

CS 270 Group Project Proposal

Carter Larsen (clars32@byu.edu)

October 8, 2024

"Predicting Weekly Fantasy Football Scoring Output"
My Interest Level: 10/10

1. Description

Fantasy football is a popular game that allows competitors to assemble their own virtual team of real-life NFL players. These players earn points for their team based off of their real-life performance (e.g. 4 points per passing TD, 1 point for every 10 rushing yards, etc.). Teams are matched up against each other, and whichever team finishes with the higher total score is said to have won the matchup that week. This continues over the course of a season until a league champion (and league loser) is crowned.

This project aims to use machine learning to predict the fantasy score of an NFL player on a given week. Using the wealth of data publicly available both as actual NFL stats and expert predictions specific to fantasy football, the model will aim to use past performance and current projections to create a reasonable estimate for the fantasy score of the player.

2. Possible Features

Some possible features the model could include are:

- Fantasy points scored the previous week/3 weeks/season/etc.
- Average points given up to the player's position by the opposing defense
- Expert projected points for that week
- If the player is playing at home or away
- Average points scored the previous season
- Individual rushing/passing yards, touchdowns, etc. instead of total points

Here is what an example of an instance could look like:

SeasonAvgPts	Last3WeeksAvg	LastWeekPts	ExpertProj	Home/Away	Target: Fantasy Points
14.445	15.04	13.24	15.37	A	8.40

3. Gathering Data

Data for both player stats and fantasy projections is widely available. For example, pro-football-reference.com has the complete game log for all NFL players, including all stats (features for the model) and fantasy points scored (the targets for the dataset). fantasydata.com provides a multitude of player projections, including historical data that can be used in training. The main task in gathering and labeling data would be to scrape the data from the different sources, combine it all into a usable dataset, clean it, and normalize it for use. Dataset size should not be a problem as data exists for the top players over the last several years.