# Optimising Fantasy NFL DraftKings Points

John Lazenby, Xiaming Jin,
Aarushi Bagga, Shaun Fendi Gan
December 30, 2020

***Abstract*—Prediction and prescription models were tested to find the most effective method for winning at daily DraftKings double up competitions. Utilising free online expert projections incorporated with a variance threshold on previous point performances was the best approach, yielding an expected return of $160 for every $100 wagered. Using the combination of a player's 10$^{th}$ percentile performance and 60$^{th}$ performance as a scoring method with a variance threshold was also an effective and interpretable method, returning $140 for every $100 wagered.**

## I. Introduction

IN 2019, almost 46 million Americans participated in daily fantasy sports (DFS). Traditional fantasy sports have been around since the 1960's but DFS has only recently exploded in popularity [1]. In both traditional and DFS, participants select athletes from various sports and accumulate points according to those athletes on-field performance. The popularity of DFS has been attributed to a participant's ability to bet on contests without needing to wait the duration of an entire sports season to see a payoff [2]. Our goal for the project was to create a data driven strategy with a positive expected return from daily fantasy football.

There are a variety of NFL DFS contests in terms of both the contest's payout structure and rules for accumulating points. Most contests use the "classic" scoring rules (appendix Fig. 2-3). There are many ways to accumulate points, but the general principle is simple; an athlete accumulates fantasy points when they do something positive on the field. The second important aspect of the "classic" rules is participants are constrained in which players they are allowed to pick. Participants must satisfy the position requirements (appendix Fig. 4), additionally each player has a salary and the participant can only select a roster of players with a total salary under $50,000 (appendix Fig. 5). There are a wide variety of payout structures that we generalize into two types. In one type (we call top heavy) only a small percentage of the top players receive any payout, but their payout is large relative to the amount they wager. In the other type (we call double up) roughly the top half of players double the amount of money they wager. We decided we were more likely to be successful in the short term in the double up contests because we need perform only in the top 50% of contestants to be profitable instead of the top 95%.

We investigated four strategies to compete in the double up style contests described above. The first strategy relies only on point prediction and optimization. The other three strategies attempt to reduce the innate variability of this problem.

Advisor: Hari Bandi

## II. Method

This section describes our four strategies: DraftKings Point Predictions, Yards per Game Predictions, Variance Tolerance, and Percentile combinations. Data can be found in [3].

### A. DraftKings Point predictions

An initial point prediction model was developed to predict DraftKings Points (DKP) for each player. This was the obvious first choice as the goal was to score a high enough DKP to clear a required threshold for being in the top 50%. When selecting a lineup, participants are required to select a combination of different positions. The possible positions included:

1) Quarterbacks (QB),
2) Wide receivers (WR),
3) Running backs (RB),
4) Tight ends (TE) and
5) A defensive line up (DST)

where at least one defensive lineup, quarterback, and tight end must be selected, as well as two running backs and three wide receivers.

Since each position has different factors that contribute to their performance and fantasy point score, this problem can be divided into four key components. Wide receivers and tight ends have the same goal in an offensive lineup so we modeled them using the same covariates. The important covariates for each position can be seen in Table I, where the team, opponent and previous match scores were also used for all positions.

This is a time series problem and a player's performance throughout the season was used in our modeling approach. We created lag variables on the key features that described a player's form seen in Table I. A four game lag was used. Lag variables for a team's previous scores to track the form of the team were also used. Training sets were formed by taking all data points in 2019 before April 8, 2019, and the testing sets were all points between April 8, 2019 and April 28, 2019.

TABLE I
Unique Features for Each Positions

| | |
|---|---|
| Quarterback | Pass.Att, Pass.TD,Pass.Yds, Fumbles, Rush.Att, Rush.TD Rush.Lng, Rush.Yds |
| Runningback | Rush.Att, Rush.Yds, Rush.TD, Rush.Lng, Rec.Tgt, Rec.Rec, Rec.Yds, Rec.TD, Fumbles |
| Defense | Pass.Sk, Pass.Int, DST.points.allowed, Fumbles lost |
| Wide Receiver + Tigthend | Rec.Tgt Rec.Rec, Rec.Yds, Rec.TD, Fumbles |

Additional lag variables such as the season average on points for all games before the upcoming date and the form across the previous eight games instead of four were also explored. Both of these variables did not contribute to a significant impact on improving out of sample $R^2$ and were hence excluded.

