



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423



Student Performance Analysis Using Linear Regression

Mayur Khandve¹, Ashish Chikne², Ganesh Mendhkar³, Ajnkya Salunkhe⁴,
Dnyaneshwar Aher⁵, Prof Ekta Patel⁶

Student, Department of Computer Engineering, Shree Ramchandra College of Engineering Pune,
Maharashtra, India¹

Professor, Department of Computer Engineering, Shree Ramchandra College of Engineering Pune,
Maharashtra, India⁵

ABSTRACT: Student performance analysis plays a crucial role in educational research and practice, enabling educators and policymakers to gain insights into factors influencing academic outcomes. Among the various data mining techniques, linear regression has emerged as a widely used and interpretable method for analyzing student performance. This review paper provides a comprehensive overview of the application of linear regression in student performance analysis, exploring its advantages, limitations, and implications for educational research and practice.

The paper begins with a thorough examination of the existing literature, summarizing studies that have utilized linear regression to analyze student performance. Key methodologies, datasets, and findings are synthesized to provide a comprehensive understanding of the current state of research in this area. Subsequently, the advantages of linear regression are discussed, highlighting its interpretability, simplicity, and ability to quantify the impact of independent variables on student outcomes. The flexibility of linear regression in handling diverse student-related factors, including both numerical and categorical variables, is also emphasized.

KEYWORDS: Student performance, Data mining, Linear regression, Educational data analysis, Predictive modeling, Academic outcomes

I. INTRODUCTION

In the field of education, understanding and predicting student performance is of paramount importance for educational institutions, policymakers, and educators. The ability to identify factors that influence student outcomes can aid in designing effective interventions and tailored educational strategies. With the increasing availability of educational data and advancements in data mining techniques, researchers have turned to methods such as linear regression to gain insights into student performance. This review paper provides a comprehensive examination of the application of linear regression in student performance analysis, highlighting its strengths, limitations, and implications for educational research and practice.

II. OBJECTIVE

The objective of this paper is threefold. Firstly, we aim to review and synthesize the existing literature on student performance analysis using linear regression. Secondly, we will examine the advantages of linear regression over other data mining techniques in the context of student performance analysis. Lastly, we will discuss the limitations and potential challenges associated with using linear regression and propose recommendations for future research in this domain.[3]



III. REVIEW OF LITERATURE

Mayur Khandve et al in," Review on Different Types of Algorithm for Data Mining" In this paper, we studied different types of algorithm used for data mining such as linear regression, logistic regression, decision trees, random forests, and naive Bayes. linear regression algorithm has many advantages like simple and easy to interpret, handle all types of independent variables and provide information about variables due to this algorithms linear regression algorithm can be best suited for analysis of student performance application

Asstprof.MohamedSadiq B B, Spoorthy H P R, STUDENT PERFORMANCE ANALYSIS SYSTEM,This paper deals with Learning based performance analysis of results a system that strives for excellence at different levels and dimensions in the area of student's interest. This system is designed to analyze and predict student Performance only

SomyaMishra ,MrunalLokare , AniketPatil ,STUDENT PERFORMANCE ANALYSIS SYSTEM,Student an easy and accurate data about projects and academic percentages. Students can view all the information in just one click which saves a lot of time and effort.

DurgeshUgale, JeetPawar,PERFORMANCE PREDICTION USING DATA MINING TECHNIQUES,CGPA is the most influence attributes indetermining the survival of students in their study, whether they can complete their study or not.

IV. REQUIUREMENT

- **Software Requirement**

- **HTML**

HTML (Hypertext Markup Language) is the standard markup language used for creating the structure and presenting the content of web pages. It is the backbone of web development and is responsible for defining the elements, layout, and formatting of a webpage.

- **Python**

Python is a computer programming language often used to build websites and software, automate tasks, and conduct data analysis. Python is a general-purpose language, meaning it can be used to create a variety of different programs and isn't specialized for any specific problems.

- **CSS**

CSS is used to define styles for your web pages, including the design, layout and variations in display for different devices and screen sizes.

