Evaluation of Model Iteration 1

*Date: 08/03/2021*

**Work Completed**

*Filename: Model Iteration 1*

A comprehensive evaluation of the importance of different NFL statistics as performance indicators has been conducted through a review of current literature written by industry professionals. Two of these statistics, ‘First Down Yards’ and ‘Sacks’, are included in the first model iteration.

We have created the ‘Model’ class in Python. This takes the results of games as well as two predictive statistics in the form of a Pandas Data frame as an input. The class creates a rating for each team, which can be interpreted as the score they are likely to contribute to a game. The creation of the ratings is done using simple ridge regression techniques through the Scikit Learn module. Three different methods for producing the ratings have been included in the class, and the user is given the option to choose which is employed.

The Model class has a ‘Winner’ method which outputs the predicted score for a game between to inputted teams. An ‘Accuracy’ method provides statistics concerning the model’s performance.

**Evaluation of Method**

The model shows predictive potential. However, the method used to produce it has some major flaws. Most notably, no model diagnosis is conducted. We have overlooked the possibility of ill-conditioning, and other assumptions of the linear model. Poor understanding of ridge regression is also a factor which may be hindering performance.

The use of ‘First Down Yards’ and ‘Sacks’ in the model added little to its performance. I have decided to focus on perfecting the scores aspect of the model in the second iteration.

**Second Iteration Plan**

Before a second model is coded, some knowledge gaps need to be filled. First, a comprehensive understanding of the Scikit module needs to be gained to use it correctly. It is recommended an online course is taken to fill this requirement. Also, revision of multiple linear regression, and research into ridge regression needs to be conducted. A statistics textbook should suffice to fill this requirement.

The second model file should be properly commented, and the model properly diagnosed. This of vital importance, as it was one of the original goals of undertaking this project.