Predicting Acceptance to Graduate School Based on Test Scores and GPA

Minkyu Lee

Introduction

Graduate School is an opportunity and privilege to enhance individuals' knowledge, experience, and resume. However, individuals most possess certain skills to be admitted into different skills. Some requirements are GRE, TOEFL, and undergraduate GPA to assess the students' capabilities and potential.

This project will analyze the correlation between these scores and the chance of admittance into graduate schools. By training the model based on the collected data, it will provide a speculation and prediction of one's admittance to a graduate school.

This project will interests and benefit students looking to assess their levels.

Data Acquisition

The data used for this project was from https://www.kaggle.com/mohansacharya/graduate-admissions, compiled by Mohan S Acharya. The data contains the GRE Score, TOEFL Score, university rating, statement of purpose, letter of recommendation, cumulative GPA, research, and chance of admittance for 400 students. Initially, when imported, the dataframe looked like this:

```
df.head()
```

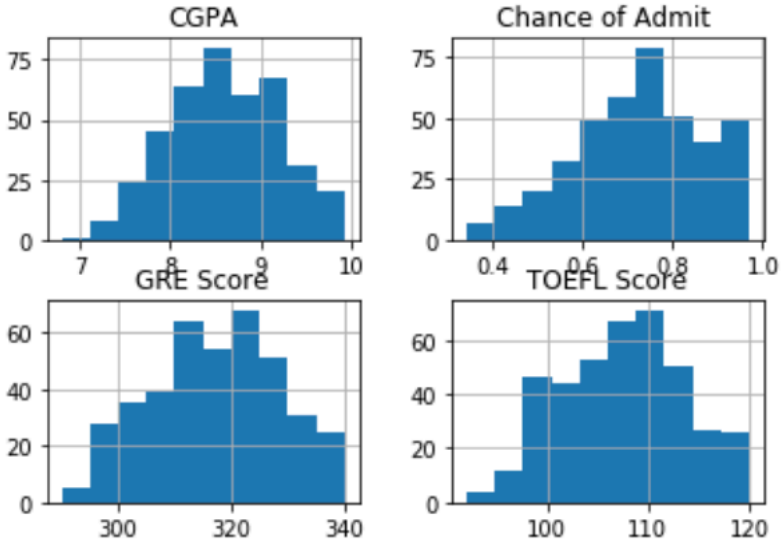
:

	Serial No.	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Research	Chance of Admit
0	1	337	118	4	4.5	4.5	9.65	1	0.92
1	2	324	107	4	4.0	4.5	8.87	1	0.76
2	3	316	104	3	3.0	3.5	8.00	1	0.72
3	4	322	110	3	3.5	2.5	8.67	1	0.80
4	5	314	103	2	2.0	3.0	8.21	0	0.65

For the purpose of this project, only the columns GRE score, TOEFL score, cumulative GPA, and chance of admittance was kept. Cumulative GPA is out of 10, the chance of admittance is a percentage, GRE Score is out of 340, and TOEFL score is out of 120.

Data Exploration

Cumulative GPA, GRE score, and TOEFL score are independent variables while chance of admittance is the dependent variable. Below is the histogram of the data:



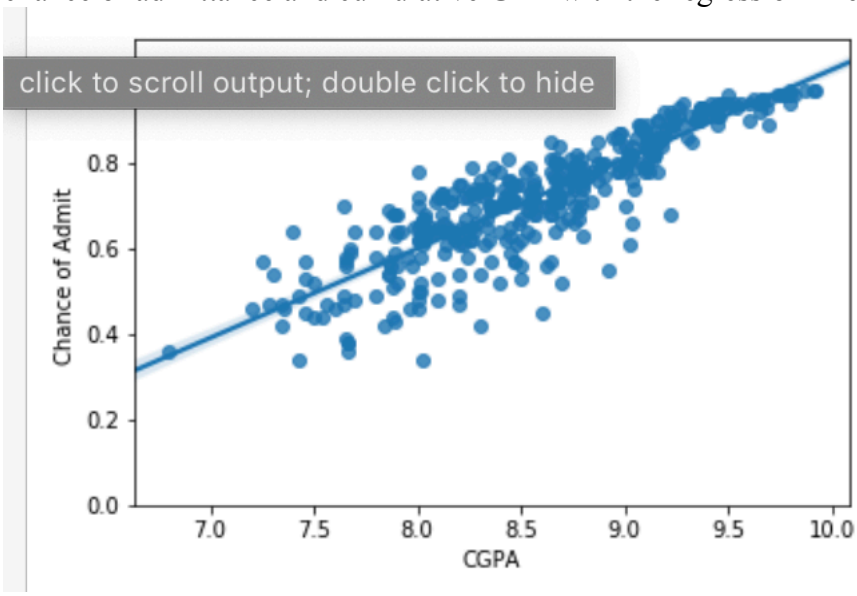
The data correlation seen is:

```
graph1.corr()
```