V. SYSTEM ARCHITECTURE

System architecture in machine refers to the overall structure of the components and processes involved in a machine learning system. It encompasses the arrangement of hardware, software, data, and algorithms to enable the training, evaluation, and deployment of machine learning models. [2]

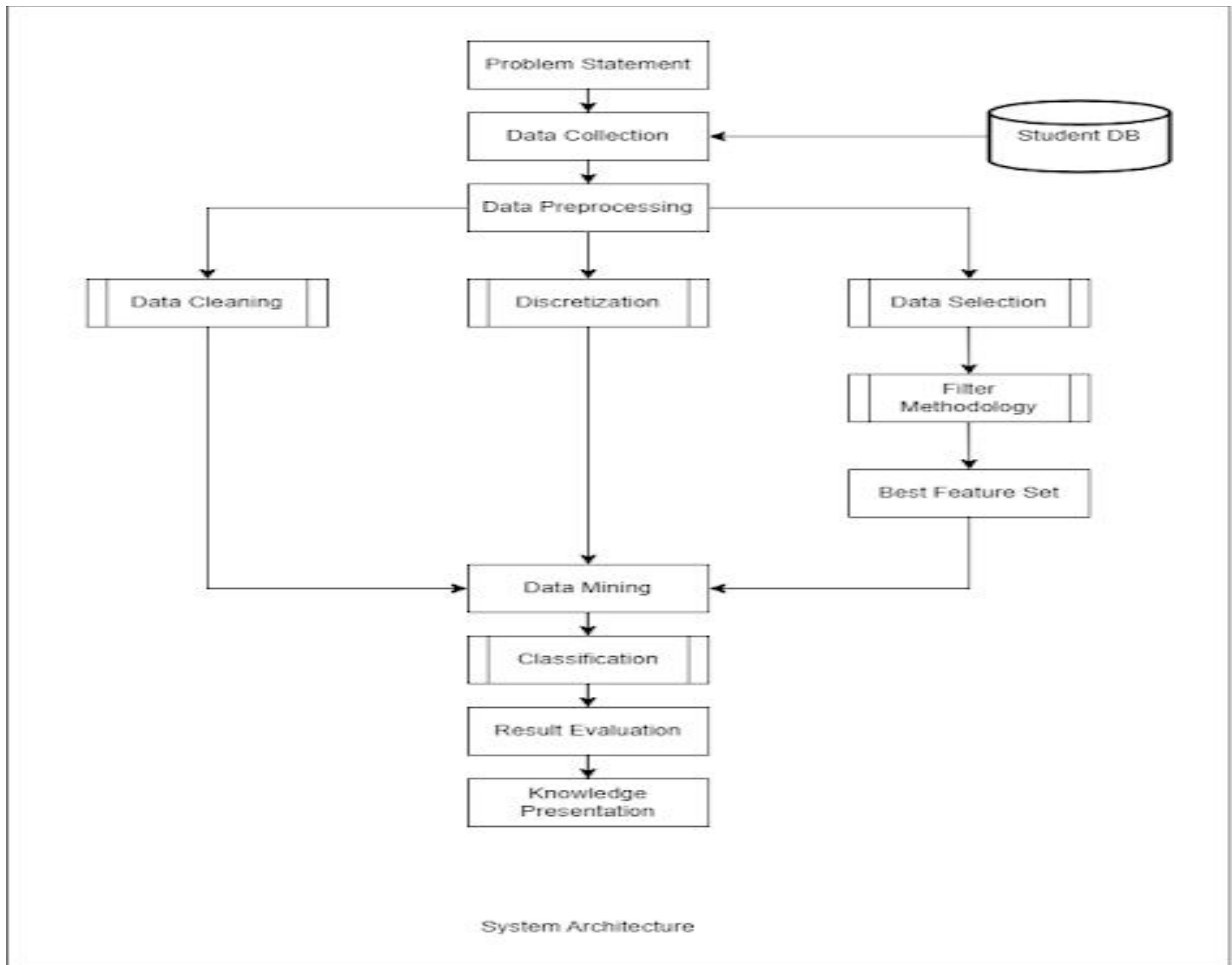


Figure 1.1 System Architecture

- A. **Data Collection:** The system begins by collecting relevant data from various sources, such as student records, academic databases, and educational platforms. This data may include information about student demographics, previous grades, attendance, study habits, and socio-economic factors.
- B. **Pre-Processing**
The collected data undergoes pre-processing to clean and transform it into a suitable format for analysis. This stage involves tasks such as data cleaning, handling missing values, feature selection, normalization, and encoding categorical variables.
- C. **Feature Engineering:**
In this step, additional features are derived or modified from the existing data to enhance the predictive power of the model. Feature engineering techniques may include creating new variables, aggregating data, or extracting meaningful features that can capture relevant patterns in student performance..
- D. **Model Training:** The training set is used to train the linear regression model. The model learns the relationship between the independent variables (such as previous grades, study time, etc.) and the dependent variable (student performance). During training, the model adjusts its coefficients to minimize the difference between the predicted and actual student performance

