Fantasy Football Predictions

Vincent Broda(vbroda@vols.utk.edu), Dillon Frankenstein (dfranke2@vols.utk.edu),
Fort Hunter(lhunte21@vols.utk.edu), and Austin Smith(zyr546@vols.utk.edu)

***Abstract*— Picking the best fantasy football players is high stakes for many individuals. With the rising popularity of punishments for the worst players in each league, it is more important now than ever to pick the best team. The goal of this paper is to outline a machine learning model that can accurately predict the best fantasy players for each week.**

I. OBJECTIVE

The main objective of this project is to do an analysis of historical NFL player data to make a machine learning model that can predict the top fantasy football performers for each week. We plan to achieve this by obtaining game stats for each current NFL player and defense. Our current plan is to use a dataset from kaggle. However, if we find that this project takes less time than expected, there are some APIs that we could use that would make obtaining the data a bit harder and give us more to do. There is also the possibility of us scraping the data from somewhere online such as ESPN's website. After we obtain the data, we will clean it up by taking out players that we don't need. This will likely be defensive players, non-skill players, and retired players.

Once we have cleaned the data and have it formatted the way that we need, we will apply this data to various machine learning models to try to predict the top fantasy performers each week. The model will take into account historical player data, the strength of the defense that the player is going against, and how the teams have fared against each other in the past and vice versa for predicting the best defenses each week. Once we find out what is working the best we will spend the remaining time fine tuning that model to try to achieve the best results.

By the end of the semester we hope to be able to be as close as possible to predicting the best performers each week. Since the NFL season is currently going on we will be able to see how well our model is working almost immediately.

II. MOTIVATION

Selecting fantasy football players each week holds a great deal of risk and can quickly become overwhelming for some players to do successfully. We think that augmenting fantasy football players with a machine learning model can greatly improve safety and chances of success. Helping players be safer and more confident in their choices can increase enjoyment as increased success rate is expected. Furthermore, most fantasy football players are a part of fantasy football leagues, which often carry punishments for any players with underperforming rosters. Helping players be more competitive with the help of our model means that these sometimes grueling punishments are avoided, which will help players' mental health.

III. DATA

Kaggle NFL player stats data includes important information for fantasy football positions such as QBs, WRs, RBs, TEs, Kickers, and Defenses. This data can include catches, yards, touchdowns, turnovers, field goals they are playing against. We would also need the data of the defenses the players are playing against like how many catches, yards, touchdowns, turnovers, and field goals allowed. We can also find non-football statistical data that can affect stats, such as different stadiums' statistical data, injury data, or weather data.

IV. TECHNOLOGIES USED

There are a myriad of data processing and machine learning algorithm python libraries that are available these days. We plan on utilizing sci-kit learn to choose our machine learning algorithm. Since we have a great deal of good data we will be opting to create a supervised model. As for the algorithm, we plan on testing performance of a few different implementations to ensure the best model. Additionally, we will be using some data processing libraries such as pandas for its ease of use and modularity. We also plan to use Google Collab because of its ease of use and its free computation. As well, we plan on using an API to get more recent player statistics in the later stages of the project.

V. RESPONSIBILITIES

- Fort: Data collector. Will be responsible for collecting the raw data and coming up with the methods to store it.
- Dillon: Data Cleaner. Will be responsible for cleaning up the data. This includes removing unnecessary stats and players and formatting that data in a way that will be compatible for use with the model.
- Vincent: Model Creator. Will be responsible for making the machine learning model to calculate statistics and return an accurate fantasy score prediction.
- Austin: Model Optimizer. Will be responsible for making sure the machine learning model is as effective and fast as possible.

VI. MILESTONES

Firstly, we plan on cleaning up the Kaggle dataset. Then we want to make and find an initial model. Our goal is to have this stage of the project done by mid to late October.

Next we want to work on using an API to get more advanced and recent data and then update the model with this new data. We hope to see several improvements in this iteration over our initial stage. We plan on having a final model by the end of November.

VII. Expected Outcomes

The model created through our process is expected to predict the most successful NFL scorers per week. This should play into the greater goal of our model being able to make an optimal layout to promote the most successful squad. We are confident that this approach will be successful as there is a great deal of highly relevant data that is made available by the NFL. Furthermore, our greater goal of creating an optimal lineup will be a great way to judge overall model performance. Having enough modularity in our model that lets us accurately predict a single player's performance and a team as a whole will be paramount. Overall, we expect to create a machine learning model that can greatly enhance any fantasy football player and their chances of winning.

References

[1] Funk Monarch, "NFL stats 2012-2023," Kaggle, https://www.kaggle.com/datasets/philiphyde1/nfl-stats-1999-2022 (accessed Sep. 26, 2024).