Identifying Undervalued Players in Fantasy Football

Christopher D. Morgan
Southern Methodist University, cdmorgan@smu.edu

Caroll Rodriguez
Southern Methodist University, carollr@smu.edu

Korey MacVittie
Southern Methodist University, kmacvittie@smu.edu

Robert Slater
Southern Methodist University, rslaterr@smu.edu

Daniel W. Engels
Southern Methodist University, dwe@smu.edu

Identifying Undervalued Players in Fantasy Football

Christopher Morgan, Caroll Rodriguez, Korey MacVittie, Robert Slater, and Daniel W. Engels

Master of Science in Data Science, Southern Methodist University
6425 Boaz Lane, Dallas, TX 75275 USA
{cdmorgan, carollr, kmacvittie, rslater, dwe}@smu.edu

Abstract. In this paper we present a model to predict player performance in fantasy football. In particular, identifying high-performance players can prove to be a difficult problem, as there are on occasion players capable of high performance whose past metrics give no indication of this capacity. These "sleepers" are often undervalued, and the acquisition of such players can have notable impact on a fantasy football team's overall performance. We constructed a regression model that accounts for players' past performance and athletic metrics to predict their future performance. The model we built performs favorably in predicting athlete performance in relation to other models, though this performance is heavily reliant upon the accuracy of estimates of athletes' workloads.

1 Introduction

Football and the NFL are a cherished part of American culture that has become a significant point of enjoyment for millions. One way for fans of the game to further participate in the culture is to play fantasy football, a game that uses athlete results from live football games to determine a "fantasy" team's performance. Due to significant accessibility improvements, over the past decade there has been an exponential growth in fantasy football participation rates [3].

Research has been performed on how fantasy football influences fan viewership of actual NFL games, and it has appears that a symbiotic relationship exists between both fantasy football and NFL game viewership rates where each benefit one another [8]. One indication that this relationship is truly positive is that a culture and language around fantasy football has developed. Participants have developed a wide variety of terms to describe the game showing a significant level of cultural development around the game [2].

With the rise of fantasy football, fantasy football players ("participants") have sought out ways of gaining advantages over their competitors. One of the more common routes is through unique athlete selection strategies, with participants looking for ways to optimize weekly lineups to ensure their team produces the most points. While there is much research and many tools available to gain insight into fantasy football strategy, there are often conflicting results when attempting analyze season data from year to year. To mitigate these variables, participants seek out what are known as "sleepers," and have the potential to make or break a fantasy football season [16].

"Sleepers" are high-value hidden talents, who may have been performing modestly in past seasons, spent significant amount of time as backup, or may simply be rookies with no prior history in the NFL [5]. Consistently each year, a few athletes emerge who become dominant forces in their position, and being able to accurately predict the likelihood of these athletes emerging could present a significant advantage to fantasy football players. Predicting such athletes, however, is volatile, and while there is much interest in the fantasy football community in predicting them, there are many cases in which a given athlete fails to meet expectations. While analysts with significant domain expertise can sporadically identify athletes with the potential to be a "sleeper," in general the results are lackluster and unpredictable.

Our hypothesis is that there is some combination of key factors of an athlete that would identify them as a potential exceptional player, but are otherwise overlooked due to lack of experience or other factors. To that end, we examine various physical, statistical, and environmental factors for athletes under consideration, which may influence a given athlete's abilities that would make them appear to be average or mediocre, but - when taken in aggregate - actually indicate an exceptional level of skill that can appear under the proper circumstances.

The purpose of this analysis is to construct a model capable of accurately predicting an athlete's performance, based upon athletic metrics and historical performance, to provide a way of determining the athlete's value. This model would allow players to predict a given athlete's performance in an upcoming season based on historical data.

Due to the nature of the sport and differing approaches to finding such hidden talent, there is much variance in how a "sleeper" is defined. In this work we define a sleeper as an athlete whose average draft position occurs after the first four rounds in a fantasy draft, whose performance in that draft's season puts them in the top twenty athletes for their position: this is a working definition and is subject to position-specific variance.

Our analysis includes examination of several of the complex relationships and player attributes that lead to a given athlete emerging from fantasy football irrelevance to dominance, as well as how an attempt to accurate predict said emergence. Our work specifically focuses on wide receivers and running backs, as other positions include factors that further complicate analysis, have much simpler analysis, or rely much more heavily on the team as a whole.

The wide receiver position has been described as the "most complex, compelling, and talked-about position in all professional sports" [4]. A successful running back will have speed, quickness, tackling power, and the ability to make important plays [13]. It is our position that we can predict the emergence of high-performance players who were otherwise underestimated or overlooked, in the case of wide receivers; and that - for some roles - there exists a knowable combination of physical characteristics that can make the "sleeper" status predictable, in the case of running backs.

While we focus our analysis primarily in the realm of fantasy football, we are of the opinion that the results generated here could also be used for actual NFL scouting and analytics as identifying high performing players who might soon become available as draftable rookies or free agents to be signed could present a significant advantage to NFL teams. It is also possible that the results we generate here may provide the groundwork for similar analysis for other physically-demanding work, including other sports or possibly other professions.

2 Understanding Fantasy Football

2.1 History

The initial concept of the first fantasy football league was created by three men in a Manhattan hotel in 1962. Bill Winkenbach was the well-known co-owner of the Oakland Raiders, based out of California. Scott Stirling worked as a newspaper columnist for the Oakland Tribune, and Bill Tunnell was the Oakland Raiders' public relations person.

The first league consisted of an eight-man team that was called the GOPPL (Greater Oakland Professional Pigskill Prognosticators League). In its inception, the owners of the team were various people in the Oakland Raiders' organization [9].

2.2 Main Rules

Like the NFL, fantasy football is a season long competition played online, and both have the goal of accumulating the most points to win each game. However, point acquisition in fantasy football is based upon the performance of athletes in NFL games, rather than a direct correspondence: for example, a player whose actions do not result in points in the NFL game may still yield points in fantasy football. Each week of the NFL season, statistics are gathered for each athlete, which are then fed into the fantasy football leagues, whose games are determined by the statistics of the athletes involved. Fantasy football players draft individual athletes at the beginning of the season, with their overall goal being to pick the athletes who will net them the most points in each game [1].

2.3 The Draft

Fantasy drafts can take several different forms. Initially drafts were performed by hand in person by sports enthusiasts. With advancements in online fantasy sports, nearly all drafts are completed using some type of online tool. The most common type of drafts are the auction, linear, and snake drafts.

Each draft has its own nuances, but snake and auction drafts are typically seen as having the most potential for fairer outcomes, as linear drafts put later pick order teams at a significant statistical disadvantage, and are becoming increasingly phased out of fantasy football leagues.