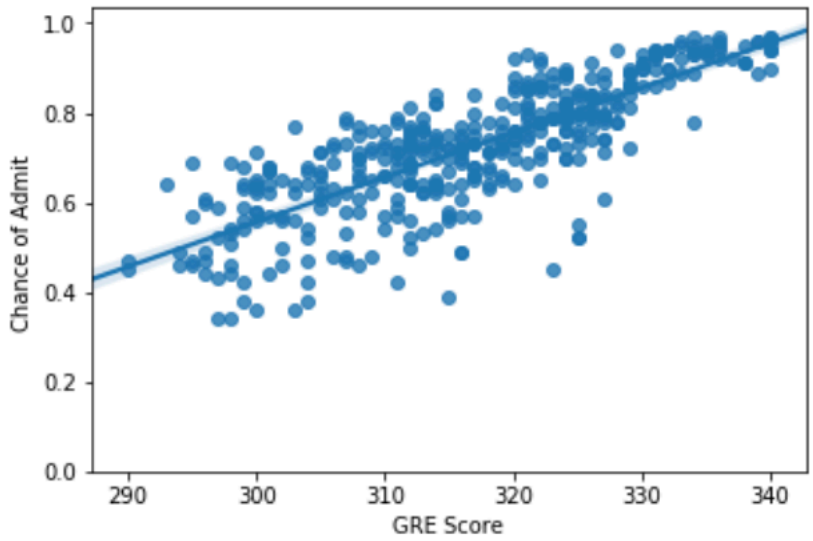
:

	GRE Score	TOEFL Score	CGPA	Chance of Admit
GRE Score	1.000000	0.835977	0.833060	0.802610
TOEFL Score	0.835977	1.000000	0.828417	0.791594
CGPA	0.833060	0.828417	1.000000	0.873289
Chance of Admit	0.802610	0.791594	0.873289	1.000000

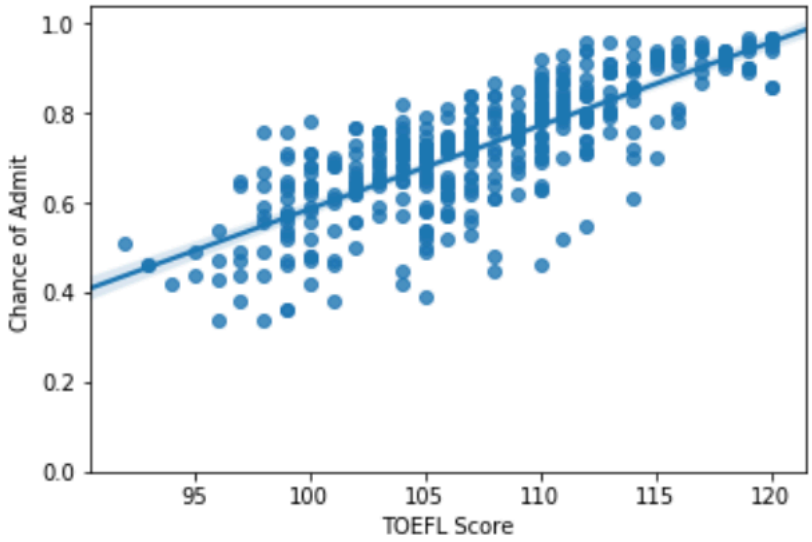
Linear regression will be used to analyze the correlation. The linear correlation between the chance of admittance and cumulative GPA with the regression line is:



The linear correlation between the chance of admittance and GRE score with the regression line is:



The linear correlation between the chance of admittance and TOEFL score with the regression line is:



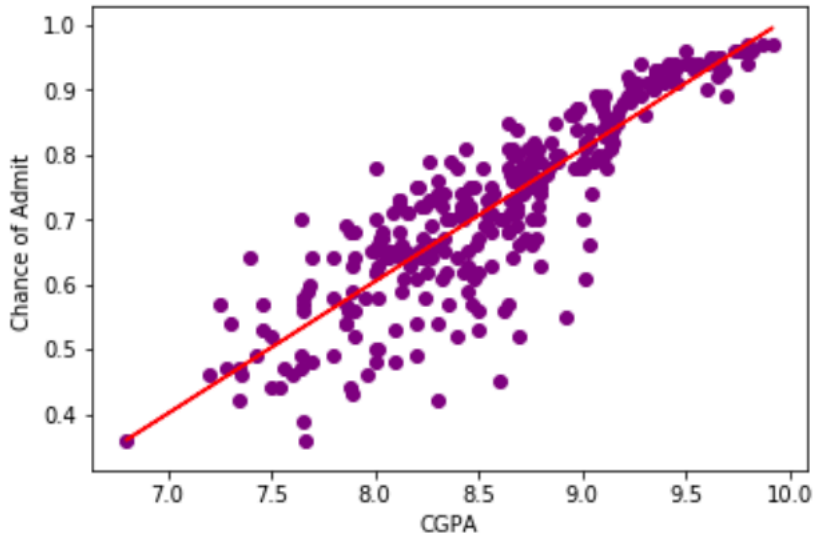
With these data, we found:
The Pearson Correlation Coefficient between cumulative GPA and chance of admittance is 0.8732890993553003 with P = 2.336514000498104e-126;
The Pearson Correlation Coefficient between GRE score and chance of admittance is 0.8026104595903505 with P = 2.458112414179346e-91
The Pearson Correlation Coefficient between TOEFL score and chance of admittance is 0.7915939869351047 with P = 3.6341021759974733e-87;