Several models were tested to produce the point predictions, these included Lasso, Random Forests and XGBoost. The goal of this problem is not to obtain powerful interpretability but rather predictability, so we used methods that were empirically stronger at prediction.

Once the point prediction was completed, the ideal selection of players was found subject to constraints using a prescription model. This model can be seen in (1)

$$\begin{aligned} \max & \sum_{i=1}^{n} z_i y_i \\ \text{s.t. } & \textstyle\sum_i z_i \cdot \text{DKSalary} \leq 50000 \\ & \textstyle\sum_i z_i \cdot \text{Def}_i = 1 \\ & \textstyle\sum_i z_i \cdot \text{QB}_i = 1 \\ & \textstyle\sum_i z_i \cdot \text{TE}_i \geq 1 \\ & \textstyle\sum_i z_i \cdot \text{RB}_i \geq 2 \\ & \textstyle\sum_i z_i \cdot \text{WR}_i \geq 3 \\ & \textstyle\sum_i z_i (\text{TE}_i + \text{RB}_i + \text{WR}_i) = 7 \\ & z_i \in \{0, 1\} \end{aligned} \tag{1}$$

where $y_i$ is the predicted points for player $i$ and $z_i$ is a decision variable that returns a binary vector that indicates the players that form our roster for the upcoming week. An additional constraint (2) was consecutively added to obtain lineups that were close to optimal

$$\sum_{i=1}^{n} z_i y_i \leq \text{obj} - \epsilon \tag{2}$$

where obj was the objective value of the previous optimisation, and $\epsilon$ is a small tolerance to limit the objective value. For this prescription, $\epsilon = 1$ was used and 10 rosters were produced. These rosters could be used to better analyze the prediction power of different models. Rosters that were not optimal in point predictions according to the model often outperformed on fantasy points.

Combining these two components of prediction and prescription displays a foundation for an analytically informed solution at tackling fantasy competitions.

### B. Accounting for Variability 1 - Yards per Game Predictions

The outcome we attempted to predict in the previous section (DKP) is extremely variable. Outcomes are not entirely determined by the skill of the players. For example a skilled player may have a low DKP score in a given week because they were not given any opportunities to play when the ball was close to the end-zone. We hypothesize that by attempting to model a different outcome that is less dependent on luck and is a better measure of a player's actual skill we may have better results than if we model the outcome of interest (DKP) directly. We determined yards gained per game would be a more skill related metric to use as the dependent variable in our models. This metric varied slightly by position and is seen in Table II. The predicted yards was also normalised by position by dividing the prediction by the mean value of all players in that position.

TABLE II
NEW DEPENDENT VARIABLE FOR EACH POSITIONS

| | |
|---|---|
| Quarterback | Pass.Yds + $\max\{0, \text{Rush.Yds}\}$ |
| Runningback | Rush.Yds + Rec.Yds |
| Defense | Yards.Allowed |
| Wide Receiver + Tightend | Rec.Yds |

Underlying both of these strategies is a point prediction that depends on our available data. Because NFL DFS is popular there are a variety of free expert predictions available online. These expert predictions take into account additional features such as a player's injury status and the weather which were not available in our data. Therefore the outcome of using expert projections were compared to the original model, in which both utilise the same prescriptive model (1).