Auction Draft In this format, each player in the draft selects an athlete, after which each player can choose to place a bid on the athlete. Once the bidding is complete, the athlete goes to the player who bid the most.

Snake Draft In this format, players take turns drafting an athlete in a set order. When the last player has drafted, the draft order is then reversed: whoever drafted last in the first round now drafts first, and whoever drafted first in the last round now drafts last. This process continues until the draft process is completed.

Linear Draft In this format, players take turns drafting an athlete in a set order. Unlike in snake draft, however, each round follows the same order: a player's draft position is set and does not change each round.

Regardless of the particular type of draft, a drafting strategy is essential to ensuring that a player can produce a team of athletes who will perform well. Athletes with low perceived value will often be drafted later: if a given player can estimate athletes' performance, and can identify undervalued athletes, that gives the player a significant advantage. Players need to take many factors into consideration when considering how to build a team: for instance, a player with a more balanced athlete selection may out-perform a player who simply over-selects athletes at a specific position.

While these considerations are important to an overall drafting strategy, these concerns are beyond the scope of this work, as we focus on identifying undervalued athletes.

2.4 The "Sleeper" Phenomenon

The root cause for the "sleeper" phenomenon seems to be attributable to a diverse set of causes, which interact on a seemingly case-by-case basis to produce exceptional results. Each of the possible factors provides a reasonable explanation upon inspection: poor coaching, poor team support, mediocre college-level performance, and natural progression are often cited as reasons for a given athlete being underestimated.

Performance measurements used in football have been extracted from player statistics, and a study determined which of those factors are most responsible for the success of a given team, with the results suggesting that defensive players have a stronger impact on results than offensive players [10].

Combined with our research, these studies provide a framework against which we construct models for building optimized draft selection lists.

For some fantasy football leagues, participants are given money with which they purchase players, and it is in this context that being able to identify "sleepers" becomes quite important. If a given athlete is undervalued, but his performance is outside the bounds of expectations for the indicated price, this can give the team builder an edge [15]. Similarly, in Daily Fantasy Football (DFS), weekly drafts are performed using an auction system in place of a more traditional draft, which athletes are given costs: in such scenarios, the optimal strategy may be to identify “sleepers” so that a player can acquire athletes who are undervalued but have a higher “true value” than would be indicated by their listed cost [6].

This suggests that the same may hold in football, so we examined the source of this increased performance, with the intent being to find a means for a player to identify “sleepers” based upon a variety of metrics.

3 Data

Our data consists of football athlete data taken from 2012 to 2018, which consists of a combination of athletics metrics (such as from training camps) as well as metrics from games those athletes participated in.

This data is then divided into two categories, running backs and wide receivers. Each group has a different set of metrics that are important to that role. In addition to role-specific metrics, there are a variety of metrics relevant to both those roles as well as other athletes.

The metrics for the roles are outlined in Table 1, Table 2, and Table 3.

3.1 Data Source

To begin understanding the problem and quantifying different player attributes, we began by seeking out one of the more comprehensive data source for fantasy football which is the *playerprofiler* website[1].

playerprofiler is often utilized as an “on-the-fly” data source for visual breakdowns of individual athlete performance and metrics. For the purposes of our research, we purchased access to the site’s full data library to begin modeling and studying factors for identifying high potential athletes. While we will be incorporating other data sources, this will be our primary source for our analysis.

Data for each athlete corresponds to their position, and therefore running backs and wide receivers will not share entirely the same metric list, but there will be some overlapping attributes. Each dataset is designed such that each row represents a unique athlete, with each column representing an attribute or play statistic from a selected NFL season. After extensively pruning unnecessary columns and those with missing data, we are left with each dataset containing roughly 300 attributes.

[1] https://www.playerprofiler.com/data-analysis/

Table 1. Shared Metrics

Metric	Description
Age	Calculated Age of Player
Position Type	Specific versions of Running Back and Wide Receiver in NFL (i.e. Full Back, Half Back, Tail Back, Grinder, Flanker, Split End, Slot back, No Class)
Years Experience	Total number of years in NFL of Player
Targets	How frequently the ball is thrown to the player
BMI	Body Mass Index of the player
Draft pick	Overall draft pick number for player's draft year
Height (Inches)	Height converted to inches only
Weight	Players weight in lbs
Arm Length	Players arm length in inches
Handsize	Players hand size in inches
20-Yard Shuttle	Drill time from NFL combine
3-cone drill	Drill time from NFL combine
40-yard dash	Drill time from NFL combine
Agility Score	The sum of a players 20-Yard Short Shuttle time and 3-Cone Drill times
Athleticism Score	Aggregate of 40 yard dash, burst score, & agility score normalized for size (created by player profiler)
Bench Press	Number of bench press reps performed at 225 lb presses during the combine
Broad Jump	Number of inches player jumped in Broad jump
SPARQ-x	An approximation of Nike's proprietary SPARQ athleticism scoring system
Burst Score	Sum of a player's vertical jump height & Broad jump distance equally weighted (provided by player profiler)
Vertical Jump	Player Vertical jump at the NFL combine
Dominator Rating %	Outlined by Frank DuPont. Represents player's percentage of the team's offensive production. Historical results have show: 35% level as elite; 20-35% as high potential; $<$ 20% as point of concern.
PBLK	Grade for Pass Blocking of Team
RBLK	Grade for Run Blocking of Team
DEF	Grade for Defense of Team
RDEF	Grade for Run Defense of Team
TACK	Grade for Tackling of Team
PRSH	Grade for Pass Rush of Team
COV	Grade for Defensive Coverage of Team
SPEC	Grade for Special Teams
PASS	Passing Grade of Team.
Had A Breakout	Fantasy football player who has been ineffective or marginal in the past, but who suddenly has high performance

Table 2. Running Back Only Variables

Metric	Description
Carries	Number of times a rushing player attempts to advance the ball.
College YPC	Measures a running back's averaged yards they carried the ball during a rushing attempt.
Speed Score	(weight*200) / (40-timeˆ4). Formula developed by Bill Barnwell to predict running back success. Biased towards heavier running backs.
College Target Share	Measures the percentage of all passing targets directed at a particular wide receiver or tight end.
RECV	Receiving Grade of Team.

Table 3. Wide Receiver Only Variables

Metric	Description
Height-Adjusted Speed Score	Player speed, accounting for average WR height.
Catch Radius	Calculated by squaring a players 40-yard time, 20-yard shuttle, and 3 cone time, multiplied by the square of a player's height, arm length, and vertical jump.
College YPR	Passing yards per reception of Player during college years

While many common football statistics are included in our dataset for each season dating back to 2013, we have detailed a small selected portion of attributes whose meanings would likely not be obvious to readers unfamiliar with athletic metrics. Table 1 list metrics that appear in both the running back and wide receiver datasets, whereas tables 2 and 3 represent attributes unique to running backs and wide receivers, respectively.