- E. Prediction and Analysis: Once the linear regression model is trained and evaluated, it can be used to predict student performance for new or unseen data. Given the relevant student characteristics, the model provides a predicted performance value. These predictions can be further analyzed and interpreted to identify influential factors and gain insights into the key determinants of student success.
- F. Visualization and Reporting: The system may include visualization components to present the analysis results in a user-friendly manner. Graphs, charts, and summary statistics can be generated to visually communicate the findings. Additionally, comprehensive reports summarizing the analysis and insights can be generated for further analysis and decision-making by educators and administrators.

VI. METHODOLOGY

- Data mining: data mining is the process of discovering patterns, relationships, and insights from large datasets. Data mining plays a crucial role in the analysis of student performance using linear regression. It involves extracting patterns, relationships, and insights from the collected student data to predict their performance and understand the factors influencing it. That's why we are using linear regression algorithm.
- **Linear regression:** Linear regression is a statistical modelling technique used to establish a linear relationship between a dependent variable (student performance) and one or more independent variables (predictor variables) such as previous grades, attendance, or study time. It predicts the numerical value of the dependent variable based on the values of the independent variables. Linear regression is a well-established and widely used technique in the field of predictive analytics. It has a strong theoretical foundation and is extensively studied, making it easier to interpret and understand the results. This interpretability is crucial in educational settings, where educators need to comprehend and explain the factors influencing student performance [1]

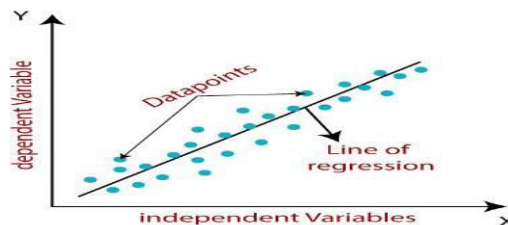


Fig2.1. Linear regression

- To evaluate the performance of our prediction model, we chose to implement the linear regression data mining technique. Linear regression offers several advantages that make it a suitable choice for this project compared to other data mining techniques.
- Firstly, linear regression is a well-established and widely used technique in the field of predictive analytics. It has a strong theoretical foundation and is extensively studied, making it easier to interpret and understand the results. This interpretability is crucial in educational settings, where educators need to comprehend and explain the factors influencing student performance.[1]

Flowchart

In order to enhance the clarity and visual representation of the processes involved in the student performance analysis using linear regression, a flowchart is included in this paper. The flowchart serves as a graphical illustration of the sequential steps and decision-making processes employed in the analysis[3]

