Machine Learning Project 1, Due July 31

This is the first project for this class. For this project, you will work in groups of 3 on using regression to understand and make predictions about graduate school admissions. You will make an iPython Notebook and submit it to Canvas. In this notebook, you will answer the following questions. **You should elaborate on each question as much as you can using markdown (text) cells!** You will be graded on your presentation and your responses.

You should use the notebook `Regression_2020` as guidance. You will need to import the `Admission_Predict_Ver1.1` file from Canvas as well.

1. Make the names of every group member clearly indicated in the notebook, either in the title, filename, or a markdown cell at the start of the document.

2. Answer the following questions about the data.

   (a) Import your data. What are the features of the data?

   (b) Make 3 scatter plots of one feature vs. another. Be sure to label your axes clearly in each case.

   (c) Comment on each scatterplot. Which features are highly correlated? Does this make sense to you or not? Explain.

      (i) For those features that you think are highly correlated, back up your argument by finding the correlation coefficient.

   (d) Which features do you expect will not be well represented by a linear fit?

3. Make a prediction for the chance of someone to be admitted to graduate school assuming that they have a GRE score of 315. You should use what you have learned in this class to do so.

4. Answer one of the following questions (your choice). You can answer both for extra credit!

   (a) Make a prediction for the chance of someone to be admitted to graduate school assuming that they (1) have a TOEFL score of 109 and (2) a GRE score of 325.

   (b) Assume that you have a limited amount of time to study for the TOEFL exam and the GRE exam before applying to graduate school. Right now, your estimated GRE score is 300 and your estimated TOEFL score is 100. Which exam should you spend more time studying for? Should you study for only one of the exams, or both? Discuss and back up your arguments **using what we have learned in this class**.

5. If you have any other ideas or discussion, implement/discuss them for extra credit!