Public Sentiment Toward the NFL

Can it predict weekly Fantasy Football outcomes?

- Ryan Timbrook
- Diego Vales
- David Madsen

Data Science to help you win…

PYTHON R

Using Data Science to Help Win Fantasy Football Games

BY GORDON YUN
NOVEMBER 6, 2018

COMMENTS 8

ENGAGEMENT

Social media and the rise of fantasy sports

by Kayla Lee

How Top Fantasy Draft Companies Use Social Media

Fantasy draft sites emerged with the rise in fantasy sports popularity—particularly, the popularity of fantasy on social media.

Fantasy Insights with Watson

Use AI to make better decisions in Fantasy Football

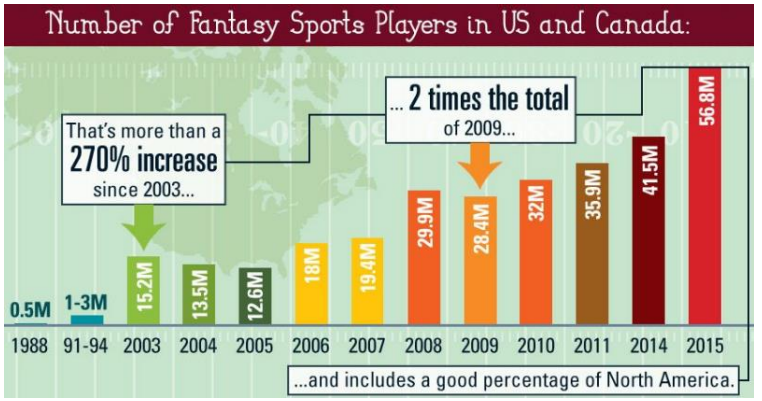
How it works

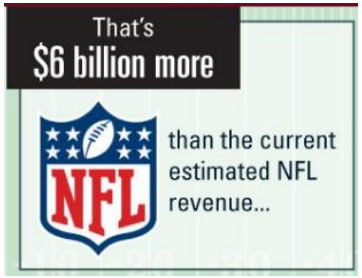
ESPN Fantasy Insights draws upon the latest in machine learning techniques to turn unstructured data into valuable insights. Nearly 10 million players rely on the combined resources of Watson Discovery and Watson OpenScale running on the IBM Cloud to give them a competitive edge.

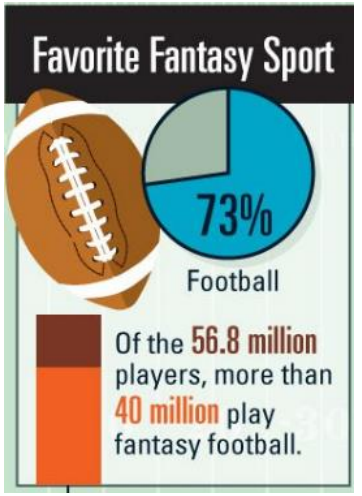
Get the latest updates on Twitter @IBMSports

Fantasy Football Industry

Fantasy Football is an $18.6 BILLION Market.



