Capstone Two Project Proposal

Problem Statement

Context: In order to make a profit, casinos need to accurately forecast the probability of a team winning its game. If the game spreads are not accurate, consumers will be able to exploit the inaccurate spreads to win money against the casinos, thus costing the casinos potential profits. Accurately predicting the chances of a team winning are vital for casinos to remain profitable

Criteria for Success: My model will be considered a success if the percentage of times the favored team wins falls within the 99% confidence interval for our predictions (see below for calculations).

99% confidence interval = mean probability of favored team +- standard error of mean * 2.595 (2.595 is critical value for 99% confidence with 255 degrees of freedom)

Scope of Solution Space: The scope will purely be the probability of an NFL Team winning its game. The project will try to find the most impactful statistics and figure out how influential the statistics are.

Constraints Within Solution Space: The biggest constraint is the lack of available data. Due to the nature of the NFL, a lot of in-depth data is proprietary so my project will mostly use more basic and public data.

Stakeholders to Provide Key Insights: Any manager or supervisor who oversees the NFL Game spreads

What Key Data Sources Are Required: The main and primary data source will be NFL Team Stats (2002-2019). If needed/applicable, I may also use NFL Statistics. Both data sources can be found on Kaggle.