

Predicting Fantasy Football Production Leveraging Domain Expertise

Motivation

Over the last decade, fantasy sports have gathered a significant following. Recently, gambling on fantasy football has taken off – in particular, weekly competition sites such as FanDuel have become incredibly popular. These weekly leagues provide each user a bundle of \$60,000 in-game currency with which to purchase players in order to set their line-up for weekly betting competitions. The salaries are tied to the average fantasy points per game generated by the player – i.e., a player like Tom Brady, currently averaging 28.4 fantasy points per game, has a salary of \$9,100, whereas someone like Chase Daniel, yet to play this season, has a salary of \$5,000.

Given this, the goal is to produce regression models for all fantasy positions to find an exploitable market inefficiency. Particularly, we want to create predictive regression models for: Quarterbacks, Running Backs, Wide Receivers, Tight Ends, Place Kickers, and Defensive Units, to beat the market, as defined by FanDuel, and gain an edge in fantasy betting.

Data

We will be scraping historical fantasy point production data under FanDuel's scoring system for each position from RotoGuru.com, a site that tracks weekly salary and production values back to 2011. Similarly, we will be scraping actual yardage, snap counts, turnovers, and touchdown production for all players from the 2011 season on from the NFL's statistics portal.

However, noting that each team is different and weekly match-ups play an incredibly important role for a player's production, we will also be gathering data for the team a player belongs to and their weekly opponent. For this, we will use Football Outsider's metrics for player rankings, defensive and offensive team rankings, and team efficiency rankings.

Features

For offenses:

- Primary box score statistics for
 - last X games (i.e. yardage gained, touchdowns scored, turnovers committed, yards per attempt) – where X will be at most 16 games, i.e. a season's worth.
 - opposing defense over last X games (i.e. yards allowed, turnovers caused, sacks, tackles for loss)
- Current Football Outsider(FO) ranking for
 - player
 - supporting cast
 - opponent's defensive units

These features will be adjusted depending on the specific position group being modelled (i.e. we don't care much about fields goals scored for running backs)

For defenses:

- Primary box score statistics for last X games
- Current FO rankings for
 - individual members of defense
 - total defense
- Opponent's primary box score statistics for last X games
- Opponent's FO rankings for
 - individuals
 - total offense

Evaluation

Since the current NFL season is underway, we have a weekly opportunity to test our predictions. Moreover, due to the popularity of fantasy sports, most sports journalism sites provide their own prediction models for fantasy players. We will be evaluating our models not only in terms of how well they perform on the 'live' test each week, but also how they compare relative to commercial models currently in use.