STUDENT PERFORMANCE PREDICTOR

Harika Boppana,  
 Department of Computer Science,   
 Manipal Institute of Technology,  
 MAHE,  
 Manipal, India.  
 [harikaboppana1@gmail.com](mailto:harikaboppana1@gmail.com)

Karthik Reddy,  
Department of Computer Science,  
Manipal Institute of Technology,  
MAHE,  
Manipal, India.  
[karthikcule@gmail.com](mailto:karthikcule@gmail.com)

*Abstract*—Periodic testing and introspection about one’s performance are essential in learning. It gains added significance in a school or college environment where students are required to take tests that determine their final grades. It can be challenging to analyse one’s performance by the students themselves objectively. An application that analyses past performance and predicts future performance would greatly aid the habit of introspection for students. This paper focuses on doing that precisely, considering several factors that could influence the student’s performance. Students can gain insights into their past performance and get an easy-to-understand report on where their current learning methods could lead them. We use machine learning techniques for regression and classification to predict a student’s qualitative and quantitative results and extensive data analysis to obtain insights into students’ past performance.

Keywords—data analysis, data visualisation, regression, classification, insights, prediction.

# Introduction

# Literature review

# Methodology

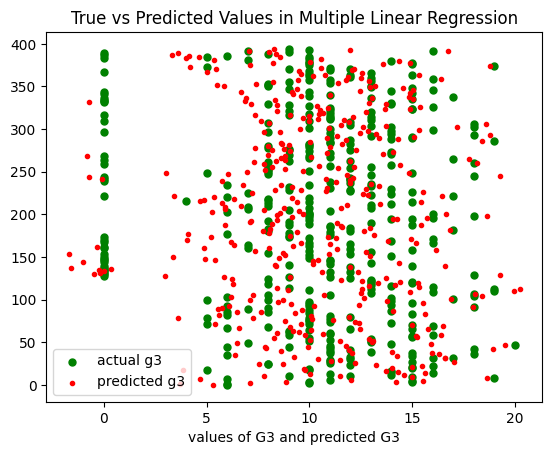
# Results and discussions

To demonstrate the purpose of giving a student a prediction based on several factors, we pick a random row and predict the G3 values for that row. The actual value for the randomly selected is 9.

1. Multiple linear regression

The RMSE value for the multiple linear regression is 1.88950.

The model predicted the value for random row to be 8.0309



1. Polynomial regression

The model was trained using 2nd ,3rd and 4th degrees to see what degree works best. The 3rd and 4th degrees seem to be overfitting for the data.

RMSE for 2 degree:1.295911432156298

RMSE for 3 degree:2.311849317612015e-13

RMSE for 4 degree:3.864185668074331e-13

The predictions for the randomly selected row:

degree 2:10.42008594903156

degree 3:9.00000000000015

degree 4:9.000000000000417

A graph of a graph of a graph

Description automatically generated with medium confidence

A graph of a graph with red dots

Description automatically generated with medium confidence

A graph of a graph with red dots

Description automatically generated with medium confidence

1. Logistic regression without SGD
2. Logistic Regression with SGD
3. KNN

# CONCLUSION

The best model for this data seems to be [model name].

# FUTURE WORK

We aim to deploy the project as an application students can download on desktop/mobile devices. The data to predict their performance can then be changed to suit the student, starting with a questionnaire to give some initial estimate. Henceforth, the application would have daily tracking of habits that influence performance, all combined to provide real-time predictions. This way, the student can see how they are progressing towards their goals as they make more changes to their study methods. The data will also be used to give visual insights into the student’s current habits, allowing them to decide on changes accordingly. To increase personalisation, the application would enable the student to choose habits that they think are affecting their performance, and predictions and insights are then given based on those habits. To scale it further, we could add a teacher-oriented feature where the teacher would be able to track their student’s progress and identify students that need more attention, as well as change their teaching methods according to current student performance and predicted results,

##### Acknowledgement

We express our gratitude to Dr Narendra V G for his dedicated guidance in navigating essential Machine Learning concepts, enabling us to construct the project and Dr Manjunath K N for constantly supporting us as we learned to implement the concepts that our project was built upon. Existing research and study of the field was necessary for our own understanding and ability to apply concepts learned. We sincerely appreciate the authors of referred scientific papers for their valuable contributions directly to the field of computer science and indirectly to this project

##### References

[1] A. A. Author, "Title of Paper," in \*Title of Conference\*, Place of conference, Year, pp. xxx-xxx.

[2] paper 2

[3] paper 3

[4] paper 4

[5] paper 5

[6] S. Gupta and J. Agarwal, "Machine Learning Approaches for Student Performance Prediction," in *10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2022, pp. 1-6*

[7] I. Singh, A. S. Sabitha and A. Bansal, "Student performance analysis using clustering algorithm," in *6th International Conference - Cloud System and Big Data Engineering (Confluence), Noida, India, 2016, pp. 294-299*

[8] C. Li Sa, D. H. b. Abang Ibrahim, E. Dahliana Hossain and M. bin Hossin, "Student performance analysis system (SPAS)," in *The 5th International Conference on Information and Communication Technology for The Muslim World (ICT4M), Kuching, Malaysia, 2014,* pp. 1-6

[9] H. M. R. Hasan, A. S. A. Rabby, M. T. Islam and S. A. Hossain, "Machine Learning Algorithm for Student's Performance Prediction," in *10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-7*

[10] H. Gull, M. Saqib, S. Z. Iqbal and S. Saeed, "Improving Learning Experience of Students by Early Prediction of Student Performance using Machine Learning," in *IEEE International Conference for Innovation in Technology (INOCON), Bengaluru, India, 2020, pp. 1-4*