Chances of Graduate Admission

Dr. Sulove Koirala

4/6/2020

Introduction

The dataset contains GRE score, TOEFL score, GPA, SOP rating etc. of prospective graduate students. We are going to see which paramater might play a greater role in the chance of getting admission.

Loading Packages

```
library (ggplot2)
library (gridExtra)
library (ggcorrplot)
```

Loading and Cleaning the data

There is not much cleaning to do. We are just going to remove the Serial Number as it has no role in the analysis.

```
Admission_Predict <- read.csv("R:/Datasets/Admission_Predict.csv")
head(Admission_Predict)
```

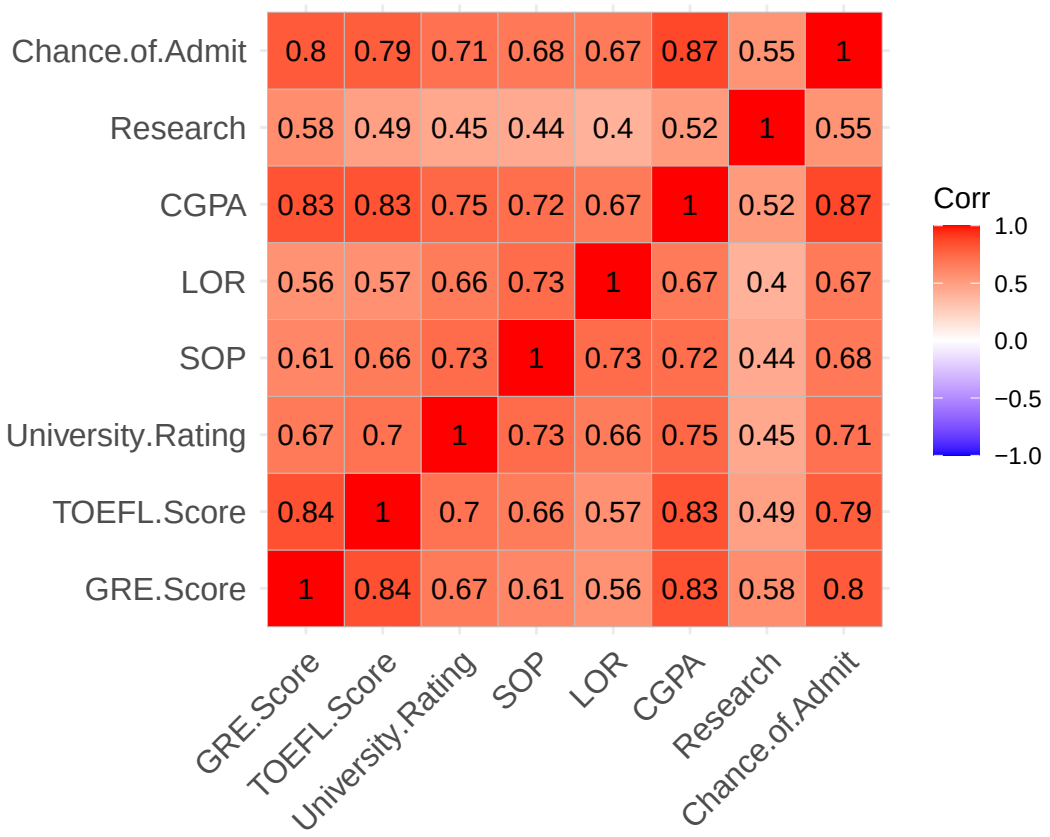
```
##   Serial.No. GRE.Score TOEFL.Score University.Rating SOP LOR CGPA Research
## 1         25       336         119                 5 4.0 3.5 9.80        1
## 2        144       340         120                 4 4.5 4.0 9.92        1
## 3        203       340         120                 5 4.5 4.5 9.91        1
## 4        204       334         120                 5 4.0 5.0 9.87        1
## 5         72       336         112                 5 5.0 5.0 9.76        1
## 6         82       340         120                 4 5.0 5.0 9.50        1
##   Chance.of.Admit
## 1            0.97
## 2            0.97
## 3            0.97
## 4            0.97
## 5            0.96
## 6            0.96
```

```
Admission_Predict$Serial.No. = NULL # That should do it
```

Correlation Plot

This should provide the straight answer to what we are looking for. We are going to use ggcorrplot, which will create a matrix of correlation.

```
admission = cor(Admission_Predict)
ggcorrplot(admission, lab = T )
```



It is obvious from the graph that there is greater correlation between Chances of Admission and CGPA. So, to have a higher chance of admission we need to have a good CGPA. Also, GRE and TOEFL plays a significant role. The least (although not insignificant) graded during selection seems to be Letters of Reocmmendation. Other things we can note here is higher GRE score is correlated with high TOEFL and CGPA.

