

Predicting Outcomes of NFL Contests

Using regression techniques to find betting market value

Introduction

Every NFL contest has a number of ways for fans and bettors, both recreational and professional, to get involved in attempting to profit from the game. Such ways include picking players from the game for a fantasy football lineup, betting on a player to perform well or poorly, betting on a team to win the game, or betting on the two teams in the contest to score above or below a certain number of points. The motivation of this project is to explore ways to implement machine learning into NFL contest outcomes with a goal of profiting from the NFL betting market.

Problem Statement

The aim of this project is to take a mathematical approach to betting on NFL contests. We want to predict the total number of points scored between two teams playing against each other in any particular NFL game. We will likely be using regression techniques for numeric predictions, and will be comparing our predictions to the game totals that professional sports books put out for the public to view and bet on. Comparing our predictions to the betting market will paint a clear picture on which games present us with an opportunity to profit. Once we have identified a game or games with glaring values in comparison to the betting market, we can recommend these plays to a potential client, friend, or even play them ourselves.

Data

Attribute Selection

The attributes that will be used for this project will mainly be NFL team statistics such as passing yards per attempt, total yards per play, offensive rating, defensive rating, and points allowed per game. We will also need weather and venue data, as NFL games are played in many different types of conditions that would seem to affect the game. Referee data would be another important aspect of NFL games that is worth collecting and exploring, as all referees have their own style of officiating that most likely affects game totals. Another important piece of information is data on game totals, which will tell us what values the sports books set the betting market for each game, and can serve as a marker from which to compare our predictions and make recommendations.

Obtaining the Data

For all of the attributes described, there are many sources from which to obtain the data. One source is pro-football-reference.com, which has a database that contains nearly every relevant football statistic for every NFL game played dating back to 1920, which includes weather, venue, and referee data. Another source for obtaining the needed data will be fantasydata.com, which has defensive and offensive rankings for every team in each contest, and data from sports books, which includes the totals for each game.