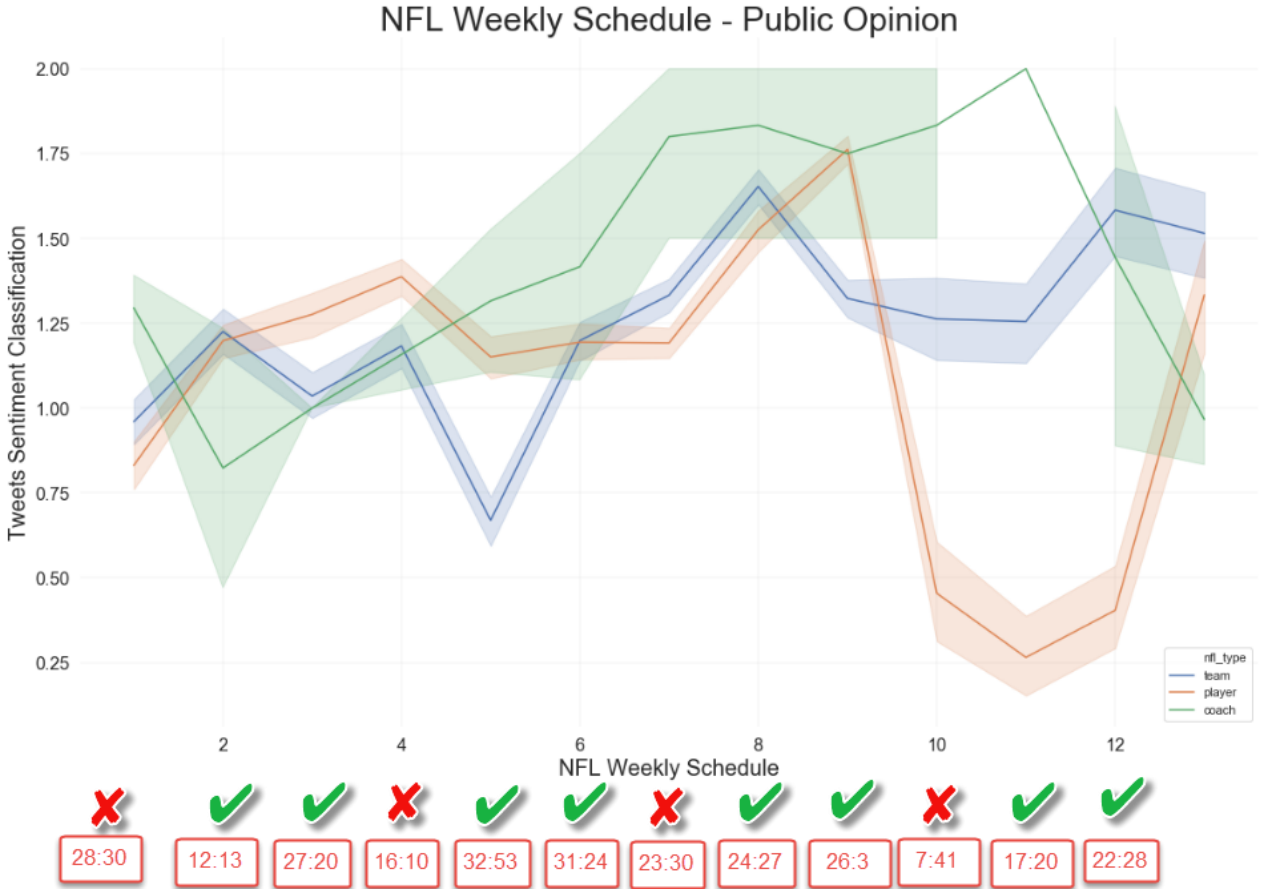




… Mapping Public Opinion To Weekly NFL Schedule …

Developer

Houston Texans
1st in AFC South



✘	✔	✔	✘	✔	✔	✘	✔	✔	✘	✔	✔
28:30	12:13	27:20	16:10	32:53	31:24	23:30	24:27	26:3	7:41	17:20	22:28

Team: Houston Texans

Coach: Bill O'Brian

Player: Deshaun Watson

VEDAR

1. positive sentiment: compound score >= 0.05
2. neutral sentiment: (compound score > -0.05) and (compound score < 0.05)
3. negative sentiment: compound score <= -0.05

compound score metric

1: Obtaining Twitter Data

Three Approaches Taken:

- Steaming API
  - Limited to 7 days data
- Full Archive Historical Search
  - Weekly search starting from NFL schedule week 1 to week 13
  - 10,000+ Tweets (Team, Coach, Player)
- Public Twitter Data Sets Pre-labeled
  - Twitter US Airline Sentiment Analyze how travelers in February 2015
  - First GOP Debate Twitter Sentiment Analyze tweets on the first 2016 GOP Presidential Debate