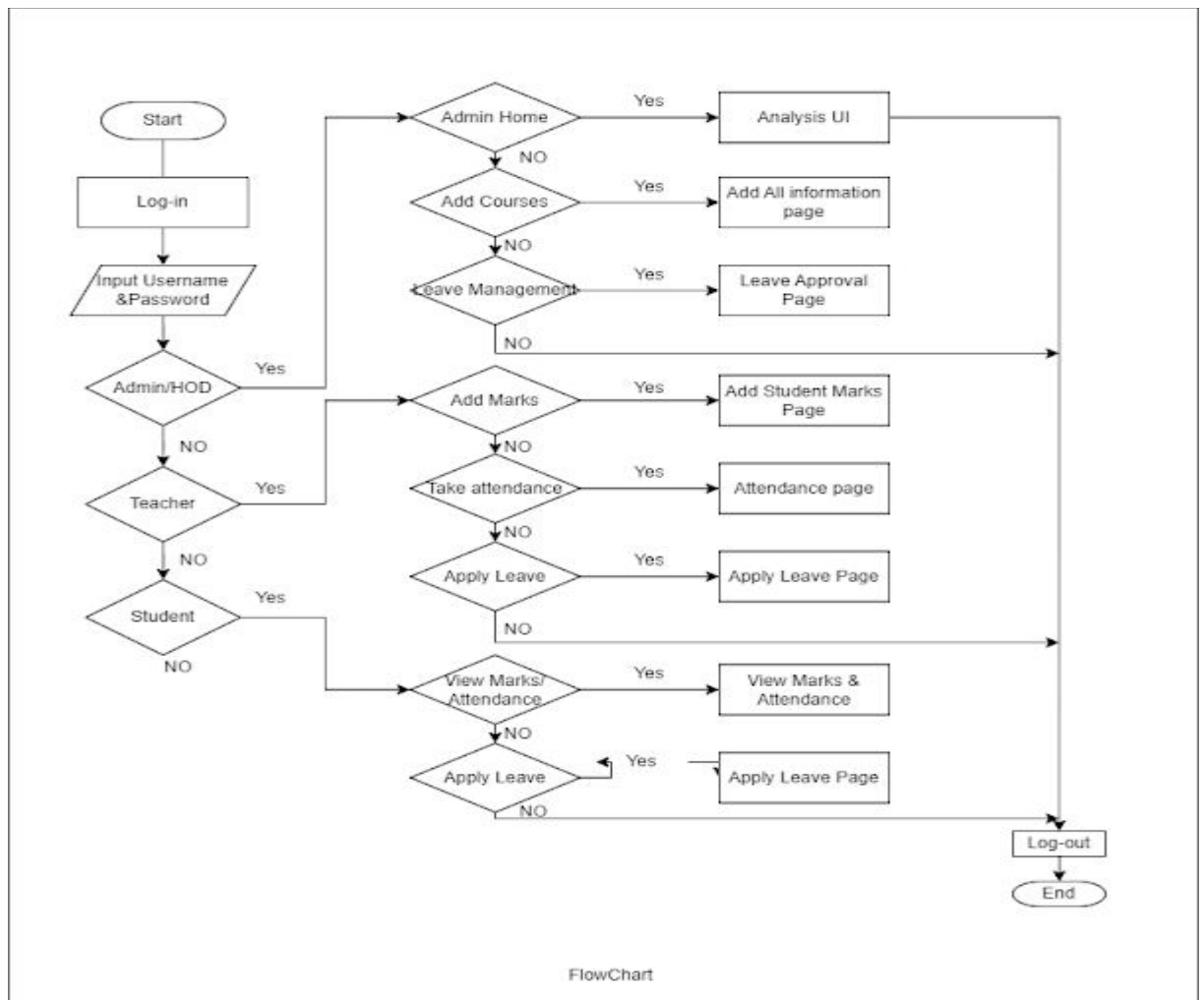


Figure 2.2 Flowchart

VII. RESULTS

□ Output

To provide concrete evidence and visual representation of the outcomes achieved in the student performance analysis project, a screenshot of the project output is included in this paper. The screenshot serves as a tangible illustration of the results obtained through the implementation of the linear regression model.

