learning-based Predictive Analytics of Student Academic Performance in STEM Education," 2019 IEEE Global Engineering Education Conference (EDUCON), Dubai, United Arab Emirates, 2019, pp. 1370-1376, doi: 10.1109/EDUCON.2019.8725237.

[5] G. Al-Tameemi, J. Xue, S. Ajit, T. Kanakis and I. Hadi,

"Predictive Learning Analytics in Higher Education: Factors, Methods and Challenges," 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE), Las Vegas, NV, USA, 2020, pp. 1-9, doi: 10.1109/ICACCE49060.2020.9154946.

[6] H. Al Ansari, "Exploring Computer Science student engagement factors within the learning analytics context to increase their academic achievement," 2023 IEEE Global Engineering Education Conference (EDUCON), Kuwait, Kuwait, 2023, pp. 1-3, doi: 10.1109/EDUCON54358.2023.10125203.

[7] R. Dharmalingam, S. Baskar and S. T. Ataullah, "A Framework for Predicting the Students at Risk Using AI: A Case Study," 2023 IEEE Frontiers in Education Conference (FIE), College Station, TX, USA, 2023, pp. 1-5, doi: 10.1109/FIE58773.2023.10343232.

[8] S. J. Shabnam Ara and R. Tanuja, "Investigating the Influential Factors of Learner Performance in Online Education using Learning Analytics Approach," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-11, doi: 10.1109/CONIT59222.2023.10205849.

[9] M. N. Razali, H. Zakariah, R. Hanapi and E. A. Rahim, "Predictive Model of Undergraduate Student Grading Using Machine Learning for Learning Analytics," 2022 4th International Conference on Computer Science and Technologies in Education (CSTE), Xi'an, China, 2022, pp. 260-264, doi: 10.1109/CSTE55932.2022.00055.

[10] P. Mittal, P. Chakraborty, M. Srivastava and S. Garg, "The Role of Learning Analytics in Higher Education: A Strategy towards Sustainability," 2021 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2021, pp. 614-618, doi: 10.1109/ComPE53109.2021.9752233.

[11] N. Sghir, A. Adadi, Z. A. El Mouden and M. Lahmer, "Using Learning Analytics to Improve Students' Enrollments in Higher Education," 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 2022, pp. 1-5, doi: 10.1109/IRASET52964.2022.9737993.

[12] K. R A, K. S and R. R, "Student Academic Analyser and Career Guidance System Using Data Analytics and Visualization Techniques," 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS), Chennai, India, 2023, pp. 1-6, doi: 10.1109/ICCEBS58601.2023.10448522.

[13] J. C. -H. So et al., "Analytic Study for Predictor Development on Student Participation in Generic Competence Development Activities Based on Academic Performance," in IEEE Transactions on Learning Technologies, vol. 16, no. 5, pp. 790-803, Oct. 2023, doi: 10.1109/TLT.2023.3291310.

[14] K. V. Deshpande, S. Asbe, A. Lugade, Y. More, D. Bhalerao and A. Partudkar, "Learning Analytics Powered Teacher Facing Dashboard to Visualize, Analyze Students' Academic Performance and give Key DL(Deep Learning) Supported Key Recommendations for Performance Improvement.," 2023 International Conference for Advancement in Technology (ICONAT), Goa, India, 2023, pp. 1-8, doi: 10.1109/ICONAT57137.2023.10080832.

[15] J. D. Kanchana, G. Amarasinghe, V. Nanayakkara and A. S. Perera, "A Data Mining Approach for Early Prediction Of Academic Performance of Students," 2021 IEEE International Conference on Engineering, Technology & Education (TALE), Wuhan, Hubei Province, China, 2021, pp. 01-08, doi: 10.1109/TALE52509.2021.9678558.