

PLAGIARISM SCAN REPORT

Words 959 Date May 04,2021

Characters 10032 Excluded URL

6%

Plagiarism

94%

Unique

3

Plagiarized
Sentences

49

Unique Sentences

Content Checked For Plagiarism

INTRODUCTION TO DATA SCIENCE PROJECT REPORT

Title of the project: Student Result Prediction Using Linear Regression

Name of the Student: CH. Rukmini

Regd. No: AP18110010298

rd th

Semester/Year: 3 year – 6 semester

Objective:

In every educational institution analyzing the performance of students and predicting the result is important to understand the capability and performance of every student. Predicting student performance can help the teachers to take steps in developing strategy for improving performances at early stages. With the help of data mining supervised and unsupervised techniques developing these kinds of applications are helping teachers to analyze students in a better way compared to existing performance. In these student result predictions using linear regression algorithms taking input as previously how many hours a student studied for the exam and their results. Based on that calculating pass and fail percentage of students.

Introduction:

Prediction of student results helps the instructors or teachers develop a good understanding of how well or how poorly the students in their class are performing whether they are able to pass or not or they are failed. So based on that result the instructors can take proactive steps to improve student learning and getting them pass and getting good marks. Based on a data set collected from students of their hours studied and their marks using a linear regression model to predict student performance.

Nowadays the most important research topic is predicting the student academic performance in many schools, colleges etc., because based on the performance teachers can improve the students with new outcomes. Instructors can easily identify the students who are performing well and who are performing average and who are performing poor in a class based on the predicted results. For example, after predicting the students' results, that shows some of the students in class are at risk in academics, instructors may help those students for achieving better results in future exams of study. Examples of taking proactive measures for better improvement by keeping extra classes, revising the syllabus again, more working hours and presenting animated videos for better understanding of lectures to improve student's improvement for better understanding and being active in classes and cooperative learning strategies, to name a few. A variety of mathematical techniques, such as linear regression has been employed to develop various models to predict student academic performance/result.

Literature Survey:

☐ In some research papers, the researchers are working to address the issue of the results of

students' analysis and their prediction. They used a classification-prediction technique to predict the exam results of the students. The author says that educational information is hidden and could play a major role in students' performance for better development of upcoming results.

□ Currently, the process of declaring and managing the students' results at the colleges, is performed manually with extensive human intervention. The students' results are obtained through a paper or excel application and kept that paper in notice boards or near classrooms or attached to a wall for declaration and then stored. Despite having an application that generates the result, the results are not very effective and the system consumes lots of time and human resources in performing various tasks, it is costly, it lacks data security and efficiency. At present, educational institutions need advanced technologies and computerized technologies. And once implemented, it will minimize all the problems mentioned.

Proposed Method:

Collecting the Pre-Processing the Analysing the data data data

Visualising the Data Science

plots Algorithm

Displaying the predicted Graph of the results result

- In this project, first we need to collect the data from source or need to write our own data based on the survey given by the students.
- After that, we need to pre-process the data before going to analyze the operation because the raw data which we have collected may be sometimes having null values or noise. So, we apply some cleaning techniques to pre-process the data.
- Getting the pre-processed data and analyzing the data.
- After that, applying the data science algorithm i.e., linear regression to the collected data.
- Then visualizing the plots and the respective graphs to predict the results.
- Finally, Graph of the predicted result.

Dataset Description:

We used our own dataset consisting of hours a student studied for exams and how many marks a student got. Based on that we will implement whether students passed or failed according to cutoff marks. It helps the instructors to develop the education of students and to motivate how to get good marks in upcoming exams.

Data Preprocessing:

It is a data science technique that transforms raw data into an understandable format. Raw data (real world data) is always incomplete and that data cannot be sent through a model. That would cause certain errors. That is why we need to preprocess data before sending through a model. For the data preprocessing techniques and algorithms, we used Scikit-learn libraries.

Steps in Data Preprocessing

Here are the steps I have followed;

1. Import libraries - We used pandas and NumPy to import libraries
2. Read data
3. Checking for missing values
4. Checking for categorical data
5. Standardize the data
6. PCA transformation
7. Data splitting

Pre-Processed Data:

Implementation of Code and the respective outputs and graphs:

-

Conclusion:

Predicting a student's performance would boost the results of a student's grades and gives the teachers a better approach for teaching the students who are at risk of failure. Regression models,

tree-based models are created to make the best predictions with high accuracy. The basic idea is to increase the efficiency of the prediction results using various algorithms. Thus, by finding the student grade whether he/she has passed or failed using a linear regression model gives optimal results.

References:

1. https://www.researchgate.net/publication/332893829_Predicting_Students'_Performance_Using_Machine_Learning_Techniques
2. <https://www.mdpi.com/2076-3417/10/3/1042/html>

Sources	Similarity
<p>GitHub - SagarGaniga/Data-Preprocessing: Data ...</p> <p>Data preprocessing is a data mining technique that involves transforming raw data into an understandable format.</p> <p>https://github.com/SagarGaniga/Data-Preprocessing</p>	3%
<p>Data Pre Processing Techniques You Should Know by ...</p> <p>It is a data mining technique that transforms raw data into an understandable format. Raw data(real world data) is always incomplete and that data cannot be sent through a model. That would cause certain errors. That is why we need to preprocess data before sending through a model.</p> <p>https://towardsdatascience.com/data-pre-processing-techniques-you-should-know-8954662716d6</p>	3%
<p>4.pdf - International Journal of Pure and Applied Mathematics ...</p> <p>The basic idea is to increase the efficiency of the prediction results using various algorithms . Thus by finding the RMSE, we observed that SVM,linear model ...</p> <p>https://www.coursehero.com/file/54885611/4pdf/</p>	2%