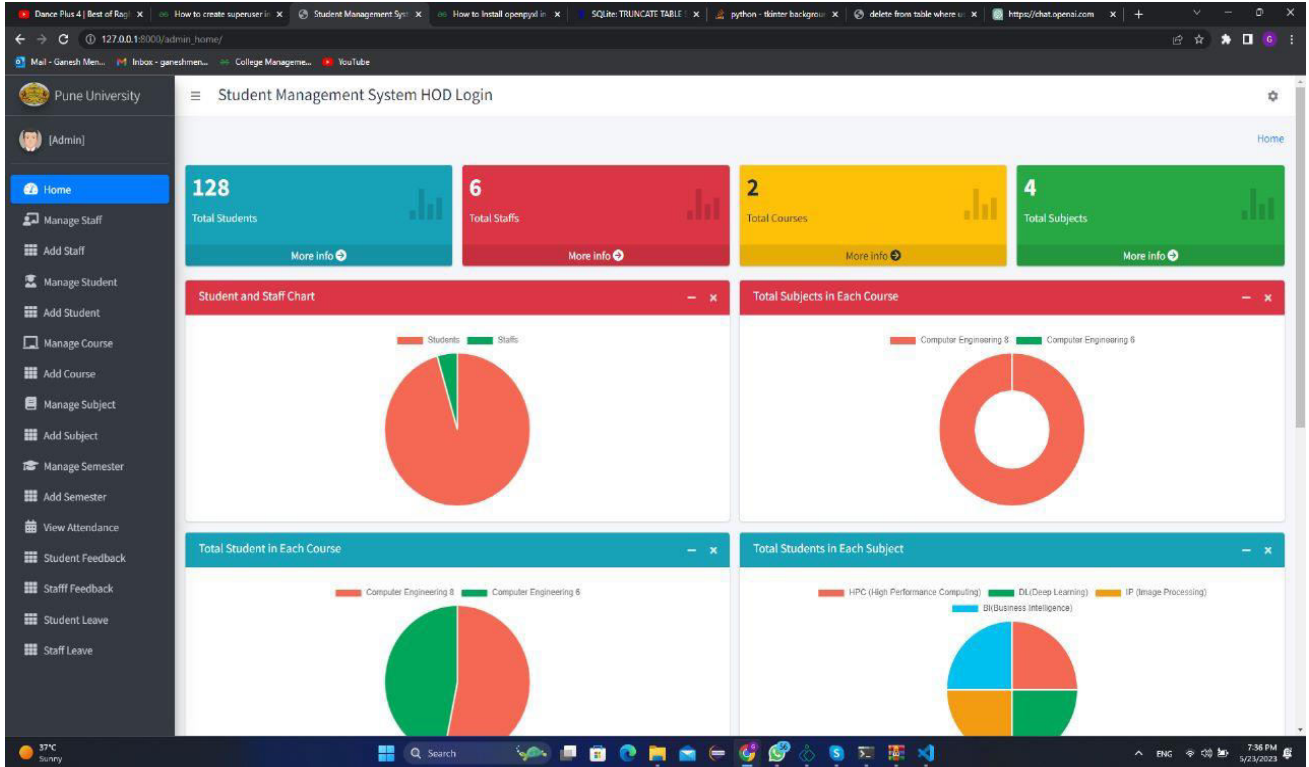


Figure3.1 HOD Login

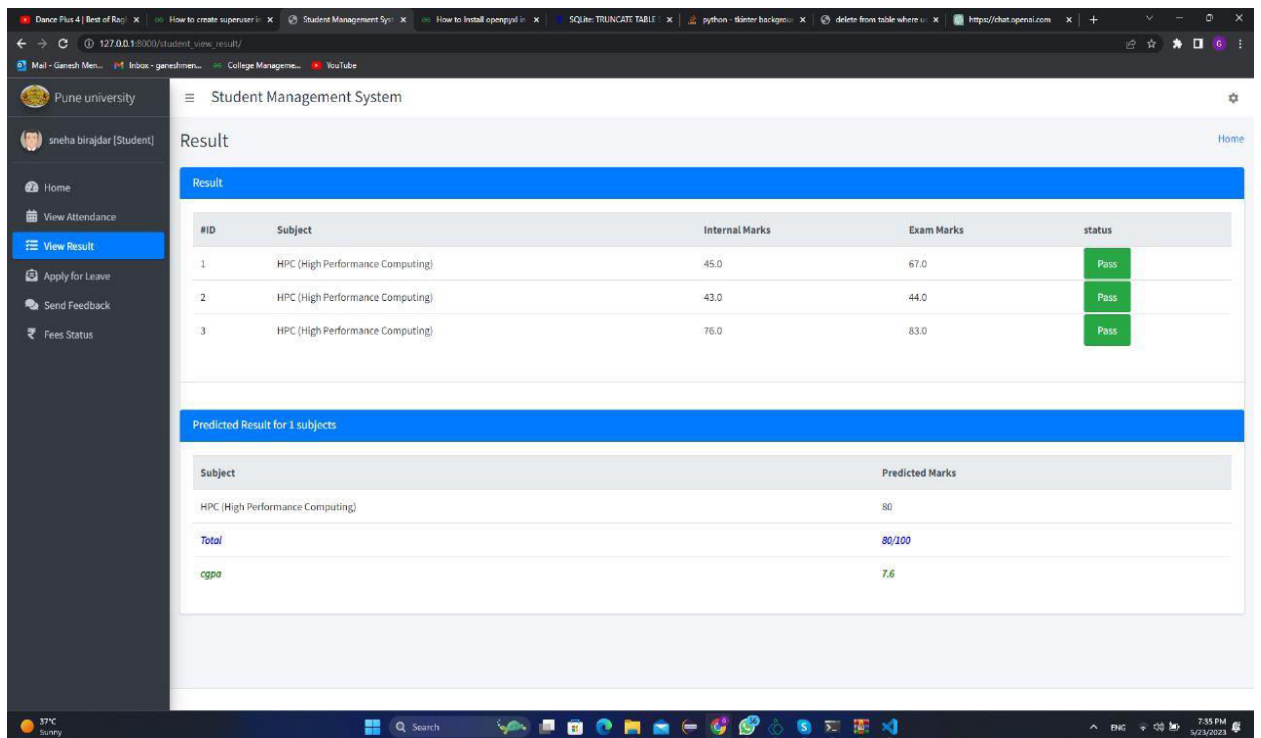


Figure 3.2 CGPA Prediction

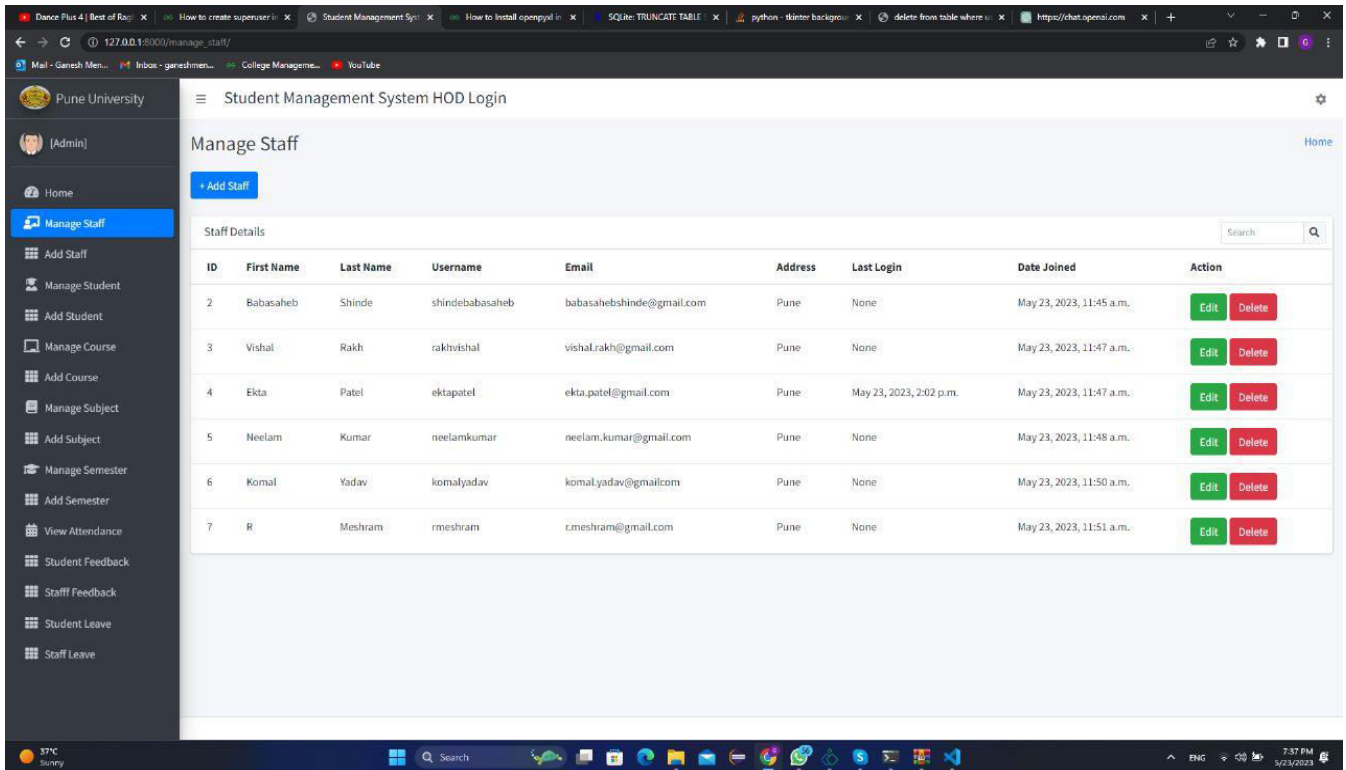


Figure 3.3 Staff Management

VIII. CONCLUSION

This paper has explored the topic of student performance analysis using linear regression. The study has highlighted the importance of understanding and predicting student performance to enhance educational strategies, identify at-risk students, and provide targeted interventions. By employing data mining techniques, particularly linear regression, and valuable insights can be gained from student data. The paper discussed the advantages of using linear regression over other data mining algorithms in the context of student performance analysis. Linear regression offers interpretability, simplicity, and the ability to identify important predictors. It allows educators and researchers to understand the relationship between independent variables, such as previous grades, study time, and socio-economic factors, and student performance.

REFERENCES

1. Mayur Khandve , Ashish Chikne , Ganesh Mendhkar , Ajnkya Salunkhe , Dnyaneshwar Aher , Prof Ekta Patel, "Review on Different Types of Algorithm for Data Mining" May 2023, International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), Volume 3, Issue 8, May 2023, PP-25-32