### C. Accounting for Variability II - Variance Tolerance

In our third strategy, a variance threshold was imposed as an additional constraint in our roster choice optimization model. Players that perform at a consistently high level are prioritised, leading to the rosters having higher and more consistent scores. Two approaches were tested: add a threshold for the total variances of our picked roster;

$$\textstyle\sum_i z_i \cdot \text{VAR}_i \leq \text{Variance Sum Threshold} \tag{3}$$

add a threshold for each individual player's variance.

$$z_i \cdot \text{VAR}_i \leq \text{Variance Threshold for i} \in 1, ..., n \tag{4}$$

where VAR is the variance normalized by position of a player's last 8 fantasy performances. The choice of threshold was tuned on four weeks of data and the threshold that gave the most wins was chosen.

### D. Accounting for Variability III: - Percentiles

By participating in double up style we need only beat 50% of the field to be profitable and any points we accumulate beyond that threshold no longer contribute to the profitability. In this percentile strategy, the aim wasn't to maximise the expected points for each football player, but instead try to identify a set of players that are able to consistently score enough points to exceed the profitability threshold. To do this, players that score many points even on their worst games were opted for. This can be accomplished by assuming a player's recent performances represent the distribution of what to expect from their future performances. Players that score many points on their worst games were identified by considering their DKP at some percentile less than 50 amongst their previous games.

Taking this approach runs the risk of completely ignoring a player's potential up side. For example if two players have the same $10^{\text{th}}$ percentile but one player has a much higher upside ($90^{\text{th}}$ percentile for example) than we would prefer the player with the higher upside. To do this we characterize players as some combination of their floor (lower percentile performance) and their up side (higher percentile performance) and then select the combination of floor and upside that leads to a roster that consistently exceeds the profitable threshold. According to our theory above, the optimal combination of floor and upside likely will consider floor much more than upside. We do not rely on this theory but instead use back-testing to learn the optimal combination.

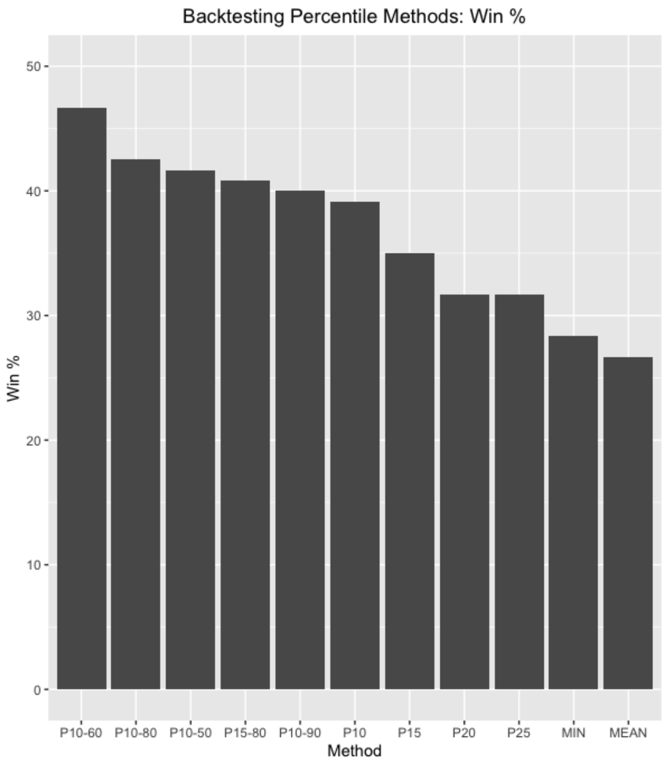



Fig. 1. Win % of 5 optimal/close to optimal rosters for various percentile methods. P10-60 indicates that the $10^{\text{th}}$ percentile had a 95% weight where as the 60th percentile has a 5% weight. P10 indicated that the $10^{\text{th}}$ percentile was weighted at 100%. Minimum and mean were also included as a baseline.

Profitability thresholds for DFS double up games from 2017-2019 were collected and expected scores based on different linear combinations of percentiles of a player's previous games were created. Five optimal and close to optimal rosters were then selected based for each percentile combination to observe the win percentage. Combinations of low percentiles ($10\text{-}50^{\text{th}}$) and high percentiles ($50\text{-}90^{\text{th}}$) were tested, where low percentiles were weighted as 100%, 95% or 90%. The results of testing using a 95% weight on low percentiles can be seen in Fig. 1. Predictions using the minimum and mean were also included as a baseline. The combination of the $10^{\text{th}}$ and $60^{\text{th}}$ percentile approach approaches a win percentage 50% which is what is required for profitability.