Lesions Learnt

Request Usage Burn…

API Pagination Lesions Learned…

Need for Search Operational Parameter Tuning

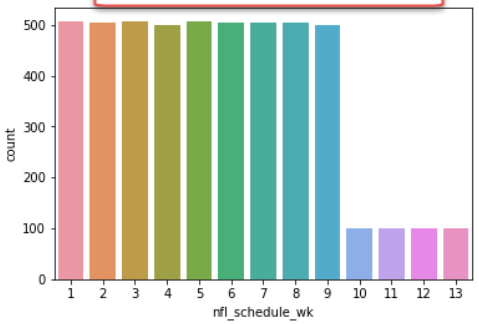
Use case for saving raw data to file – ran out of requests before report was written

Twitter API Wrappers…
-- Twython – doesn't have documentation on how to use premium vs. standard free

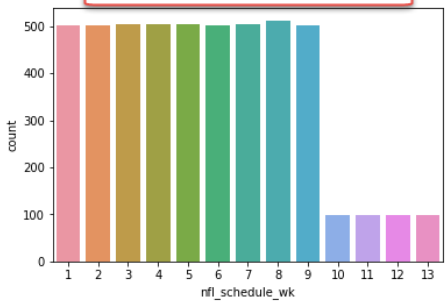
3: Twitter Data Sentiment Classification