Since the P-Value is greater than 0.05 for all three cases, none of the cases are rejected as insignificant. Observing that the Pearson Correlation Coefficient is the highest between cumulative GPA and chance of admittance, GPA seems to be the most important factor assessing a student while GRE comes second and TOEFL last. However, for both GRE and GPA, not that a high GRE and high GPA almost guarantees acceptance whereas their significance is not as high for average and lower scores. It seems that when they are lower, the acceptance rate must be compensated with other factors such as letter of recommendation and statement of purpose, which were not included in this analysis.

Modeling

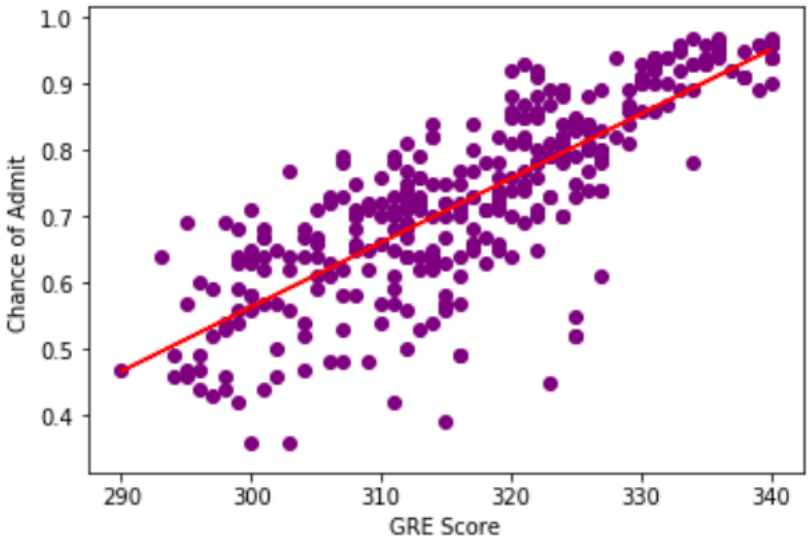
The next part of the project is modelling with the test data we already found to predict the prospects of future students. Training the relationship between GPA and chance of admittance returns a model with a r2 value of 0.54/

```
Coefficients:  [[ 0.20371613]]
Intercept:  [-1.02542191]
Mean absolute error: 0.06
Residual sum of squares (MSE): 0.01
R2-score: 0.54
```



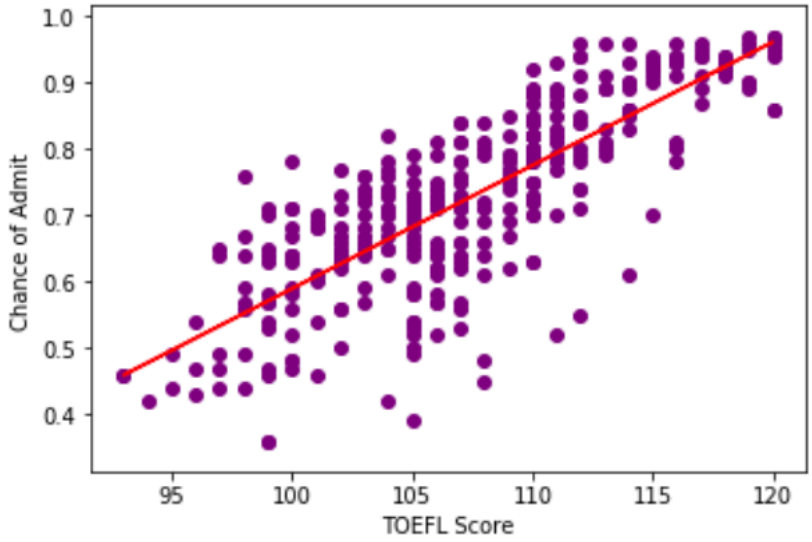
Training the relationship between GRE and chance of admittance returns a model with a r2 value of 0.36.

```
Coefficients:  [[ 0.00969772]]
Intercept:  [-2.34593557]
Mean absolute error: 0.07
Residual sum of squares (MSE): 0.01
R2-score: 0.36
```



Training the relationship between GRE and chance of admittance returns a model with a r2 value of 0.32.

```
Coefficients:  [[ 0.01858974]]
Intercept:  [-1.26971426]
Mean absolute error: 0.07
Residual sum of squares (MSE): 0.01
R2-score: 0.32
```



Conclusion

This project uses linear regression to analyze the correlation between independent variables, test scores and GPA, and the dependent variable, the chance of admittance into graduate school. One can conclude that there indeed exists a positive correlation between test scores and GPA with the chance of admittance where GPA is the most significant. This research indicates that if a student wishes to enter a graduate school, he or she must start prepping early to maintain a high GPA.