## III. Results and Discussion

In this section the results of all the predictive and different prescriptions are evaluated.

### A. Prediction Models

Four metrics were used to evaluate DKP and yards prediction models. $R^2$ , Mean Absolute Error, Median Absolute Error and percentage of observations lying in an acceptable error range. $R^2$ evaluates the model developed for each position against a baseline model that is based on the average points scored at each position. Factors unaccounted for in the model like injuries gave rise to very high discrepancies between predicted points and points scored in some cases, hence, $R^2$ alone does not summarize model performance. For the problem at hand, Mean Absolute Error (MAE) and Median Absolute Error, and percentage of observations lying in an acceptable error range better estimate prediction error distribution. The error range considered is different for all positions. These metrics are summarised in Table III and IV.

TABLE III
Performances of DKP prediction models

| Position | $R^2$ | MAE | Median Abs. Error | % Error Threshold |
|---|---|---|---|---|
| QB | 0.15 | 6.69 | 5.95 | 43.00% |
| RB | 0.32 | 6.17 | 4.48 | 49.67% |
| WR+TE | 0.35 | 4.84 | 3.40 | 57.77% |
| DST | −0.05 | 7.35 | 7.31 | 0% |

For DKP, Random Forests gave the highest out of sample $R^2$ for offensive positions. $R^2$ was $0.32$ for Runningbacks and Wide receivers + Tightends. Other models had scores below $0.3$. For Quarterbacks, out of sample $R^2$ value was $0.15$. For defensive lineups none of the models were predictive, and the best out of sample $R^2$ was $-0.05$. This out of sample performance was not sufficient, and it is believed that this was due to the large variability in DKP outcomes.

TABLE IV
Performances of Yards prediction models

| Position | $R^2$ | MAE | Median Abs. Error | % Error Threshold |
|---|---|---|---|---|
| QB | 0.39 | 64.31 | 59.05 | 49.23% |
| RB | 0.32 | 23.59 | 18.11 | 54.19% |
| WR+TE | 0.38 | 20.85 | 15.42 | 54.20% |
| DST | −0.06 | 6.89 | 6.74 | 33.33% |

For yards, a substantially higher predictive performance for Quarterbacks was obtained, increasing $R^2$ to $0.39$. Predictive performance for Runningbacks, Defence, and Wide Receivers + Tightends however, did not improve significantly.

### B. Prescription Rosters

Three metrics were used to evaluate the prescriptions of rosters. Firstly, '% Win' measures the percentage of times the optimal roster clears the money line for a selection of test weeks. This indicates the reliability of the model. 'Avg. clearance' obtains the average points the optimal roster was away from the money line. This indicates the selection power of the model. Finally, 'Std' describes the average standard deviation of the points achieved from the 10 rosters produced. This indicates the stability of the model

The performances of each model are summarised in Table V. Both point prediction models based on predicting DKP

TABLE V
PERFORMANCES OF EACH MODEL

| Model | % Win | Avg. Clearance | Std |
|---|---|---|---|
| Predicted DKP | 12.5% | $-24$ | 26 |
| Yards norm | 37.5% | $-12$ | 17 |
| Expert Projection | 70% | 15 | 16 |
| $0.95 \cdot 10\%^{\text{ile}} + 0.05 \cdot 60\%^{\text{ile}}$ | 71.4% | 2 | 18 |
| Yards norm + var | 43% | $-9$ | – |
| Expert Proj. + var | 80% | 5.4 | – |

and yards performed the worst. Predicted DKP had the lowest percentage of wins and the lowest scores, averaging a clearing of negative 24 points. Predicting DKP also had the largest variability.