Other Visualizations

```
gre = ggplot(Admission_Predict)+
    aes(GRE.Score, Chance.of.Admit, colour = GRE.Score)+
    geom_point()

Toefl = ggplot(Admission_Predict)+
    aes(TOEFL.Score, Chance.of.Admit, colour = TOEFL.Score)+
    geom_point()

CGPA = ggplot(Admission_Predict)+
```

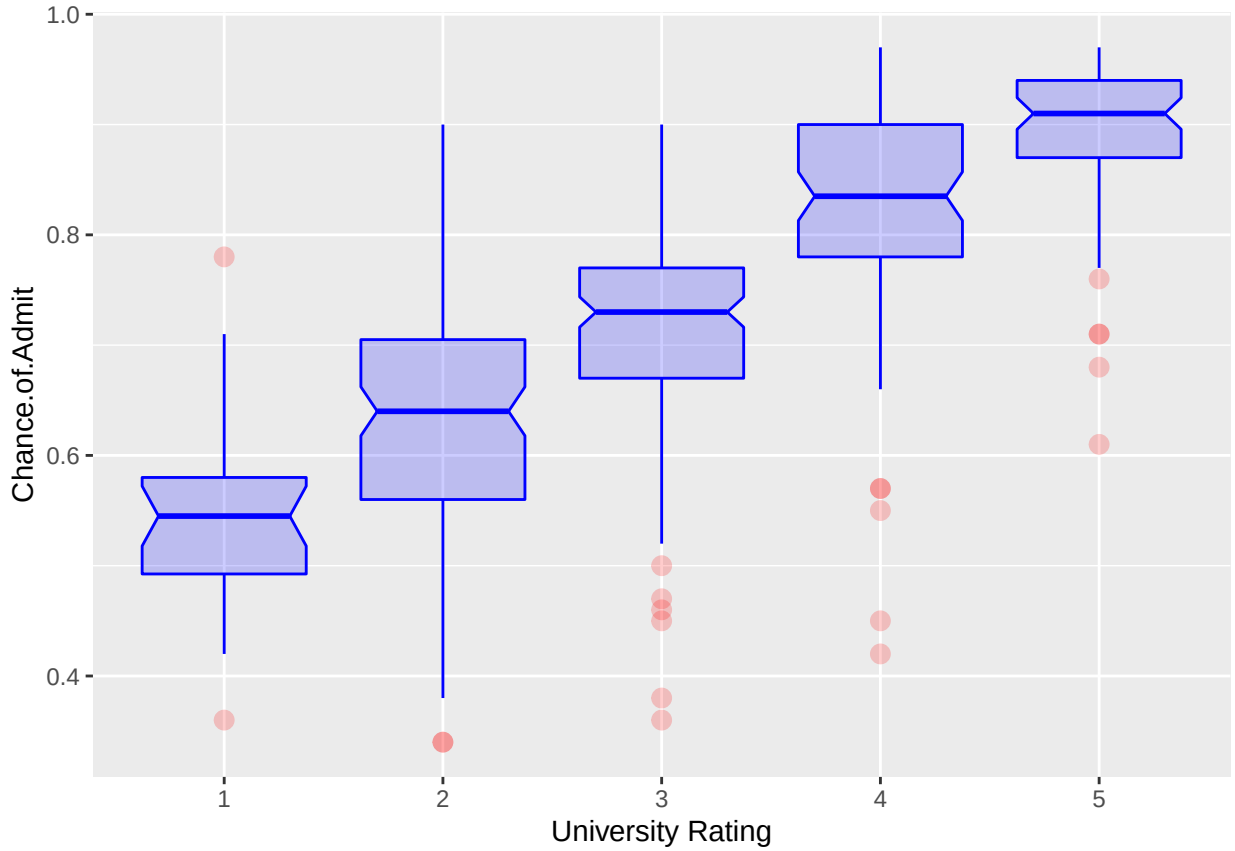
```
    aes(CGPA, Chance.of.Admit, colour = CGPA)+
    geom_point()

#We are going to combine the plots into one by using GridExtra package, which we have loaded in the beg

grid.arrange(gre, Toefl, CGPA, nrow = 3)
```

Finally, we visualize the rating of University responsible for the chance of graduate admission.

```
ggplot(Admission_Predict)+
    aes(factor(University.Rating), Chance.of.Admit, fill = factor(University.Rating))+
    geom_boxplot( color="blue",fill="blue", alpha=0.2,  notch=TRUE, notchwidth = 0.8,outlier.colour="re
    labs(x = "University Rating")
```



As expected, the chances of Admission in higher ranked universities are slim.

##Bibiliography

Mohan S Acharya, Asfia Armaan, Aneeta S Antony : A Comparison of Regression Models for Prediction of Graduate Admissions, IEEE International Conference on Computational Intelligence in Data Science 2019