VEDAR Classification on NFL Twitter Data

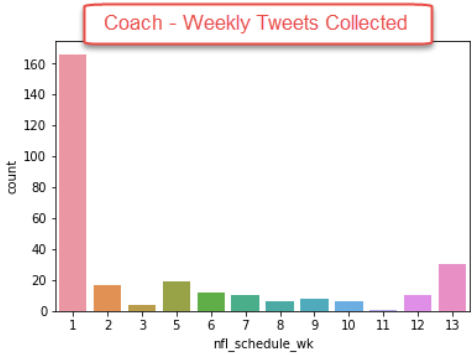
Player - Weekly Tweets Collected



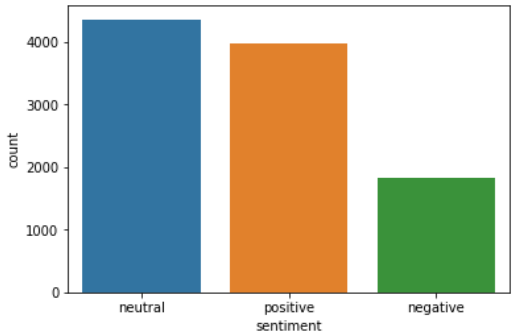
Team - Weekly Tweets Collected



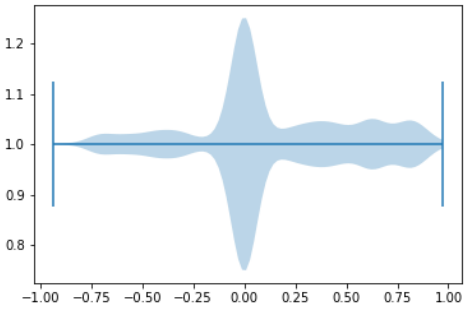
Coach - Weekly Tweets Collected

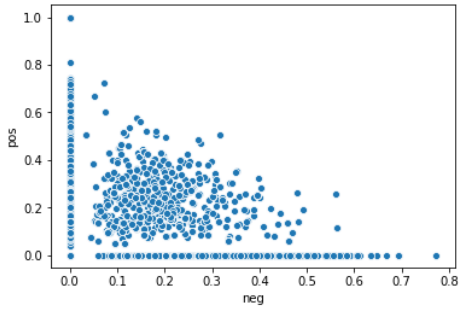


Polarity Scoring

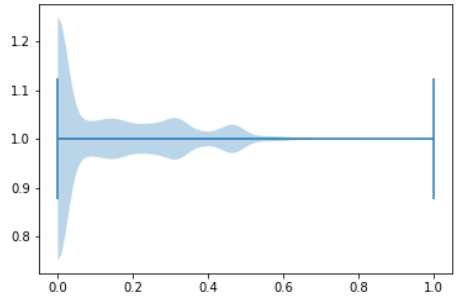


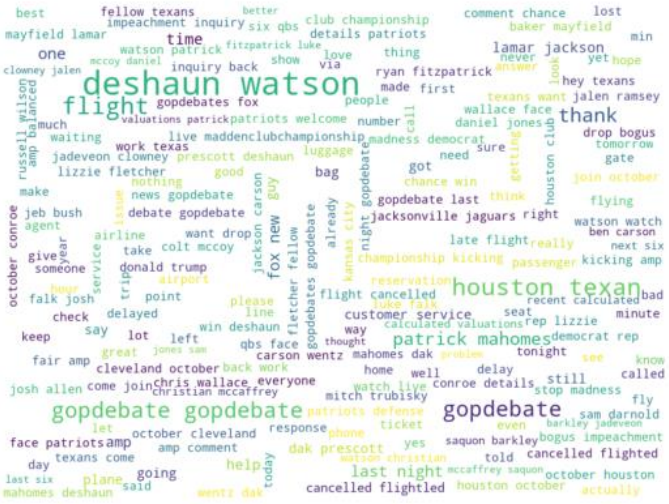
Compound Score



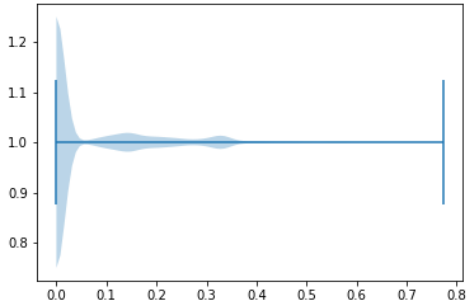


Positive Score





Negative Score



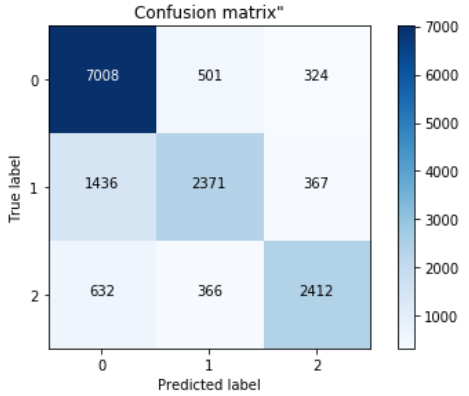
4: Sentiment Classification Model Results

50,000 Labeled Tweets

Best Fit Model: **77% Accuracy**

- Classification Type: **LinearSVM**
- Vectorizer: **TfidfVectorizer**
- N-Gram: **Bigram**
- Binary: **True**