Predicting on yards improved the win percentage to 37.5% and reduced the variability of the roster performances. However, this results was still far from profitable. Incorporating a variance threshold onto the yardage prediction improved win percentage by 6%, but again it is still not profitable.

The percentile predictions had a variability similar to the yardage prediction but outperformed it in terms of reliability winning 71.4% of the time. The average score was 2 points above the money line. This is especially compelling because it achieves this performance despite not leveraging any of the data about a player's opponent, or non-fantasy related statistics.

Expert projections used in combination with the prescription model (1) had a 70% win rate with a high average clearance of 15 points. This result approached the performance of the percentile approach. Using the same expert projections in combination with a variance threshold led to an average win rate of 80%. This was our best performing model.

## IV. CONCLUSION

We were initially surprised that our point predictions were so poor. We assumed that these analytical models of Random Forests and Lasso would perform well and help us select winning rosters. In hindsight its not surprising that our predictions under performed expert predictions considering the quality of our data and short period of time we had to create the models. Experts often have access to more features and complete data, and have spent likely improved the performance of their models over several years. As disappointing as our point prediction results were our optimization and variance reduction results were extremely compelling.

Combining freely available expert predictions with an optimization prescription is already a profitable strategy. This suggests what we initially expected is true; a majority of fantasy players make only a casual effort in their roster selection or let personal biases override the common sense advice to trust the experts. If participants are aware of expert predictions but still perform poorly it could suggest that many humans are bad intuitive optimizers.

Over time it is possible that the market will become more competitive and the profitability threshold will increase. In this case our final two strategies (variance tolerance and percentiles) are promising. The expert prediction combined with variance tolerance is promising because it outperforms the expert prediction only model in terms of win percentage and the tolerance can be adjusted as the profitability threshold changes. The percentile strategy is compelling because it matches the performance of the expert predictions by using only past DKP. If the profitability margin increases it is likely the percentile strategy can be improved by leveraging additional data.

In conclusion, using an analytic approach for roster selection can lead to profitable results on NFL DFS. Furthermore, even if the NFL DFS market becomes more competitive, an analytic approach will likely still be able to maintain an edge.

## REFERENCES

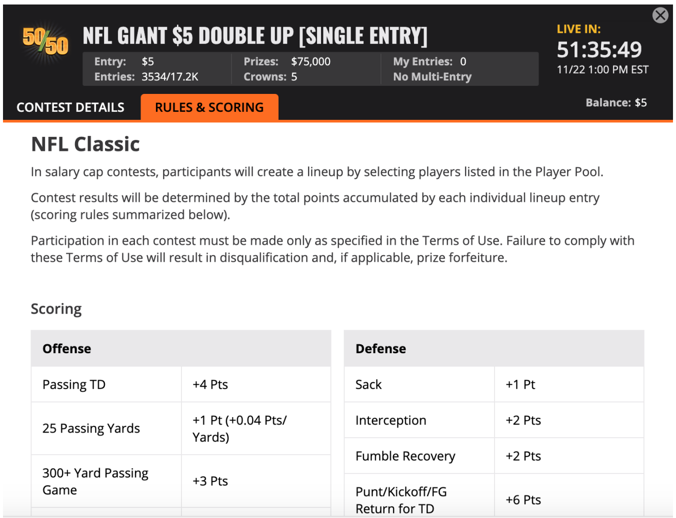
[1] Fantasy Sports and Gaming Association. Industry Demographics. https://thefsga.org/industry-demographics/, accessed on 25 Nov 2020.
[2] Fantasy Sports Net. History of Fantasy Sports https://fantasy-sport.net/history-of-fantasy-sports/, accessed on 25 Nov 2020.
[3] Advanced Sports Analytics. NFL Raw Data 2016 - 2020. https://www.advancedsportsanalytics.com/nfl-raw-data, accessed on 15 Oct 2020.

## APPENDIX

**Note:** It should also be noted that the team is currently at a \$5 net profit, losing \$10 but winning \$15 throughout the investigative process.

