

A Mini Project Report on

Predicting Student Performance using Regression Analysis

Submitted in partial fulfillment of the requirements for the award
of the degree of

Bachelor of Engineering

in

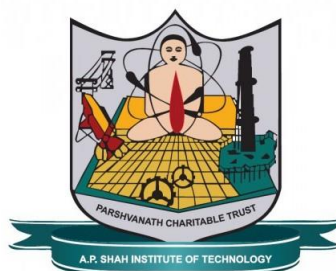
Computer Engineering

by

**Janhavi Anap (19102043)
Prathamesh Hambar (19102001)
Tejas Sheth (19102026)
Het Patel (19102005)**

Under the Guidance of

Dr. Pravin Adivarekar



Department of Computer Engineering

A.P. Shah Institute of Technology
G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615
UNIVERSITY OF MUMBAI

Academic Year 2021-2022

Approval Sheet

This Mini Project Report entitled “**Predicting Student Performance using Regression Analysis**” Submitted by “**Janhavi Anap**”(19102043), “**Prathamesh Hambar**”(19102001), “**Tejas Sheth**”(19102026) , “**Het Patel**”(19102005) is approved for the partial fulfillment of the requirement for the award of the degree of **Bachelor of Engineering** in **Computer Engineering** from **University of Mumbai**.

(Dr. Pravin Adivarekar)
Guide

Prof. Sachin H Malave
Head Department of Computer Engineering

Place:A.P.Shah Institute of Technology, Thane
Date: 9th August 2021

CERTIFICATE

This is to certify that the mini project entitled “**Predicting Student Performance using Regression Analysis**” submitted by “**Janhavi Anap**” (19102043), “**Prathamesh Hambar**” (19102001), “**Tejas Sheth**”(19102026), “**Het Patel**” (19102005) for the partial fulfillment of the requirement for award of a degree **Bachelor of Engineering in Computer Engineering**, to the University of Mumbai, is a bonafide work carried out during academic year 2021-2022.

(Dr. Pravin Adivarekar)
Guide

Prof. Sachin H Malave
Head Department of Computer Engineering

Dr. Uttam D.Kolekar
Principal

External Examiner(s)

1.

2.

Place: A.P. Shah Institute of Technology, Thane

Date: 9th August 2021

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that We have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Signature)

(Janhavi Anap 19102043)
(Prathamesh Hambar 19102001)
(Tejas Sheth 19102026)
(Het Patel 19102005)

Date: 9th August 2021

Abstract

The project focuses on predicting the final grade by analyzing the performance of the student. The prediction will be based on educational factors like cumulative GPA, personal factors like number of hours spent by a student for studying and social factors such as economical and educational background of their family and many more factors that may affect the GPA of the student.

It is important to find patterns in the student performance to be able to provide the necessary, accurate and timely diagnosis to the student. It also serves as a basic criterion for institutions to monitor the quality of education provided.

The Conventional Statistical Analysis and Artificial Neural Network prediction approach is necessary for the prediction. Conventional statistical evaluations help in identifying the multiple factors that actually affect the student performance. With these factors as input variables an Artificial Neural Network is modelled. Artificial Neural Network helps in analyzing large datasets which are not easily simplified through the conventional statistical techniques. It also helps to detect non-linear relationships between dependent and independent factors.

With the confirmation of the outputs from conventional statistical analysis, the training and testing of the model will be done for accurate prediction of the student performance. Performance of this neural network model is evaluated through various techniques. In this project one of the major techniques used for analysis is regression.

Keywords: Statistics, Analysis, Academics, Regression, Neural Networks

Introduction

The issues of how to prepare students to compete in a global economy are primary in education today. Prediction of student academic performance has long been regarded as an important research topic in many academic disciplines because it benefits both teaching and learning. Prediction of student academic performance will help instructors develop a good understanding of how well or how poorly the students in their classes perform, so instructors can be proactive. Non-Academic skills such as critical thinking, decision-making, and extra-curricular activities are as well essential for success in any walk of life. Predicted results are also important for the successful development of young people in society. Students who do well in academics will be better able to make the transition into adulthood and have a higher possibility of having employment opportunities and economic success.

Machine learning is a branch of artificial intelligence and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, uncovering key insights on student performance. Machine Learning will help us to predict students' grades with reasonable and significant accuracy, based on a given dataset.

Regression analysis is a predictive modelling technique that analyses the relation between the target or dependent variable and independent variable in a dataset. The different types of regression analysis techniques get used when the target and independent variables show a linear or non-linear relationship between each other, and the target variable contains continuous values.

The factors will be used in different regression algorithms and the most accurate model will be chosen for predicting student performance in upcoming semesters. The predictor/independent variables (namely, the inputs X_1 , X_2 , X_3 , etc.) of the regression models include a student's

X1: Cumulative GPA

The previous academic performance of the student heavily influences the cumulative GPA mathematically. Also, students with higher GPAs are more likely to study more to maintain a higher GPA.

X2: Attendance

It gives an insight into a student's dedication towards the course.