Given that players may change teams, it is critical to take into account the effect of possible new opportunities afforded a player based upon the change in their support. To do this, additional data was pulled in from the NFL official stats[2]. This further allowed us to make a deeper connection between players on a given team. To do this, we tied the quarterback's performance to those of the running back and wide receiver, to attempt to account for the effects of the team, which also required our use of corresponding seasonal quarterback statistics.

While quarterback performance can influence the performance of those positions, the performance of the team as a whole can play a significant role in influencing player performance. To account for this, it is necessary to quantify, at least roughly, the performance of the remainder of the team. We hypothesized

[2] http://www.nfl.com/stats/

that the offensive line strength could be a strong influence on performance, but there are few consistent methods in extant works for accounting for this. While stats such as "pancakes" - which indicates a player knocking a defensive player to the ground and out of play - are somewhat indicative of performance, many play successes can be dictated by simply containing a defensive rush for a fraction of a second longer than what other teams may be capable of.

Pro Football Focus[3] has developed a black box method for scoring different components of the team, based on sampling and reviewing each players performance on a play, and whether they achieved what is the perceived as the desired outcome of the play. Scores are given in positive and negative values for each play for each part of the team. These individual scores are compiled to create a team component score for different parts of the team and their activities, such as offensive line rush blocking, pass blocking, all measured on a scale of zero to one hundred. Admittedly, this is not ideal for all purposes, in that it is a black box for scoring team components - however, pulling in similar items such as offensive line rankings would not have to be explainable as they are often assigned by arbitrary judgments from writers.

3.2 Data Preparation

One challenge presented by the data we used is the significant amount of missing data for both the running back and wide receiver positions. In football, it is natural for some players to have missing data in terms of seasonal performance, as not all players will be heavily active during a given season, and some may not be active at all. To account for this, we replaced null values for seasonal statistics with "0" values to represent a player's lack of activity.

There is also missing data for certain athletic metrics captured by the NFL Combine and predraft workouts, as this data is only officially captured before the player is drafted, and even then, some metrics or measurements may not be collected. In some cases, this is caused by a player injury and their inability to complete the workout, while in others it may be do to a player's choice to not participate in a given workout. As these metrics are important to include for modeling performance, we used imputation to fill in the small sets of sporadic missing data.

To determine our imputation values, we investigated the options of using lasso regression and random forest models, using both available athletic metrics and player size metrics. After testing performance, random forest was determined to provide the best method of imputation, utilizing 3-fold cross validation. While imputation is less than ideal, given the possibility of outliers, the imputation allowed us to retain more records for modeling by increasing the overall size of our dataset.

[3] https://premium.pff.com/nfl/

4 Modeling

4.1 Scope & Why limit to WR/RB

Teams - both professional and fantasy - are made of more than just running backs and wide receivers, so it stands to reason that some may question our choice of focusing on these particular positions. Here we explain our reasoning regarding the various positions and why we chose the roles that we did.

Quarterbacks The quarterback position is generally regarded as one of the most complicated positions in sports, as it requires an extreme level of game intelligence, awareness, experience, precision, and athleticism. While many of these attributes can be loosely correlated to several factors, it is extremely hard to quantify how well a quarterback plays in the NFL, and even more challenging to model, given that there are only a few elite quarterbacks. It is unusual to see heavy competition for the starting quarterback role, with changes to the leading position typically caused by injury.

Further, many drafting strategies specifically avoid drafting a quarterback early, as there is comparatively less scarcity in the position in fantasy football, giving players many options for the position that will consistently return similar results. As such there is little need to aggressively draft for this position unless an individual player feels that they could potentially draft an elite quarterback. However, making such a determination also requires analysis of other players on a given quarterback's team, as they are highly dependent on the performance of other athletes, and therefore would be beyond the scope of this paper.

Tight Ends The tight end position was initially considered in our analysis. However, it was found that the results from the position were underwhelming, and would not be a good area to place focus. While there are clearly elite athletes in the position, there is little variance between them when examined from starting weekly performance. This is magnified by the fact that the NFL has shifted away from the position, and in its place more teams have begun using wide receivers in the modern, fast-paced passing. Because of these factors, tight ends have fallen out of our scope.

Defense Defensive positions are particularly complex for modeling performance, given that there are a 11 different starting athletes in the position. This would prove to be incredibly cumbersome to model, and so we opted against using this position.

Kicker As kickers are typically drafted at the very end of a draft, we opted against the investigation of kickers as a potential "sleeper" position. Examination of their scoring indicated little differentiation between most kickers, which indicates that the kicker position is most likely not a consistent way to gain additional points, providing only marginal results. In light of this, we opted to not investigate the position.

4.2 Defining a Sleeper From the Data

One area that is somewhat up to debate within the scope of fantasy football is how to properly define a "sleeper." The term is not well-defined in the community, and its usage varies between individuals.

Within the context of this work, in which we focus on running backs and wide receivers, a large percentage of running backs are drafted in the first three rounds of drafts. Given the importance of play opportunity, which we examine in more detail later in this paper, there is a limited volume of players likely to receive sufficient opportunity workload to place them in the top twenty for their role. Receivers who perform the strongest in their role are often drafted in the first four rounds, as there are slightly more players who perform well in their position.

Figure 1 illustrates the relationship between average draft pick position and fantasy point generation for players of the specific roles we are examining for the 2017 and 2018 seasons.

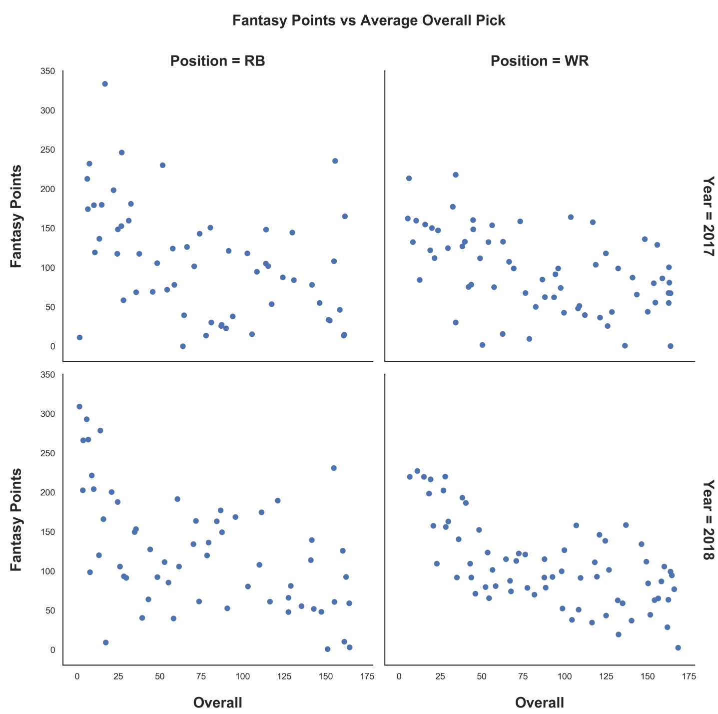

Fig. 1. Point Generation by Role, Ranking, and Year

As mentioned earlier, for purposes of our work we define sleepers as players who reach within the top twenty within their role who are drafted in the fourth round or later.

