Amar Srinivas
November 5, 2018

Fantasy Football Projections: Are ESPN's fantasy football projections accurate? Which site has the best fantasy football projections?

Fantasy football is a huge part of sports in the United States, and help casual fans understand and become more involved in a sport that they previously didn't feel fully invested in. We base our starting lineups based on the projections that were given to us by various fantasy football sites, however from week to week there are always projections that fall short and end up affecting our chances of winning negatively. This project is an attempt to figure out whether or not ESPN gives accurate enough fantasy projections as well as seeing which site among various fantasy football sites provides the most accurate projections.

The data has already been given thanks to the work of previous people, here is a sample of how it looks in jupyter:

	CMP	FG	FGA	FL	FPTS	INTS	PASS_ATT	PASS_TD	PASS_YDS	POSITION	Player	REC	REC_TD	REC_YDS	RUSH_ATT	RU
1	27.0			0.0	22.7	0.7	40.0	2.9	315.0	QB	Peyton Manning (DEN)				2.0	0.0
2	27.0			0.0	21.9	0.9	42.0	2.6	311.0	QB	Drew Brees (NO)				2.0	0.1
3	24.0			0.0	21.9	0.6	39.0	2.3	276.0	QB	Nick Foles (PHI)				4.0	0.2
4	20.0			0.0	20.6	0.7	33.0	1.6	254.0	QB	Colin Kaepernick (SF)				6.0	0.3
5	24.0			0.0	19.6	0.9	41.0	1.7	282.0	QB	Andrew Luck (IND)				4.0	0.2
6	19.0			0.0	19.1	0.7	30.0	1.8	225.0	QB	Russell Wilson (SEA)				6.0	0.2
7	26.0			0.0	19.0	1.2	44.0	2.0	305.0	QB	Matthew Stafford (DET)				2.0	0.1
8	23.0			0.0	18.1	0.7	36.0	1.9	278.0	QB	Tom Brady (NE)				2.0	0.1
9	20.0			0.0	17.6	0.6	30.0	1.7	250.0	QB	Aaron Rodgers (GB)				3.0	0.1
10	23.0			0.0	17.4	0.8	36.0	1.8	269.0	QB	Matt Ryan (ATL)				2.0	0.1
11	21.0			0.0	17.2	1.1	36.0	1.9	259.0	QB	Jay Cutler (CHI)				2.0	0.1
12	17.0			0.0	16.5	1.2	29.0	1.2	225.0	QB	Cam Newton (CAR)				6.0	0.3
13	20.0			0.0	16.4	1.1	34.0	1.5	241.0	QB	Ryan Tannehill (MIA)				4.0	0.2
14	18.0			0.0	16.4	0.6	29.0	1.2	212.0	QB	Robert Griffin III (WAS)				6.0	0.2
15	20.0			0.0	16.4	0.8	34.0	1.3	224.0	QB	Geno Smith (NYJ)				5.0	0.3
16	22.0			0.0	16.2	0.8	34.0	1.6	253.0	QB	Ben Roethlisberger (PIT)				2.0	0.1
17	19.0			0.0	16.2	0.7	31.0	1.5	214.0	QB	Alex Smith (KC)				4.0	0.2
18	22.0			0.0	15.7	1.2	33.0	1.7	261.0	QB	Tony Romo (DAL)				2.0	0.1
19	20.0			0.0	15.5	0.9	32.0	1.6	236.0	QB	Andy Dalton (CIN)				3.0	0.1
20	23.0			0.0	15.2	1.1	36.0	1.6	253.0	QB	Philip Rivers (SD)				2.0	0.1
21	22.0			0.0	14.9	1.1	36.0	1.3	252.0	QB	Ryan Fitzpatrick (HOU)				3.0	0.1

At first, I would being to look at relative and absolute error between what the projection for the player is and what the actual value of their fantasy points ended up being. I would test for statistical significance to measure whether or not there is bias within these projections or not. For example, when looking at various projections we can see someone like Matt Forte is projected to score 20 points while Ka'Deem Carey (Forte's backup) was projected to score one point. If Forte scores 19 points and Carey scores zero, both players have an absolute error of one point. However, one could argue that the near-miss on Forte's total is more impressive. We could also use relative error to measure the amount of biases in these projections.

$$\text{Relative error} = \frac{\text{projected fantasy points} - \text{actual fantasy points scored}}{\text{projected fantasy points}}$$

Data Project:

This is how I would break up the project:

1. Do exploratory data analysis and data cleaning on the given .csv files that we have a hold of, maybe scrape other site for further data in order to augment our predictions.

2. Engineer features to make sure the data frames we are working with have features that are of significance.

3. Rank features via correlation coefficients and single variable classification

4. Find the set of top features by using bootstrapped samples

5. Apply graphical causal modeling algorithms on the subset of features to search for the most causal relationships

6. Build a web app for visualization and strategies recommendation using Flask and AWS