Classification Trials

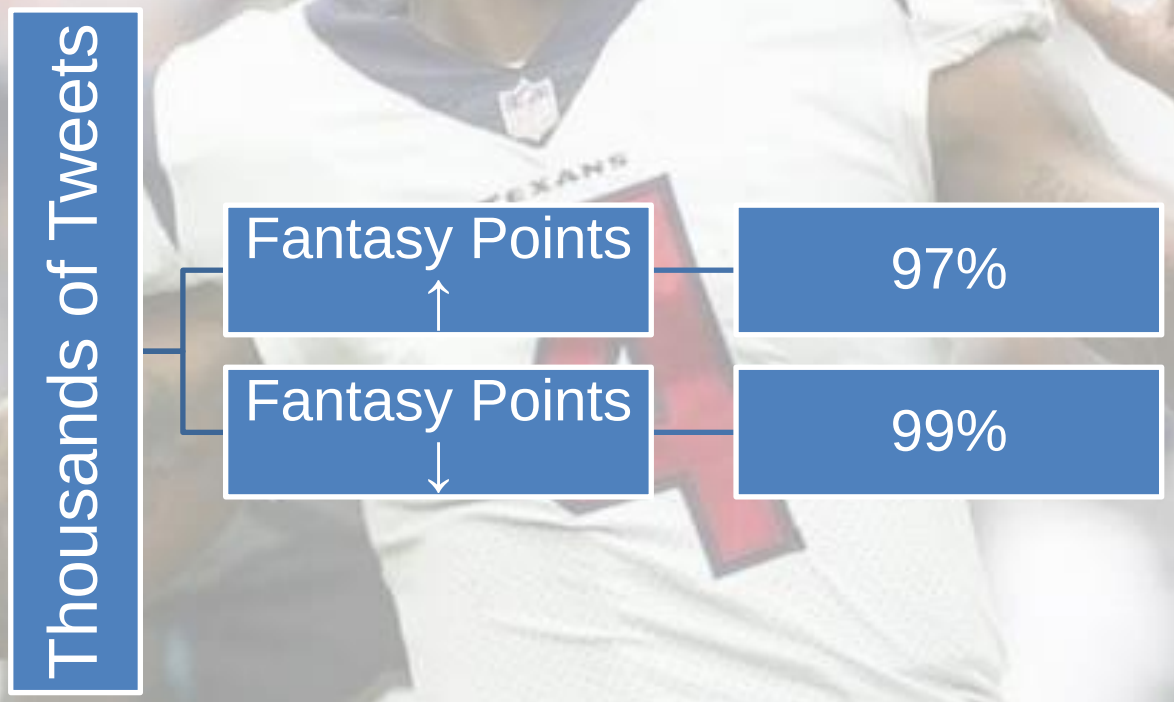


	precision	recall	f1-score	support
negative	0.77	0.89	0.83	7833
neutral	0.73	0.57	0.64	4174
positive	0.78	0.71	0.74	3410
micro avg	0.76	0.76	0.76	15417
macro avg	0.76	0.72	0.74	15417
weighted avg	0.76	0.76	0.76	15417

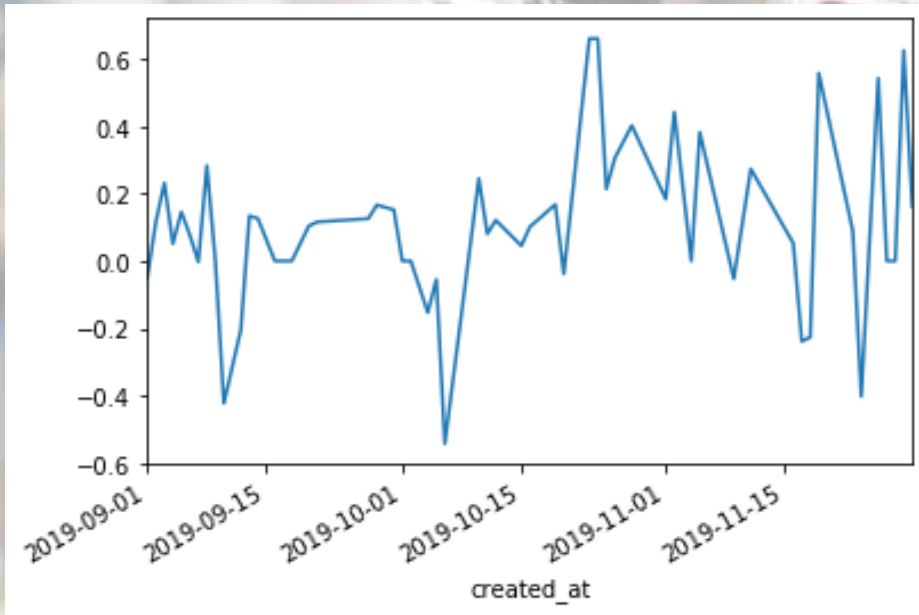
Model_Type	Vectorizer	N_Gram	Cross_Fold	Prediction_Accuracy	Total_Prediction_Points	Test_Recall_Score_Avg	Test_Precision_Score_Av	Train_Recall_Score_Av	Train_Precision_Score_Av	Total_Build_Tim	Total_Predict_Tim	Confusion_Matri
LinearSVM	tfidf	bigram	10	76.48	15417	0.7223	0.7599	0.87	0.9029	21.5007	0.0028	[[7008 501 324] [1436 2371 367] [632 366 2412]]
LinearSVM	count	unigram	10	76.45	3855	0.7393	0.7481	0.8745	0.8889	34.694	0.0018	[[1611 206 116] [260 717 93] [128 105 619]]
LinearSVM	count	bigram	10	75.77	3855	0.7383	0.7426	0.941	0.9457	62.9269	0.0011	[[1585 258 100] [255 701 81] [132 108 635]]