4.3 Model development

In order to build an algorithm capable of identifying sleepers, we had two goals to accomplish.

First, we must properly scope non-obvious picks for fantasy football participants - specifically, we are looking to find a breakpoint at which the draft becomes much more random in selection process, which would indicate a breakdown in the typical draft selection metrics and is the point at which that sleepers become relevant. For this analysis, each year's average draft picks (ADP) are imported into our dataset and matched with projected rankings, taken Fanasy Football Calculator[4]. This allows us to examine natural divides in projected performance that may exist and when performance of the selected players becomes increasingly volatile. By establishing a breakpoint in volatility, we can ensure our prediction metrics are not stating the obvious selection and we are creating a meaningful analysis to guide the reader in a more potentially statistically favorable outcome.

Second, we generate a forecast for fantasy points produced by the athlete to properly predict how much benefit the athlete will yield for the player. This is a rather complex task, as this is not as simple as just points, but also includes specific athlete metrics such as catches, yards, and touchdowns; this problem has further complexity by virtue of the game itself and the presence of others players, which drastically increase the complexity of the problem space. However, if we are able to generate even a rough forecast of the athlete's "raw performance," we can properly translate those results into any fantasy football scoring format, such as standard or point-per-reception scoring. Further, by developing projected scoring in this manner, we can ensure that our results correspond with logical breaks in the draft pick analysis.

To accomplish this, we will model and generate performance predictions using the general flow outlined in Figure 2. To have the potential for success we have to adequately account for the matter that a large portion of players in our dataset will not have available data for many of our attributes. This is particularly true for NFL rookies who have never taken an NFL snap, but it is also true for many NFL veterans who may spend much of their professional career on the bench, waiting for an opportunity to play. This requires us to accept that missing data must not be eliminated but worked to usability.

Practically speaking, this requires us to adapt our process and we plan to maximize performance by using a series of fitted models for cases where some or the majority of attributes are not present for the selected player. This will consist of a base model that includes attributes that are always present among our dataset and models that grow in complexity for available data.

[4] https://fantasyfootballcalculator.com/adp/standard/12-team/all/2018

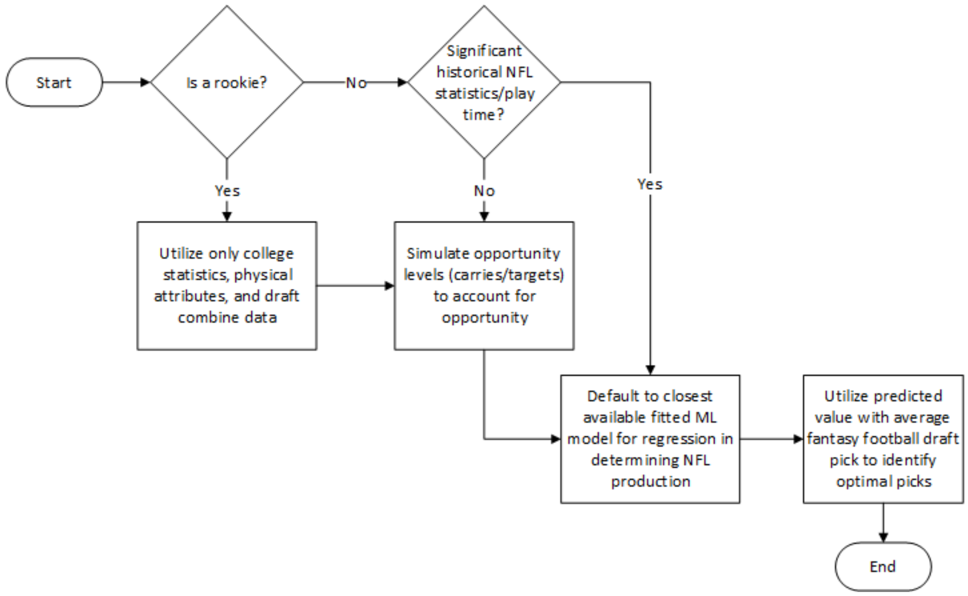

Fig. 2. Using ML for sleeper detection flowchart

Given that our dataset is large in attributes but low in row count, a level of experimentation will be required to optimize our model selection process. For the purposes of our analysis, we use lasso regression, random forest, and XGBoost for fitting our models. We then combined those into an ensemble model.

While with these models we can handle greater complexity with potential for greater predictive power, one major issue is accounting for the effects of opportunity in biasing our dataset. A reality of the NFL is that many players who have potential to perform at a high level are relegated to the bench, as they are behind an elite talent at their position on their team's depth chart. To compensate for this, we use simulated values as an input to help predict outcomes by generating different levels of opportunities (rush attempts/receptions) to balance predicted outcomes.

However, there are a significant number of players who are used purely for high-risk big plays and would not likely ever be utilized at large rates in an NFL game. To counteract this, we build different levels of simulated values and attempt to identify these players to help prevent extreme values.

4.4 Imputation

Examining our player data revealed a significant challenge in that several athletic metrics were not available from our data. Many of these athletic metrics (e.g. 40 yard dash, bench press, 3 cone drill) are captured at the NFL combine where aspiring NFL athletes perform drills, so prospective NFL teams can evaluate the players potential and determine if the player is someone they wish to pursue in the NFL draft.

Predictably for the sport of football, this doesn't always work smoothly. Players entering the draft may be injured from their final season in college, which will result in a lack of available data on the player's performance. This may take the form of missing a few workout metrics, or possibly all. In addition, some athletes may choose not to participate in certain workouts. To counteract this, imputation was performed for missing player performance

Imputation was performed by using multilinear regression and random forest modeling with a 3-fold cross validation method based on available values for the player with missing data and trained on available metrics. This was performed by developing multiple models that imputed records from most complete to least complete to ensure the maximum amount of player records could be salvaged for usage. Less than 10 running back and wide receivers who played during this period were dropped entirely due to lack of available metrics; however, these data points consisted of players who had extremely limited playing time and would not have been of interest for our purposes.

