Student Grades Prediction

Introduction

The idea of this project is to find a linear regression model that would predict future grades for students based on their gender, nationality, which parent is responsible for student, how many times the student raises his/her hand on classroom, etc.

Outline of the Program

Part 1: Fit the linear regression model.

1. Data Processing

a. drop unnecessary columns—drop the columns that are not strongly related to the students' grades.

b. convert some of the columns into binary columns—turn the quantitative data into qualitive data so they can be used in the model.

c. find the means of grouped data for further analysis—divide the values by the total numbers of each value.

d. a simple linear check by plotting scatter plots— When all times the student participates on discussion or raises his/her hand on classroom or visits a course content and the student's grade, are plotted on a scatter plot, it's easier to find out if there's a linear relationship between the predicators and students' grades.

2. Fit A Linear Regression Model

calculate the intercept and slope of each predicators so we can make predictions with this model.

Part 2: Add a Graphical User Interface(GUI) to gather information from students, and display their predicted grades .

1.Gather Information from Users

2.Make Predictions with the Inputs

Information about the Variables

Source

Elaf Abu Amrieh, Thair Hamtini, and Ibrahim Aljarah, The University of Jordan, Amman, Jordan, http://www.Ibrahimaljarah.com www.ju.edu.jo

Attributes (1)

1 Gender - student's gender (nominal: 'Male' or 'Female')

2 Nationality- student's nationality (nominal:' Kuwait',' Lebanon',' Egypt',' SaudiArabia',' USA',' Jordan',' Venezuela',' Iran',' Tunis',' Morocco',' Syria',' Palestine',' Iraq',' Lybia')

3 Place of birth- student's Place of birth (nominal:' Kuwait',' Lebanon',' Egypt',' SaudiArabia',' USA',' Jordan',' Venezuela',' Iran',' Tunis',' Morocco',' Syria',' Palestine',' Iraq',' Lybia')

4 Educational Stages- educational level student belongs (nominal: 'lowerlevel','MiddleSchool','HighSchool')

5 Grade Levels- grade student belongs (nominal: 'G-01', 'G-02', 'G-03', 'G-04', 'G-05', 'G-06', 'G-07', 'G-08', 'G-09', 'G-10', 'G-11', 'G-12 ')

6 Section ID- classroom student belongs (nominal:'A','B','C')

7 Topic- course topic (nominal:’ English’,’ Spanish’, ‘French’,’ Arabic’,’ IT’,’ Math’,’ Chemistry’, ‘Biology’, ‘Science’,’ History’,’ Quran’,’ Geology’)
8 Semester- school year semester (nominal:’ First’,’ Second’)
9 Parent responsible for student (nominal: ’mom’, ’father’)
10 Raised hand- how many times the student raises his/her hand on classroom (numeric:0-100)
11- Visited resources- how many times the student visits a course content(numeric:0-100)
12 Viewing announcements-how many times the student checks the new announcements(numeric:0-100)
13 Discussion groups- how many times the student participate on discussion groups (numeric:0-100)
14 Parent Answering Survey- parent answered the surveys which are provided from school or not (nominal: ’Yes’, ’No’)
15 Parent School Satisfaction- the Degree of parent satisfaction from school(nominal: ’Yes’, ’No’)
16 Student Absence Days-the number of absence days for each student (nominal: above-7, under-7)

How to use the program

As l mentioned before, this program has two parts, one is to fit the linear regression model, the other is to add a Graphical User Interface (GUI) to gather information from students and display their predicted grades. It takes a few steps to run through the program.

Step 1: Download Files

Download and extract the zip file named ‘projects code.zip’, and make sure you save the three files in the same folder.

Step2: Library Documentation Installation

Libraries used in part 1 are ‘numpy’, ’matplotlib’, ‘pandas’ and ‘sklearn’, and part2 depends on ‘tk’ and ‘numpy’. To run this program, we have to get all these libraries installed. The installation documentation links are as follows.
Numpy https://pypi.org/project/numpy/#files
Matplotlib https://pypi.org/project/matplotlib/#files
Pandas https://pypi.org/project/Pandas3/#files
Sklearn https://pypi.org/project/sklearn2/#files
Tk https://pypi.org/project/tk64/#files

Step3: Open and Run the Model (take Wing as an example)

Files can be opened from the **Project** tool by double clicking or middle clicking on the file name, or by right-clicking and using the **Open** in **Wing** menu item(2). Open the file named *‘480project part1-build a model.py’* and press the green play button and you would see the outputs (intercept and slopes of each predicators) in the python shell. This might take you a few seconds due to the size of data. And the scatter plot will be shown on the screen, you can zoom it to see the whole picture.

Step4 : Open and Run the Predication Program
Open the file named '*480project part2 -display the prediction.py'* by repeating what you did in step3, press the green button and you would see an interface shown on your screen.

Step5: Select or Enter Information in GUI
Click 'submit' after you select or enter the student's information, and the predication pops up, e.g. 'As predicted, your grade level might be G-03'. Click the 'quit' button if you want to end this program.

Practical Use of the Program

Students or teachers can use this program to know how the predicators impact students' grades, e.g. negatively or positively, and how strong the impact is. They can also predict students' grades to identify students at risk of failing and know there needs an improvement.
There are actually lots of more factors that can affect students' grades, it's impossible to cover all of them in the program. But with this program, teachers can help them improve school performances effectively, for example, encourage them to raise hands more(answer or ask more questions) or spend more time prepare for the exams other than working on the attributes covered in this program.

Issues l had in the process

1.I used 'quit' command in the 'Quit' button(seen in the 'part2' file l submitted) in the first place, but it didn't work very well as it took long to end the program. With the command changed into 'withdraw'(seen in my presentation), the program ends as long as you click 'Quit'.

Reference

1. IbrahimAljarah. (2016, November 26). Students' Academic Performance Dataset. Retrieved from https://www.kaggle.com/aljarah/xAPI-Edu-Data.
2. Wingware. (n.d.). Wing Python IDE. Retrieved from https://wingware.com/doc/proj/navigating-to-source.