**Data**

The data set consists of all games played between the 2010-11 season up to the 2018-19 season. This included data for roughly 11,400 games played during this time. This data is publicly available from the NHL.

**Preprocessing**

To ensure the data is functional for the model, cleaning of the data is necessary to remove null values and some normalization of data will provide proper analysis. This data set was processed previously, containing no null values; the features, however, did need to be normalized as their ranges were spread out. While processing the data, it was clear that two features, power play opportunities and hits, were not significant to determine if a team will win the game. These features were then removed to help the overall capability of the model.