4.5 Forecasting Performance

To forecast performance, we forked our model development for running backs and wide receivers as the dataset for each is significantly different, and both need to measure significantly different types of athleticism and skill sets. Both datasets were combined and transformed following an identical process for a consistent data structure. A fantasy football scoring metric was created, measured on 1 point per 10 rushing or receiving yards and 6 points for a touchdown. Bonuses for long runs were not applied, and no point value was assigned to completed receptions.

An attempt was made to bring in player name as an encoded variable as well as the players corresponding team. It was hypothesized that this might allow for a better historical component for distinguishing star players, but this proved to have no clear positive impact on the model and in some cases created impossible negative values due to a lack of carries. It was also hypothesized that including the corresponding NFL team for each player would allow for other factors like coaching influence to perhaps be accounted for, but this proved to have no clear positive effect in our research. This is likely an areas with to much variance to effectively model over six years of data, as team schemes can change dramatically, and further investigation to account for coaching scheme changes may be merited in future efforts.

4.6 Importance of Opportunity

Targets and carries were included as variables for modeling athletic performance for each year. While passing targets and rushing carries will not be known before the season, including these variables can produce a dynamic ability to model athletic performance at a high level, and allow for the identification of players who may be more capable of performing well in a fantasy setting.

A reasonable point can be made against including these variables because of the unknowns with forecasting play opportunity before the season, and initially, models without carries and targets were developed with the hypothesis that top talent would naturally rise to the top. Interestingly, this proved to not be the case with our selected modeling methods. Nearly all predictions were extremely far from actual performance to the point of no reasonable use case (with R^2 values of 0.4), though this may be a point of interest, based on the variable importance listed from our models. One might hypothesize that athletic performance is not enough to push someone to a starting role with a team, but there are less easily quantifiable factors that may separate some athletes from others.

4.7 Forecast Performance and Feature Importance

We used lasso regression, random forest regression, and XGBoost as regressors to forecast fantasy performance based on historical data. A 3-fold, 5-fold, and traditional train/test split were used for developing both models, using the 2013 to 2016 data as training data, and 2017 and 2018 as test data.

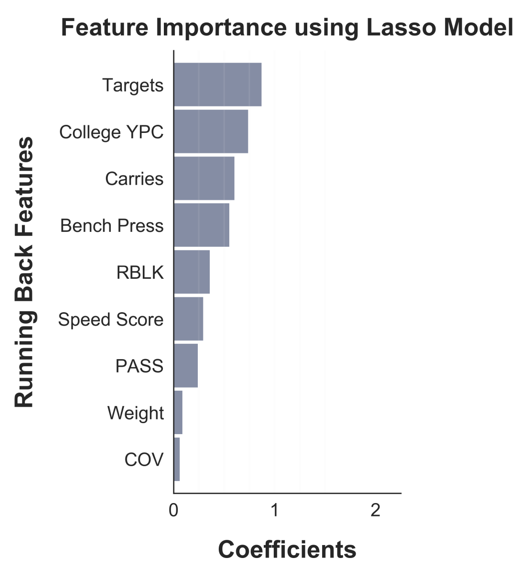

Fig. 3. Running Back Feature Importance using Lasso

Lasso regression was performed using scikit learn using an alpha of 1.0 with a "fit intercept" enabled. Our model returned a R^2 value of 0.927 for forecasting fantasy points, and an MSE of 352.03. An examination of feature importance does help explain the high score.

Figure 3 on page 14 illustrates the feature importance for this model: targets, which are the specific "opportunity" for running backs, have the highest importance in terms of predicting fantasy point generation.

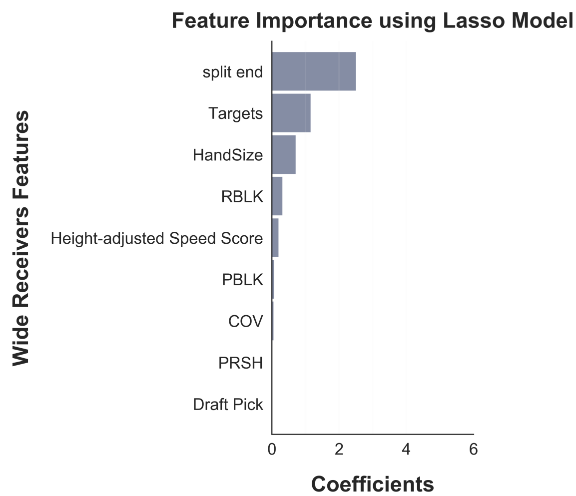

Fig. 4. Wide Receiver Feature Importance using Lasso

Figure 4 on page 15 illustrates the feature importance for this model: "split end" is the most important feature, which indicates that a player is typically the primary wide receiver. It is likely that this plays into opportunity: if a given player is given primacy over other players of a similar position, they will likely have more opportunities to score points, so it is sensible that the model is taking this into account. Also high in importance is targets, which – as with running backs – is a factor of opportunities.

Wide receivers demonstrated similar performance with a R^2 value of .89 utilizing lasso regression, with an MSE of 288.78. Here we see again opportunity being key, but factors such as height-adjusted speed score did slightly color our model, and age and weight appeared as negative coefficients for the model.

Random forest was attempted, but our results were somewhat less compelling with an R^2 of .89, with MSE of 486.96. Even after performing an extensive random grid search to tune hyper parameters the model relied heavily on carries and targets with little else appearing for importance. This was also true for wide receivers reaching a R^2 of .90, and an MSE of 254.51, which appeared stronger than the lasso regression, but on deeper examination the prediction values showed less consistency with actual performance for elite players and was overfit to lower scoring players.

XGBoost provided our most compelling results with a R^2 value of 0.96, and an MSE of 397.85. Examining XGBoosts feature importance we see while carries and targets are again the most important metrics, other variables do seem to play a larger role in improving model prediction performance. XGBoost appears to be allowing for a stronger fit in accounting for the influence of additional factors which is allowing the model to reach slightly higher levels of performance.

Figure 5 on page 16 outlines the feature importance for this model. As with the lasso models, carries and targets are of high importance in predicting point generation by running backs using XGBoost.