X3: Hours spent studying

The number of hours spent studying enhances one's grasping ability and helps keep pace with the instructor.

X4: Economical background of the family

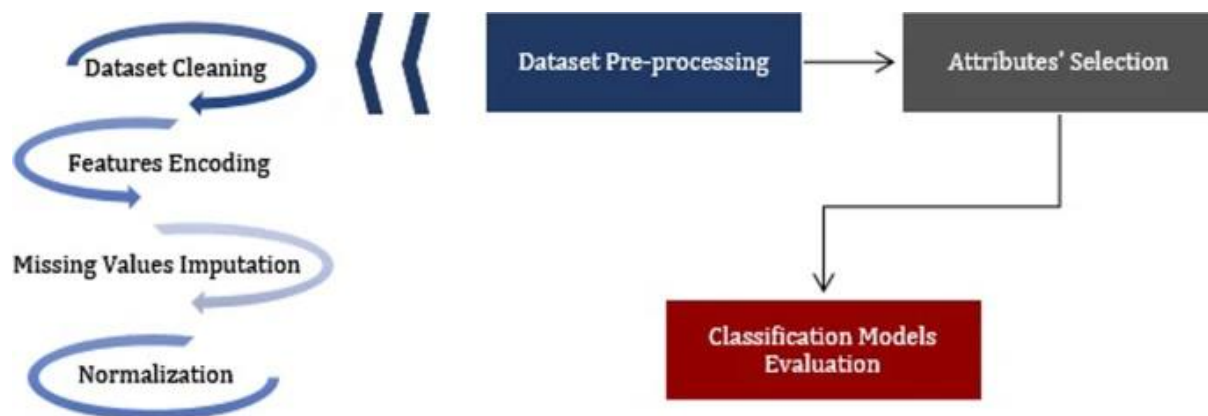
Parents with stable incomes are more likely able to afford and provide the resources required for their education.

X5: Educational background of the family

Educated parents provide motivation to the students to perform well since they know the importance of education.

The scope of the present study is limited in the investigation of the effects of cognitive factors on student academic performance.

Some of the limitations faced are that some students may only study with a test-based approach and not a concept-based approach. This analysis may not be able to consider factors such as mental and physical health. Also, the unpredictability of human nature and emotions is beyond the control of any other factors.



Problem Statement

To design an artificial neural network model that predicts the marks obtained by a student in their final semester via regression. The model considers factors such as previous academic record, attendance, socioeconomic background, etc.

Requirements

- **Python:** It is a General Programming Language used to code the model. It involves usage of several libraries such as numpy, pandas, scikit-learn, etc.
- **VS Code/Anaconda:** Integrated Development Environment to run, train and test the ML model
- **Git and Github:** A Version Control System used for collaboration
- **Google Analytics/Tableau/Power BI:** Data Visualization Tool to visualize the dataset and gain some insights
- **Google Workspace:** Project Management Tool to improve team coordination and file sharing

Bibliography

1. Huang, Shaobo, and Ning Fang. "Regression Models For Predicting Student Academic Performance In An Engineering Dynamics Course." *ASEE PEER Document Repository*, 2010, <https://peer.asee.org/15759>.
2. Sujatha, G. & S, Sindhu & Savaridassan, P.. (2018). Predicting students' performance using personalized analytics. *International Journal of Pure and Applied Mathematics*. 119. 229-237.
3. Hussain, S., Gaftandzhieva, S., Maniruzzaman, M. et al. Regression analysis of student academic performance using deep learning. *Educ Inf Technol* 26, 783–798 (2021). <https://doi.org/10.1007/s10639-020-10241-0>
4. Lau, E.T., Sun, L. & Yang, Q. Modelling, prediction and classification of student academic performance using artificial neural networks. *SN Appl. Sci.* 1, 982 (2019). <https://doi.org/10.1007/s42452-019-0884-7>
5. Abu Zohair, L.M. Prediction of Student's performance by modelling small dataset size. *Int J Educ Technol High Educ* 16, 27 (2019). <https://doi.org/10.1186/s41239-019-0160-3>

Acknowledgement

We have great pleasure in presenting the mini project report on **Predicting Student Performance using Regression Analysis**. We take this opportunity to express our sincere thanks towards our guide **Dr. Pravin Adivarekar** Department of Computer Engineering, APSIT thane for providing the technical guidelines and suggestions regarding line of work. We would like to express our gratitude towards his constant encouragement, support and guidance through the development of project.

We thank **Prof.Sachin Malave** Head of Department, Computer Engineering, APSIT for his encouragement during progress meeting and providing guidelines to write this report.

We also thank the entire staff of APSIT for their invaluable help rendered during the course of this work. We wish to express our deep gratitude towards all our colleagues of APSIT for their encouragement.

Student Name1: Janhavi Anap
Student ID1: 19102043

Student Name2: Prathamesh Hambar
Student ID2: 19102001

Student Name3: Tejas Sheth
Student ID3: 19102026

Student Name4: Het Patel
Student ID4: 19102005

