Abstract: Cross Validation of Data Simulations

Using a provided data set containing one original data matrix and 99 other data matrices simulating the contents of the original data, I perform a 5 fold cross validation of the data to determine if differentiating the original data from the simulation data is an adequate classifier. I evaluate the efficiency of this classifier by analyzing the mean AUC value across all simulations and modeling this information to determine its significance.

In this section you will find my code used to achieve this objective as well as a summary of my findings and analysis using the model I produced. Each of these are contained in separate files, **Code** and **Report**, on this github.