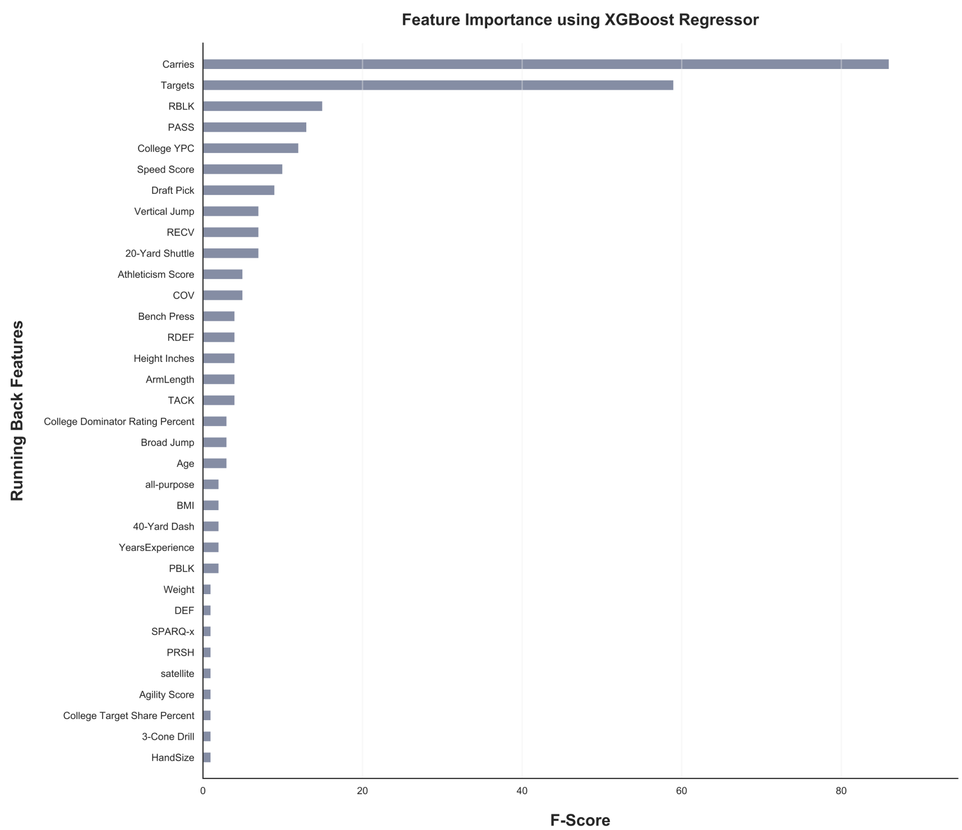

Fig. 5. Running Back Feature Importance using XGBoost

This also appeared true for wide receivers as XGBoost allowed us to reach an R^2 value of .91, and an MSE of 255.58. We see a similar chart for feature importance compared to running backs again reflecting the importance of opportunity, however, we do see a larger emphasis on blocking in the rblk and pblk variables.

Figure 6 on page 17 outlines the feature importance for this model. Opportunity again features prominently as a primary indicator of fantasy point generation for wide receivers.

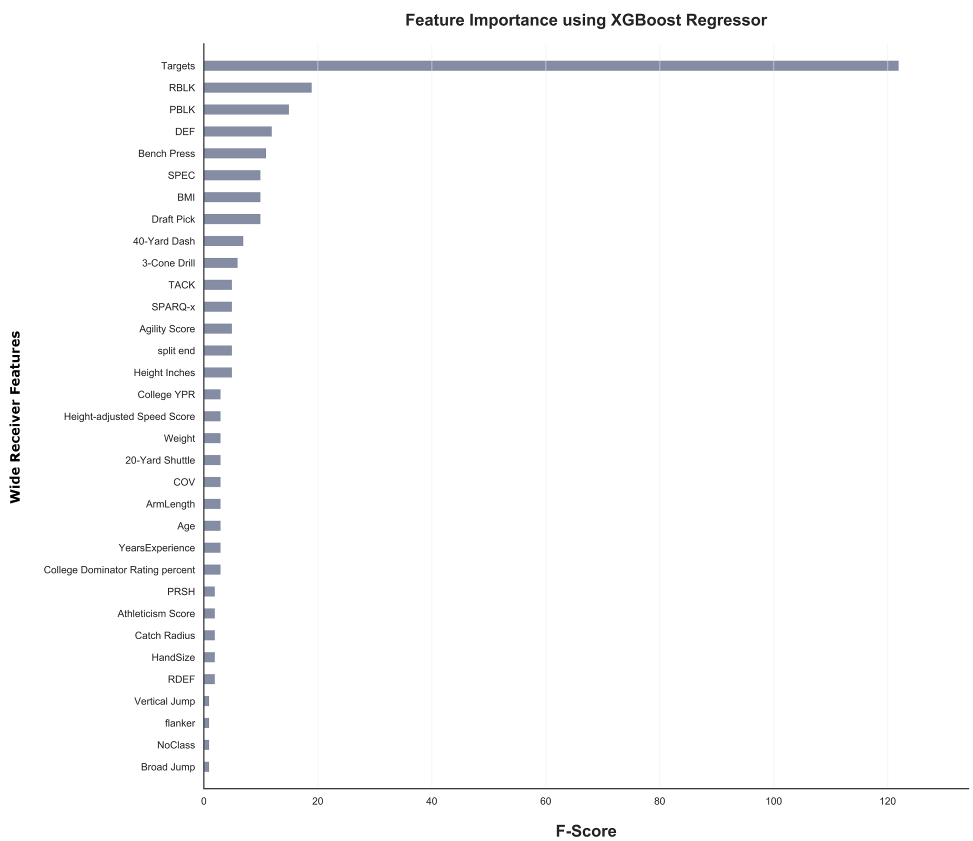

Fig. 6. Wide Receiver Feature Importance using XGBoost

Figure 7 and 8 on page 18 demonstrates the behavior for lasso regression, random forest, and XGBoost for predicted versus actual performance for the 2017 and 2018 season. Both charts appear to show while variation is present, particularly for higher scoring players, our models do produce similar values relative to actual performance with accurate forecasts for workload.