Data keys are found as

- .Att - Attempts
- .Cmp - Completions
- .Yds - yards
- .TD - related to touchdowns
- .Int - Interceptions
- .Lng - Longest pass
- .Sk - Sacks
- .Sk.Yds - Sack yards
- .Ret - Returns
- .YPR - Yards per Return
- .Ctch - Catch
- .Tgt - Target (how many times a player is thrown to)
- Rec.Rec - Times a player catches a pass
- Rec - Receives
- DST.points.allowed - in game points allowed by a defense


NFL GIANT \$5 DOUBLE UP [SINGLE ENTRY]

Entry: \$5 Entries: 3534/17.2K Prizes: \$75,000 Crowns: 5 My Entries: 0 No Multi-Entry

LIVE IN: 51:35:49 11/22 1:00 PM EST

CONTEST DETAILS RULES & SCORING Balance: \$5

NFL Classic

In salary cap contests, participants will create a lineup by selecting players listed in the Player Pool.

Contest results will be determined by the total points accumulated by each individual lineup entry (scoring rules summarized below).

Participation in each contest must be made only as specified in the Terms of Use. Failure to comply with these Terms of Use will result in disqualification and, if applicable, prize forfeiture.

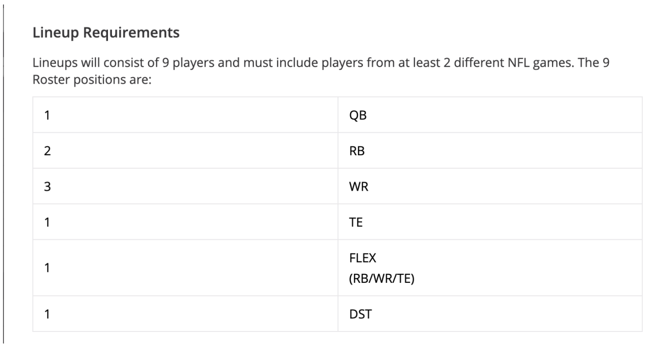
Scoring

| Offense | | Defense | |
|---|---|---|---|
| Passing TD | +4 Pts | Sack | +1 Pt |
| 25 Passing Yards | +1 Pt (+0.04 Pts/ Yards) | Interception | +2 Pts |
| 300+ Yard Passing Game | +3 Pts | Fumble Recovery | +2 Pts |
| | | Punt/Kickoff/FG Return for TD | +6 Pts |

Fig. 2. "NFL Classic" Scoring I.

| | | | |
|---|---|---|---|
| Rushing TD | +6 Pts | Interception Return TD | +6 Pts |
| 10 Rushing Yards | +1 Pt (+0.1 Pts/Yard) | Fumble Recovery TD | +6 Pts |
| 100+ Yard Rushing Game | +3 Pts | Blocked Punt or FG Return TD | +6 Pts |
| Receiving TD | +6 Pts | Safety | +2 Pts |
| 10 Receiving Yards | +1 Pt (+0.1 Pts/Yard) | Blocked Kick | +2 Pts |
| 100+ Receiving Yard Game | +3 Pts | 2 Pt Conversion/Extra Point Return | +2 Pts |
| Reception | +1 Pt | 0 Points Allowed | +10 Pts |
| Punt/Kickoff/FG Return for TD | +6 Pts | 1 – 6 Points Allowed | +7 Pts |
| Fumble Lost | -1 Pt | 7 – 13 Points Allowed | +4 Pts |
| | | 14 – 20 Points | |

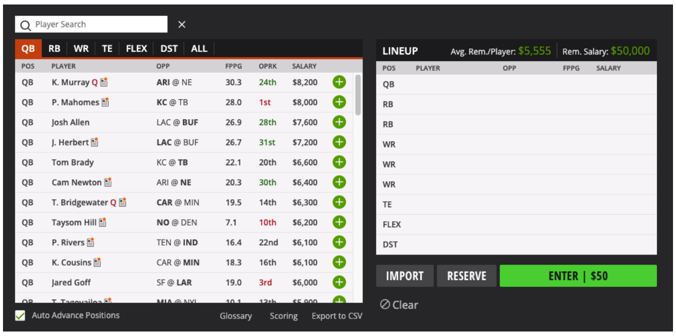
Fig. 3. "NFL Classic" Scoring II


Lineup Requirements

Lineups will consist of 9 players and must include players from at least 2 different NFL games. The 9 Roster positions are:

| | |
|---|---|
| 1 | QB |
| 2 | RB |
| 3 | WR |
| 1 | TE |
| 1 | FLEX (RB/WR/TE) |
| 1 | DST |

Fig. 4. Lineup requirements for "NFL Classic".


Player Search

QB RB WR TE FLEX DST ALL

| POS | PLAYER | OPP | FPPG | OPRK | SALARY |
|---|---|---|---|---|---|
| QB | K. Murray Q | ARI @ NE | 30.3 | 24th | \$8,200 |
| QB | P. Mahomes | KC @ TB | 28.0 | 1st | \$8,000 |
| QB | Josh Allen | LAC @ BUF | 26.9 | 28th | \$7,600 |
| QB | J. Herbert | LAC @ BUF | 26.7 | 31st | \$7,200 |
| QB | Tom Brady | KC @ TB | 22.1 | 20th | \$6,600 |
| QB | Cam Newton | ARI @ NE | 20.3 | 30th | \$6,400 |
| QB | T. Bridgewater Q | CAR @ MIN | 19.5 | 14th | \$6,300 |
| QB | Taysom Hill | NO @ DEN | 7.1 | 10th | \$6,200 |
| QB | P. Rivers | TEN @ IND | 16.4 | 22nd | \$6,100 |
| QB | K. Cousins | CAR @ MIN | 18.3 | 16th | \$6,100 |
| QB | Jared Goff | SF @ LAR | 19.0 | 3rd | \$6,000 |

Auto Advance Positions Glossary Scoring Export to CSV

LINEUP Avg. Rem./Player: \$5,555 Rem. Salary: \$50,000

| POS | PLAYER | OPP | FPPG | SALARY |
|---|---|---|---|---|
| QB | | | | |
| RB | | | | |
| RB | | | | |
| WR | | | | |
| WR | | | | |
| WR | | | | |
| TE | | | | |
| FLEX | | | | |
| DST | | | | |

IMPORT RESERVE ENTER | \$50

Clear

Fig. 5. Player Salaries and salary cap for "NFL Classic".

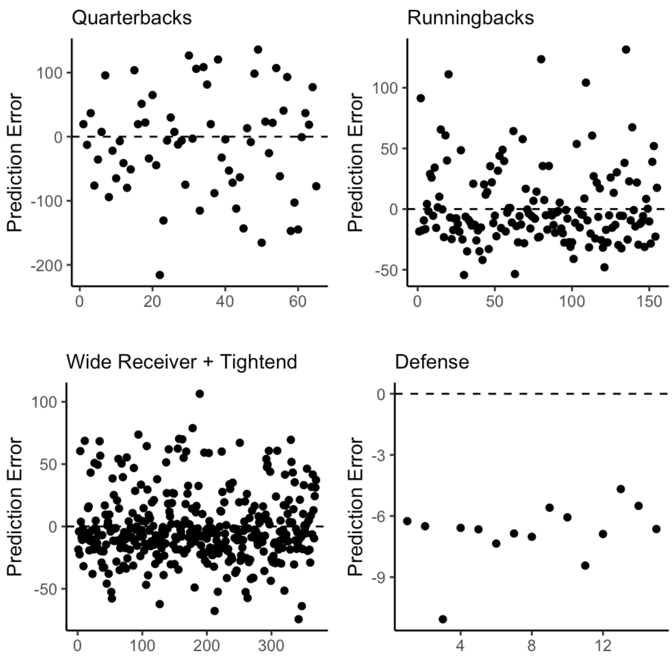



Fig. 6. Yards Prediction errors for different positions