LinearSVM	tfidf	unigram	10	75.57	15417	0.7152	0.7556	0.8259	0.8682	11.2072	0.002	[[6905 510 343] [1496 2349 390] [692 335 2397]]
LinearSVM	count	bigram	10	74.53	15417	0.7243	0.7268	0.9643	0.9656	50.5876	0.0044	[[6305 1045 461] [1034 2671 424] [535 428 2514]]
LinearSVM	count	unigram	10	74.44	15417	0.7142	0.7177	0.9141	0.922	38.1779	0.0023	[[6316 977 470] [1123 2698 384] [544 442 2463]]
LinearSVM	count	unigram	10	74.31	15417	0.7168	0.7191	0.9143	0.9212	31.2185	0.0045	[[6304 985 532] [1050 2660 410] [566 418 2492]]
Multinomial_Naive_Bayes	count	bigram	10	72.89	15417	0.6698	0.7502	0.746	0.8294	0.4043	0.006	[[7299 283 193] [1889 1936 305] [1154 355 2003]]
Multinomial_Naive_Bayes	count	unigram	10	71.51	15417	0.6531	0.7245	0.7147	0.7924	0.2743	0.0036	[[7102 343 304] [1893 1893 443] [1070 340 2029]]
Multinomial_Naive_Bayes	count	unigram	10	71.51	15417	0.6468	0.7238	0.7078	0.7903	0.2601	0.003	[[7133 320 292] [1923 1870 415] [1139 304 2021]]