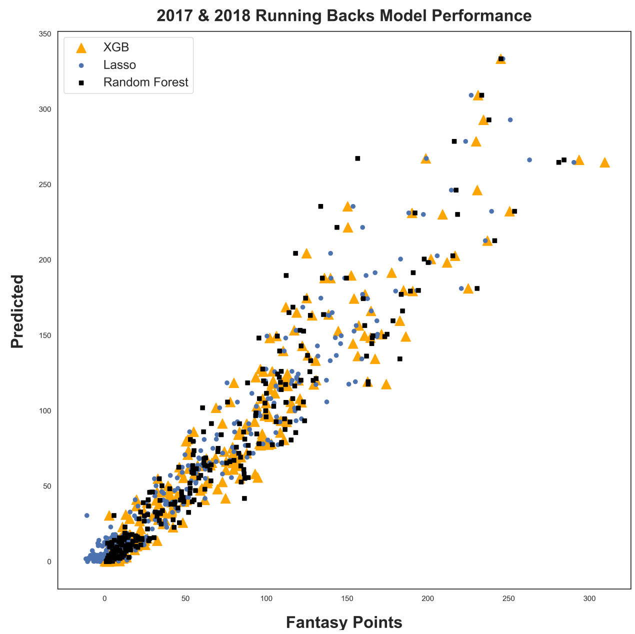

Fig. 7. Running back predicted versus actual for selected models

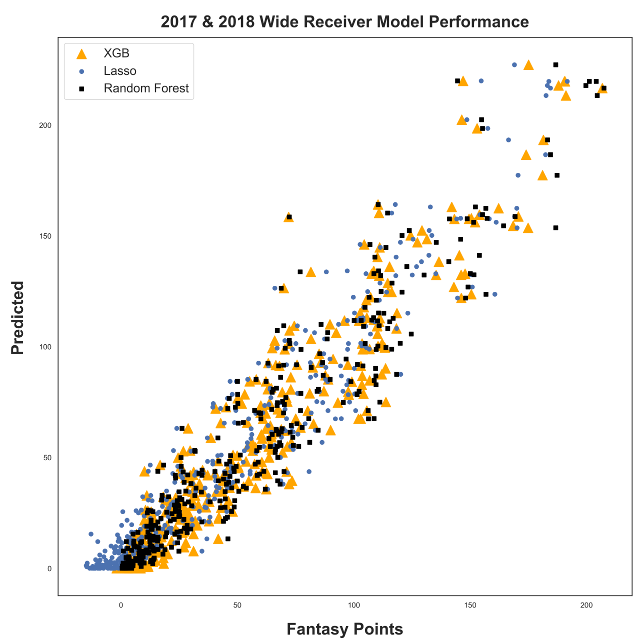

Fig. 8. Wide receiver predicted versus actual for selected models

Once we developed each of these models, we then developed an ensemble model to combine them into a more powerful model. For running backs, the ensemble model's weights were 82% lasso, and 18% XGBoost; while for the wide receivers, its weights were 19% lasso, 41% random forest, 40% XGBoost. The overall construction of this model can be seen in Figure 9 on page 19.

The ensemble models had marginally improved MSEs: the model for running backs had an MSE of 347.17, and wide receivers had an MSE of 246.73.

5 Simulation for Sleeper Identification

While these models do predict player performance effectively with accurate carry and pass targets estimates, predicting actual play opportunity remains a clear gap to reaching an entirely predicted process that is repeatable without human intervention requiring significant knowledge of current NFL rosters.

Forecasting playing opportunity was attempted, but this proved to be an area of required additional research and far beyond determining which players had the potential to play at high levels and were being undervalued.

It appears that a focus on determining team play pace and playing style combined with defensive effectiveness may allow for estimation offensive time on field, with which one may be able to model player opportunity relative to passing target/carry shares, but this still requires a level of tuning by the modeler who studies current NFL player trends. This is a desirable component to enhancing the ability to predict sleeper level talent as are finding clearly demonstrate play opportunity is key, but such an effort would reach far beyond our initial scope and remains an opportunity for future work.

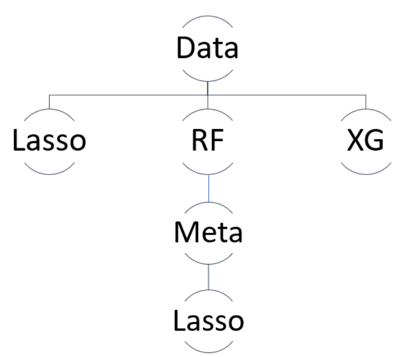

Fig. 9. Ensemble Model Construction and Data Flow

Less sophisticated but viable methods exist that will allow us to utilize the potential of our model and identify potential sleepers. One simple method would be to seek out external target and carry projections and feed them into the correct player model for the player one wishes to examine. Testing different targets and carries for each player presents a means to simulate how a player might respond to an increased workload and their potential to elevate to a strong fantasy status.

Simulation of playing opportunity levels was performed at varying levels where each running back received the same amount of targets and carries. Both models and results provided some interesting findings listing known current superstar players near the top of the model, with some notable exceptions. These exceptions were mostly players who are now reaching the sunset of their career and no longer perform at the same athletic level. This is likely indicative that the models do not fully reduce performance for age correctly and more research may be needed. A more practical remedy to this issue would be for more current sampling of athletic metrics (40 yard dash, bench press, etc.) to be taken as there is likely a significant athletic drop-off for players who have aged and are reaching the end of their career, but this is admittedly unlikely occur and be made available to the general public due to the practical constraints of measuring current NFL athletes.

Our models do illustrate some lesser-known young players with little NFL play time near the top of the projected list when carries/targets are constrained, but we can not yet say conclusively the model has identified these players for the 2019 years as true sleepers who are guaranteed to produce at high levels. While it is encouraging to see some players who were not initially highly sought after in fantasy drafts like Tyreek Hill and Michael Thomas near the top of the simulated list, other high-performance players like Davante Adams and DeAndre Hopkins were noticeably near the bottom for our wide receiver list. Both Adams and Hopkins scored lower in some workout metrics (despite Hopkins being highly-sought talent, due to his record-setting collegiate performance), which does illustrate a lack of ability to predict all individuals who will emerge as leaders at their position.

However, upon examination of the data and our models' performance, as well as discovery of the relationship between opportunity and player performance, we are not certain that these models can discover undervalued players with much statistical accuracy. The reliance of our models upon opportunity suggests that a more thorough model for opportunities that could be used in conjunction with the models developed here would be required before we could conduct a more thorough investigation of the "sleeper" phenomenon.

5.1 Benchmarking Other Sources

To test our model's performance, we compared its performance against predictions made by industry sources. We used data from the 2018 NFL season, and for comparison we used predictions generated by SportsDataio[5], a company that provides fantasy data and predictions. Using their API, we pulled JSON data

[5] https://sportsdata.io/

that showed forecasted seasonal results for the 2018 season. To normalize differences, we used the same targets and carries for running backs and wide receivers.

Both our model and SportDataio's predictions showed more sporadic predictive performance when using preseason projections for 2018, which illustrates the volatility of predictions in this space. Running back predictions showed SportsDataio reaching an R^2 of .48, and an MSE of 2572; our model produeced an R^2 of .46, with an MSE of 2692. Wide receiver predictions showed SportsDataio's predictions with an R^2 of .38 and an MSE of 1886; our model had an R^2 of .28 and an MSE of 2186.

