Team: Dominic Klusek, Jonathan Rozen

Team Contributions:

Choosing a dataset was a team effort and we both went back and fourth between dataset; main concerns were finding a dataset with features that had a linear and polynomial correlation like the dataset in the part 1. Ultimately we both agreed on the SkillCraft and Credit Risk datasets. The code for Part 1 was worked on by Dominic, and Part 2 was worked on by Jon. For analysis we both worked on it together. After running the code Dominic suggested that we add resampling to Part 1 and testing the effect of maximum tree depth for the RandomForestClassifier; Jon also suggested KFold Cross Validation to reduce the volatility of the results that were being obtained for Logistic Regression for Part 2.