2. Durgesh Ugale, Jeet Pawar, Sachin Yadav, Dr. Chandrashekhar Raut, "Student Performance Prediction Using Data Mining Techniques", international research journal of engineering and technology (irjet) volume:07 issue:05 | may 2020.
3. Somya Mishra, Mrunal Lokare, Aniket Patil, Prof. Chandrashekhar Badgujar, "Student Performance Analysis System", international research journal of engineering and technology (irjet) volume:08 issue:04 | apr 2021.
4. Syed Furkan Mehdi, Syed Rayyan, Mohd Yaseen, Rafath Samrin "Student Performance Analysis Based Machine Learning" journal of information and computational science issn: 1548-7741 C. R. SRINIVASAN, B. RAJESH, P. SAIKALYAN, K. PREMSAGAR, AND E. S. YADAV, "A REVIEW ON THE DIFFERENT TYPES OF INETWORK (INTRUSION NETWORK)," J. ADV. RES. DYN. CONTROL SYST., VOL. 11, NO. 1, PP. 154-158, 2019.



5. Kirti Gadekar , "GENERAL LINEAR REGRESSION", 2018 IJRAR January 2019, Volume 06, Issue 1 www.ijrar.org (E-ISSN 2348-1269, P- ISSN 2349-5138), PP42-46
6. "2020 JETIR September 2020, Volume 7, Issue 9 www.jetir.org (ISSN-2349-5162) 1
7. Devita Durge, 2Nikhil Bagul, 3Rushikesh Gadge, 4Siddhesh Bhavsar
8. Date of Conference: 17-18 November 2014 Date Added to IEEE Xplore: 26 January 2015 Electronic ISBN:978-1-4799-6242-6
9. INSPEC Accession Number: 14872090 DOI: 10.1109/ICT4M.2014.7020662 Publisher: IEEE Conference Location: Kuching, Malaysia
10. Qingqing Chang, Research and Application of the Data Mining Technology in Economic Intelligence System, Volume 2022 | Article ID 6439315 , Cognitive Computing Paradigms for Medical Big Data Processing and its Trends| <https://doi.org/10.1155/2022/6439315>
11. Ch. Aswani Kumar; K. Sumangali, " Performance evaluation of employees of an organization using formal cept analysis", International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012)
12. Shaomei Yang; Qian Zhu, "An Evaluation Model on Employee Performance Based on Improved BP Neural Network", 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing.
13. TANG Yu-fang, ZHANG Yong-sheng,,—Design and implementation of college student information management system based on the web servicesl. Natural Science Foundation of Shandong Province(Y2008G22), 978-1- 4244-3930-0/09 2009 IEEE.
14. TANG Yu-fang, ZHANG Yong-sheng,,—Design and implementation of college student information management system based on the web servicesl. Natural Science Foundation of Shandong Province(Y2008G22), 978-1- 4244-3930-0/09 2009 IEEE.
15. Komal Jirage , Rucha Choudhari , Varsha Gupta , Anil Patthe And Disha Maind "Student Performance Analysis Based On Learning" international journal for research in emerging science and technology, volume-2, issue-3, march-2015



Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details