**NFL Fantatsy Football Team Selection**

Principles of Data Mining

Professor Thomas Kinsman

McKay Clawson, Patrick Gleason, Kristopher Hopkins

Table of Contents

Introduction: 1

Background: 2

Descriptions and Relevance: 2

Solution: 2

Results: 2

# Introduction:

In the late 1990’s online NFL fantasy football was born. Since then the game has become one of the most popular friendly competition between football fans. Co-workers, family members and casual friends gather year after year to draft a team that they hope will finally carry them to this years championship game. Fantasy football is so popular that many people draft multiple teams, pay large league entry fees, and religiously check statistics. FX has even made a very successful TV show around the concept of friends competing to a fantasy football league. It would appear that there is no end in sight for the fantasy craze that has captivated America, especially considering the fact that NFL is enjoying an all time high in popularity and television ratings. More and more we are seeing the introduction of advanced statistics to draft strategies. Looking at complicated metrics like QBR (Quarter-Back-Rating), YAC (Yards-After-Catch) and many more are quickly becoming the modus operandi for serious players. With all of the advanced data available, data mining should be able to help players take the step to the next level. One of the hardest parts of creating a competitive team is the actual process of drafting the team. Year after year we see players get drafted in early rounds of drafts across the country only to see them have lack luster years. We think that using data science we can help alleviate the pain of drafting players that have mediocre years. The basic idea is to use a number of different clustering and classification techniques to make a simple decision: given a player name and a round number, should you draft that player.

# Background:

Fantasy football was officially created in 1962 when the first rulebook was publically published. For the next 3 decades, the game enjoyed moderate success among the football faithful. One of the biggest reasons fantasy football never reached the popularity levels the game enjoys today was how difficult it was to play the game. The draft was a manual process, where people would sit around a table and take turns selecting players. From then on players would lock in their lineups before the first game by submitting the players they thought would do best to the league commissioner. Then on Tuesday morning everyone would look at the sports section of the newspaper, find all of his or her players and look at the stats to determine how many points they had earned that week. The advent of the Internet changed the game as we know it. In 1997 CBS Sports launched the first online client and fantasy football was born. Using different websites such as ESPN, Yahoo!, or CBS Sports makes the game much more fun because it removes all of the manually intensive labor from the game. The website will track the stats and scores, manage teams, adjudicate trades, and create playoff brackets. The game has evolved so much that there are many different rulebooks and types of leagues. We are focusing on a standard league where you are allowed to have three types of offensive players: passers (Quarterbacks), receivers (Tight Ends and Wide Receivers) and rushers (Running Backs). Quarterbacks are the most valuable players on the team by the nature of the position. They score 4 points for each touchdown pass, 1 point for every 25 yards they throw, and lose points for various missteps (fumbles, interceptions, etc.…). Receivers score 6 points for every touchdown they catch and 1 point for every 10 yards they catch. Rushers are very similar, 6 points for a touch down and 1 point for every 10 yards they run for. With the understanding of how the game is scored, the draft is very simple. There are 10 rounds and each team gets to pick one player a round. The players you pick on draft day will make up the majority of your team for the rest of the season (you can do things midseason to improve your team, such as trading, but for simplicity we are ignoring these possibilities). It is not an understatement to say the draft is the single most important day of the season. With that understanding it is easy to see why we want to accurately determine the round players should be drafted.

# Descriptions and Relevance:

This problem relates to many topics in the course. The first thing that solving this problem does is demonstrate what students can do after taking the course. With an understanding of different clustering algorithms, classifiers, and attribute selection techniques students can find their own datasets and create models that are useful to them. One point that was stressed in the classroom was just how much time data cleaning takes up. We spent 3 hours sifting through gigabytes of data trying to find the most useful dataset. Once we had determined the most useful dataset there were many problems with the data. There were missing entries, entries that were supposed to be integers that had string values and extraneous rows. We spent another 2 hours cleaning all the data we wanted to use. From that point we had to determine what target variable we wanted to predict. When we realized that the dataset we had collected didn’t have the appropriate target variable we had to go and find that variable from other sources and then manually insert it into our data files. The one downside of selecting fantasy football teams is that this class won’t use it most likely in the future. It is an extremely interesting problem space with many nuances and deviations that have to be considered, but you have to have an immense amount of domain knowledge and a passion for the game in order to appreciate it. Outside of this class I think an implementation of our solution could be valuable. Many people play fantasy football and it would stand to reason that a company like Yahoo! Or ESPN would pay dearly for algorithms that consistently produce better results.

# Solution:

# Results: