Fantasy Football

Predicting High Scorers for the Upcoming Season

What is Fantasy Football?

- “Fantasy football is a game in which the participants serve as general managers of virtual professional American football teams” - Wikipedia
- Before the start of the NFL season, a group of 8-12 friends get together to hold a draft
  - During the draft, fantasy players take turns selecting real NFL players to be members of their fantasy team
- Each week of the NFL season, fantasy players are matched up against one another
  - If your players perform better than your opponent’s players, you win, feel good about life for a couple of days

The Problem

Year	Finish
2016	1st
2017	1st
2018	2nd
2019	**8th**

What Went Wrong in 2019?

2016 Draft (1st Place)

Round	Player Selected	Position Rank
1	David Johnson (RB)	1
2	Le'Veon Bell (RB)	3
3	Jordy Nelson (WR)	2

2019 Draft (8th Place)

Round	Player Selected	Position Rank
1	DeAndre Hopkins (WR)	5
2	Le'Veon Bell (RB)	16
3	Adam Thielen (WR)	64

The Challenge

- Leverage 12 weeks of learning about data science and predictive modeling to ensure fantasy football league domination
  - Gain insight into making better draft picks
  - Using a player's statistics from the previous season, create statistical model to project FF points scored for players at each position

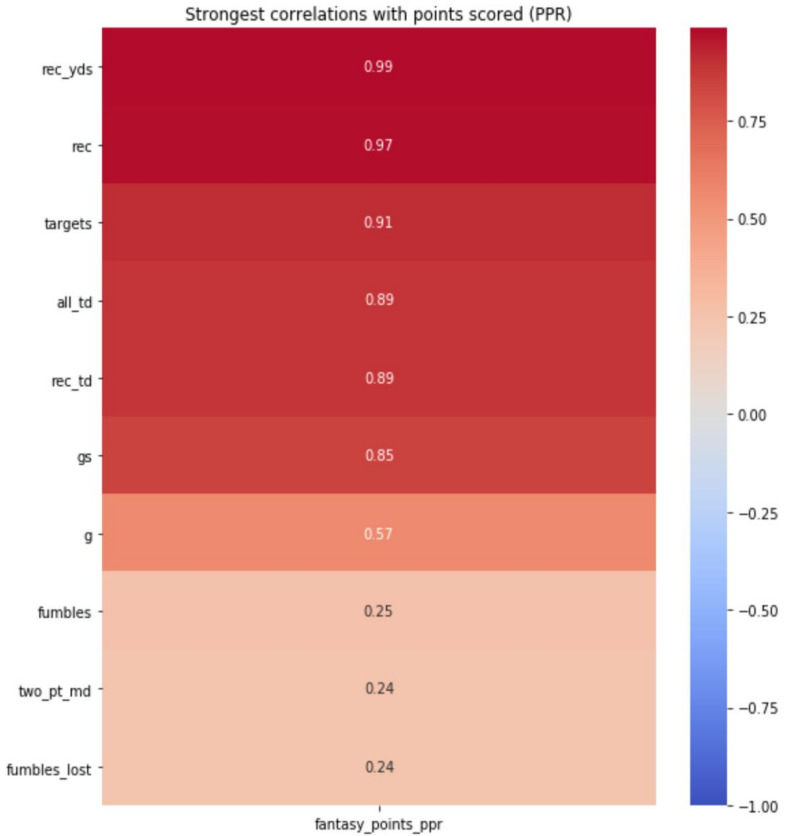
Getting Data (PFR)

					Games		Passing				
Rk	Player	Tm	FantPos	Age	G	GS	Cmp	Att	Yds	TD	Int
1	Christian McCaffrey*+	CAR	RB	23	16	16	0	2	0	0	0
2	Lamar Jackson*+	BAL	QB	22	15	15	265	401	3127	36	6
3	Derrick Henry*	TEN	RB	25	15	15	0	0	0	0	0
4	Aaron Jones	GNB	RB	25	16	16	0	0	0	0	0
5	Ezekiel Elliott*	DAL	RB	24	16	16	0	0	0	0	0
6	Dalvin Cook*	MIN	RB	24	14	14	0	0	0	0	0
7	Michael Thomas*+	NOR	WR	26	16	15	0	0	0	0	0
8	Travis Kelce*	KAN	TE	30	16	16	0	0	0	0	0
9	Nick Chubb*	CLE	RB	24	16	16	0	0	0	0	0

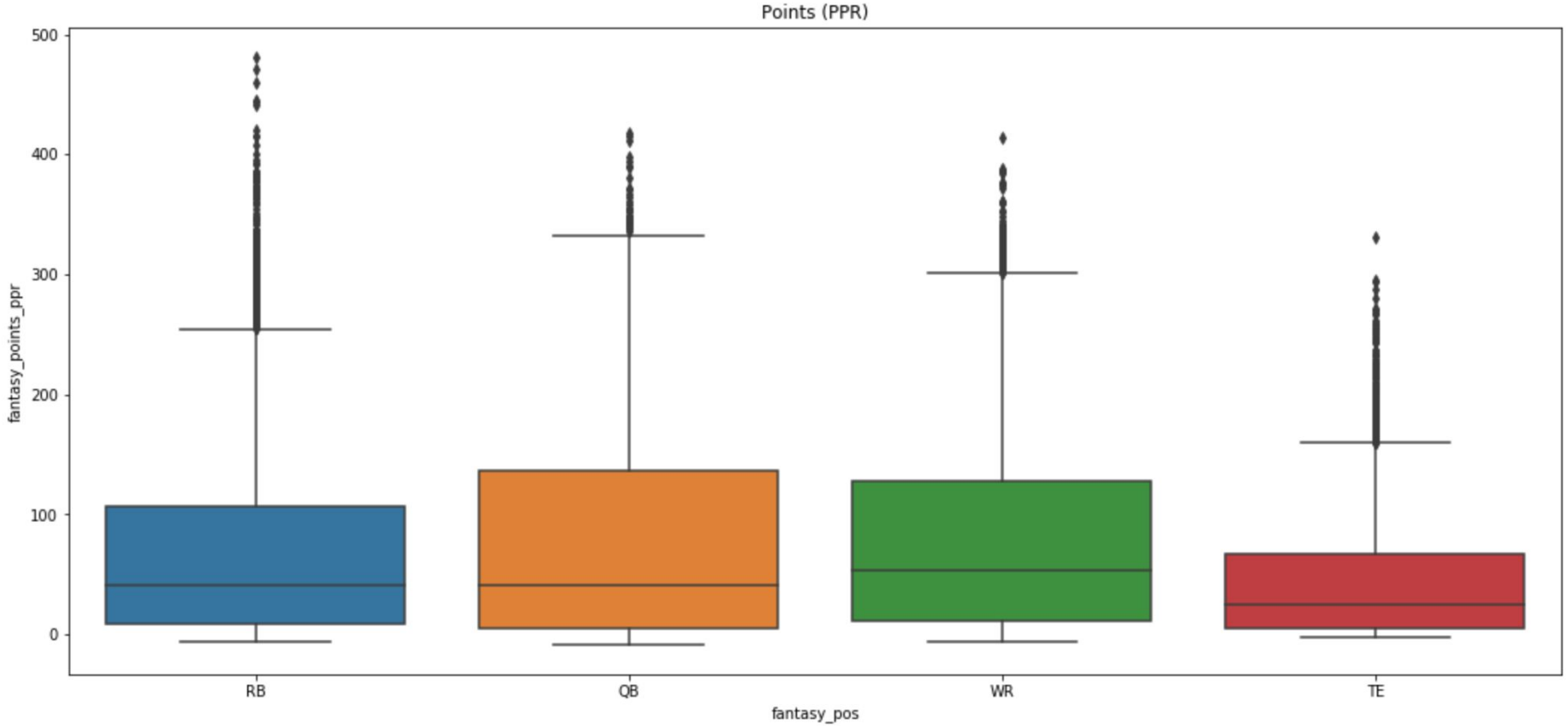
Housekeeping

- Dataframe had about 25,000 rows and 30 columns
- 100,000 missing values!
  - Many were missing because they were actually just 0
  - Others were missing because they weren't tracked until recently (e.g. targets)
- Logical imputation where possible
  - Filled in missing position if position was listed elsewhere in dataframe
  - Filled in missing targets with number of receptions (lower bound)

How to Draft Good Wide Receivers



Why You Should Draft a Running Back in the 1st Round



1st Pass at a Model

- Before trying anything more tedious, I wanted to get started with a basic linear model
- Fit a LASSO model on every feature attempting to predict the number of points scored by the player in the following season
  - Standardization with mean but not standard deviation
  - Regularization strength of 0.2
- Got R2 score of around 0.5

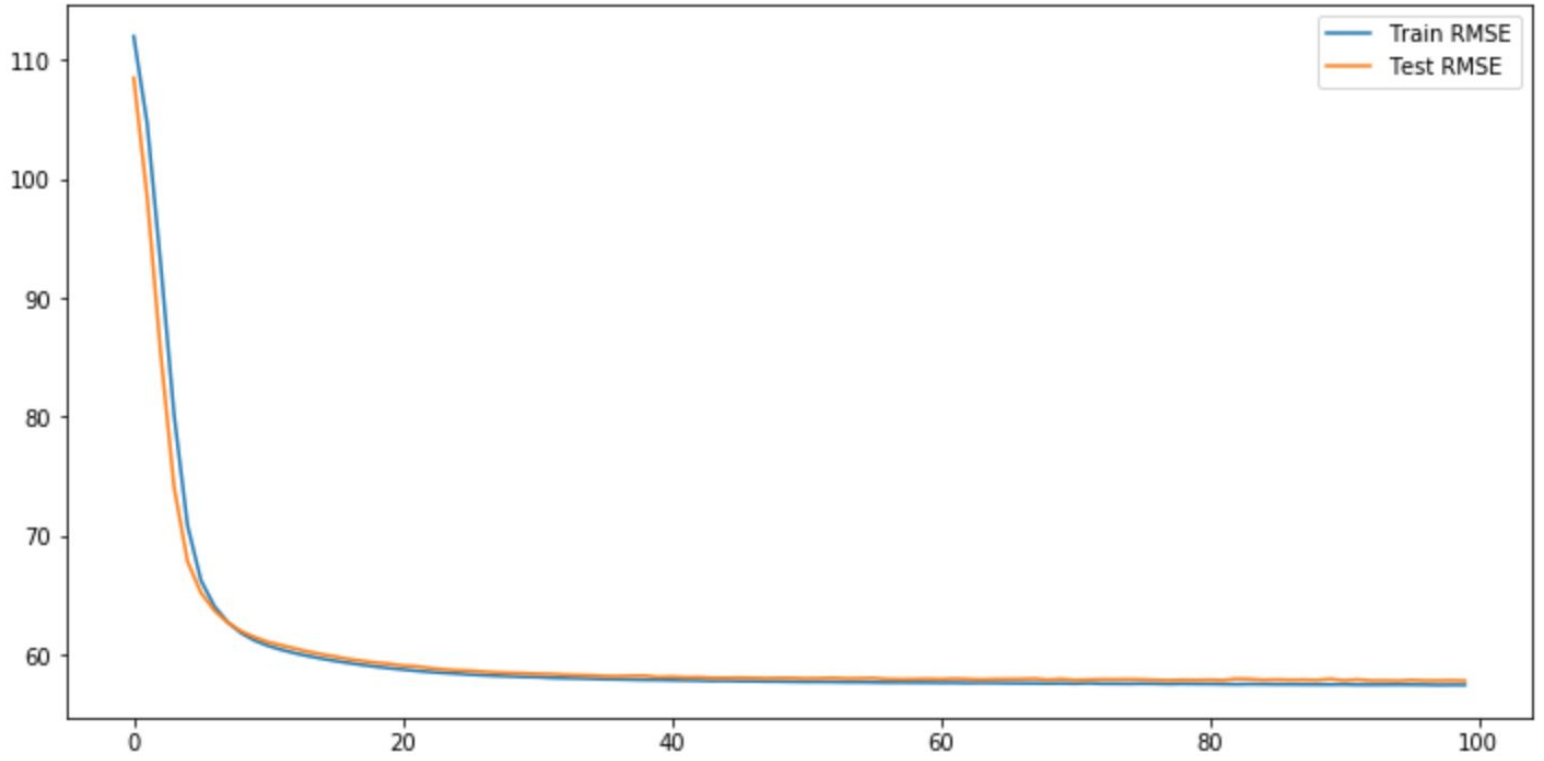
Linear Model Coefficients

Feature	Coefficient
Fantasy points (PPR) from previous season	0.52
Passing completions	0.29
Targets	0.28
Fumbles	0.26
Rushing yards	0.07

Neural Network

- Created sequential neural network with 1 hidden layer
  - Used every feature available
  - Target was fantasy points scored in the following season
  - Loss function: mean squared error
  - 100 epochs
- Final RMSE was around 60 fantasy points
- Fed 2019 data to model in order to make projections for the the upcoming 2020 season

Visualizing NN Performance



Projected Quarterback Rankings

Player	Projected Points
Lamar Jackson	321
Dak Prescott	275
Deshaun Watson	257
Jameis Winston	253
Russell Wilson	251

Projected Running Back Rankings

Player	Projected Points
Christian McCaffrey	351
Austin Ekeler	239
Ezekiel Elliot	233
Dalvin Cook	230
Derrick Henry	223

Projected Wide Receiver Rankings

Player	Projected Points
Michael Thomas	319
Chris Godwin	245
Julio Jones	231
Cooper Kupp	222
Keenan Allen	220

Projected Tight End Rankings

Player	Projected Points
Travis Kelce	194
George Kittle	192
Darren Waller	182
Mark Andrews	176
Zach Ertz	175

Reflections

- Looking at the projected rankings for next year, the model seems to project people to do almost exactly as good or bad (relative to others) as they did last year
  - This is a major limitation, no “risers” or “fallers”
- Model can’t make projections for rookies since they have no previous NFL stats
  - Also a fairly big limitation
- Model has no projections for defenses or kickers
- In summary, this model is not yet ready for production