These results illustrate the importance of forecasting opportunities: as we have indicated, accuracy in projecting targets and carries has significant impact on model performance. In these projections, there were instances of differences in projections of over 200 for carries for running backs, and 80 targets for receivers. Combined with our earlier results in training, this indicates that while our model is capable of accurate forecasting, it must be combined with opportunity forecasting, or active management by the user throughout the NFL season to update predicted opportunities for relevant players.

6 Ethics

One of the primary concerns when performing analysis of a given athlete's qualities, in an effort to predict their quality as a player, is that of discrimination. It has been demonstrated in a study that statistical discrimination, along with models of option value and productivity evaluation periods, influence player success in the NFL [12]. This would be in direct contravention of section 1.4 of the ACM Code of Ethics, "be fair and take action not to discriminate" [11].

Analysis of a given athlete's qualities in an effort to predict their quality as a "sleeper" has some potential ethical pitfalls.

In the context of fantasy football, the mathematical analysis of an athlete's attributes to determine their "sleeper-hood" is not necessarily problematic. Fantasy football has minimal impact on said athlete, and players of fantasy football may already make use of athletes' statistical information in making their decisions. Utilizing a more rigorous mathematical model fueled by machine learning or other, more intricate algorithms is not morally ambiguous.

However, what is potentially ethically problematic is the overall shape of the approach. Generalizing, what we are doing is taking an individual's properties and judging their fitness for a given goal or task based upon mathematical analysis of those properties. Due to the mathematical nature of the algorithms in use, this largely necessitates that the characteristics being used for the analysis are quantifiable in some fashion.

For selection of athletes for a fantasy football league team, this would seem to be fine. However, this same sort of approach could be used for something akin to China's "social credit score" system, in which individuals are disenfranchised from society by virtue of possibly-arbitrary characteristics. Those who implement such systems may claim that they help serve the public good by dint of using sensible metrics, but it is also ripe for abuse: using locational data, for instance, and declaring that individuals from a certain region have a lower social score simply for being from that part of the world. Ill-meaning actors could include other factors, like race or religion, and use these to manipulate the results of their algorithms, ostracizing elements of society that they see as undesirable.

While we have not implemented such variables explicitly within our algorithm, it is also entirely possible that our work could give the perception that we are using them even if we are not. For instance, it could be entirely possible for a given season to have athletes of a particular ethnicity who, for whatever reason, all show particular aptitude in a given quality we measure that we find is important to determining "sleeper-hood." While we have not explicitly coded a bias for that ethnicity, the end result of our algorithm will potentially result in having such a bias, thereby tainting future uses of the algorithm.

While the example used earlier may be extreme, this sort of examination of the potential abuses of a particular approach is necessary. Taking things to their logical conclusion before moving forward can help us in the development of our algorithms and how we sample our data, helping us to be mindful of potential biases and take the appropriate steps to avoid them. If we can see how an approach could be misused in other applications, we gain better insight into what we are doing and can use that to ensure that our conclusions are fair and balanced, avoiding potential inequities.

The emergence of daily fantasy sports contest (DFS) has increased the engagement of fans and benefiting the professional sports leagues, but this does come with points of further ethical issues relative to gambling concerns. The increase in popularity might be due to a law passed by Congress which legalized traditional fantasy sports leagues but left DFS in a vague area of interpretation. The legality issue of DFS is being proposed for amending the UIGEA law[7]. Many have argued that daily fantasy sports border much more on a game of chance and have been cited "as a threat to public welfare" because individuals are required to engage weekly [14].

While the participation in fantasy sports for cash prizes seems benign to most viewers, reasonable ethical concerns can be brought forward relative to gambling addiction. One must ask if an individual could be brought to the point of harm through addiction to fantasy football and potentially incur significant financial damages through there addiction. As such, we must recommend that all players conduct play responsibly to not put there personal well being in jeopardy through high stakes gaming.

7 Conclusions

We present a model that predicts future fantasy football performance with an MSE of 347.18 for running backs, and 246.74 for wide receivers. Opportunity is by far the most important metric for determining an athlete's fantasy football point production, and our model's utility in prediction is largely a function of the accuracy of the prediction of opportunity.

References

1. Fantasy sports trade association (fsta) (1998), https://fsta.org/about/history-of-fsta/
2. Aguirre, A.A.: Monstrosities, money, machines: A metaphoric analysis of fantasy football as a social world (2015)
3. Becker, A., Sun, X.A.: An analytical approach for fantasy football draft and lineup management. Journal of Quantitative Analysis in Sports **12**(1), 17–30 (2016)
4. Carter, C.: Going deep: how wide receivers became the most compelling figures in pro sports. Hyperion (2014)
5. Dick, M., Wellnitz, O., Wolf, L.: Analysis of factors affecting players' performance and perception in multiplayer games. In: Proceedings of 4th ACM SIGCOMM workshop on Network and system support for games. pp. 1–7. ACM (2005)
6. Dolan, P., Karaouni, H., Powell, A.: Machine learning for daily fantasy football quarterback selection. CS229 final report (2015)
7. Dourmashkin, M.: Examining the legalization of daily fantasy sports, 25 cath. u. j. l. and tech (2017)
8. Dwyer, B.: Fantasy sport consumer behavior: An analysis of participant attitudes and behavioral intentions (2009)
9. Fabiano, M.: Fantasy football 101 national football league. retrieved 2011-10-03
10. Fokoue, E., Foehrenbach, D.: A statistical data mining approach to determining the factors that distinguish championship caliber teams in the national football league (2013)
11. Gotterbarn, D., Miller, K., Rogerson, S.: Software engineering code of ethics. Commun. ACM **40**, 110–118 (1997). https://doi.org/10.1145/265684.265699, https://www.acm.org/code-of-ethics
12. Hendricks, W., DeBrock, L., Koenker, R.: Uncertainty, hiring, and subsequent performance: The nfl draft. Journal of Labor Economics **21**(4), 857–886 (2003). https://doi.org/10.1086/377025, https://doi.org/10.1086/377025
13. Horton, T.: Complete running back. Human Kinetics (2016)
14. Leishman, N.C.: Daily fantasy sports as game of chance: Distinction without a meaningful difference?, 2016 byu l. rev. 1043 (2016), https://digitalcommons.law.byu.edu/lawreview/vol2016/iss3/9
15. Steenkiste, P.: Finding the optimal fantasy football team (2015)
16. Teramoto, M., Cross, C.L., Willick, S.E.: Predictive value of national football league scouting combine on future performance of running backs and wide receivers. Journal of Strength and Conditioning Research **30**(5), 13791390 (2016). https://doi.org/10.1519/jsc.0000000000001202