Predicting Player Outcomes from Tweets

Case: Deshaun Watson

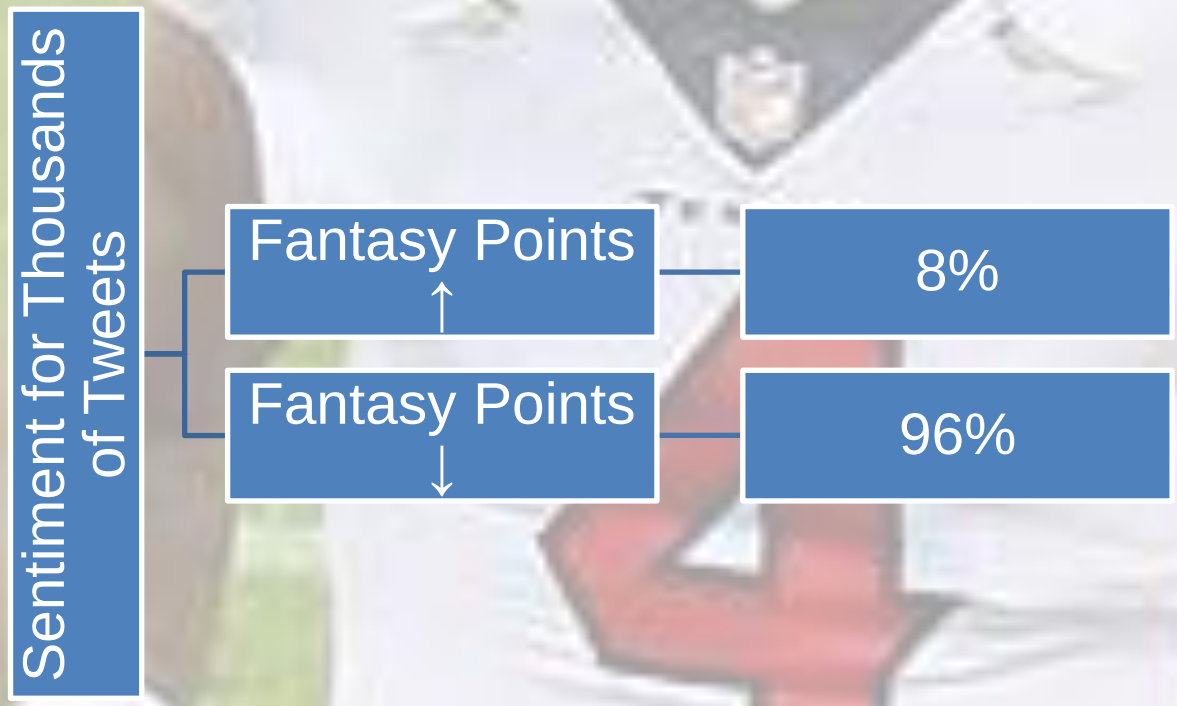


Tweet Sentiment Over Time



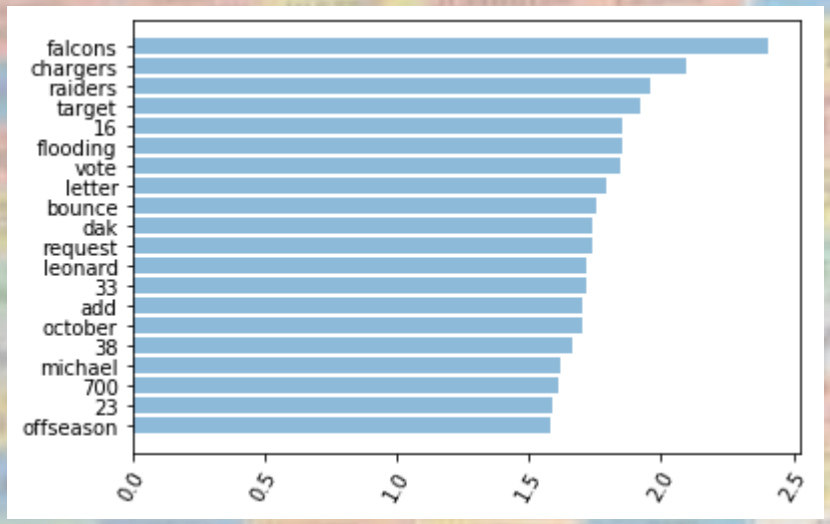
Predicting Player Outcomes from Tweet Sentiment

Case: Deshaun Watson Revisited

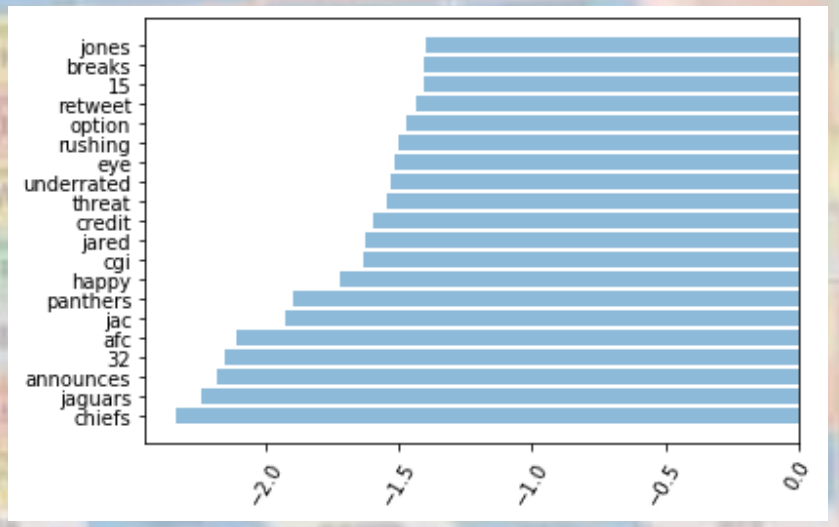


Feature Importance

Top Features



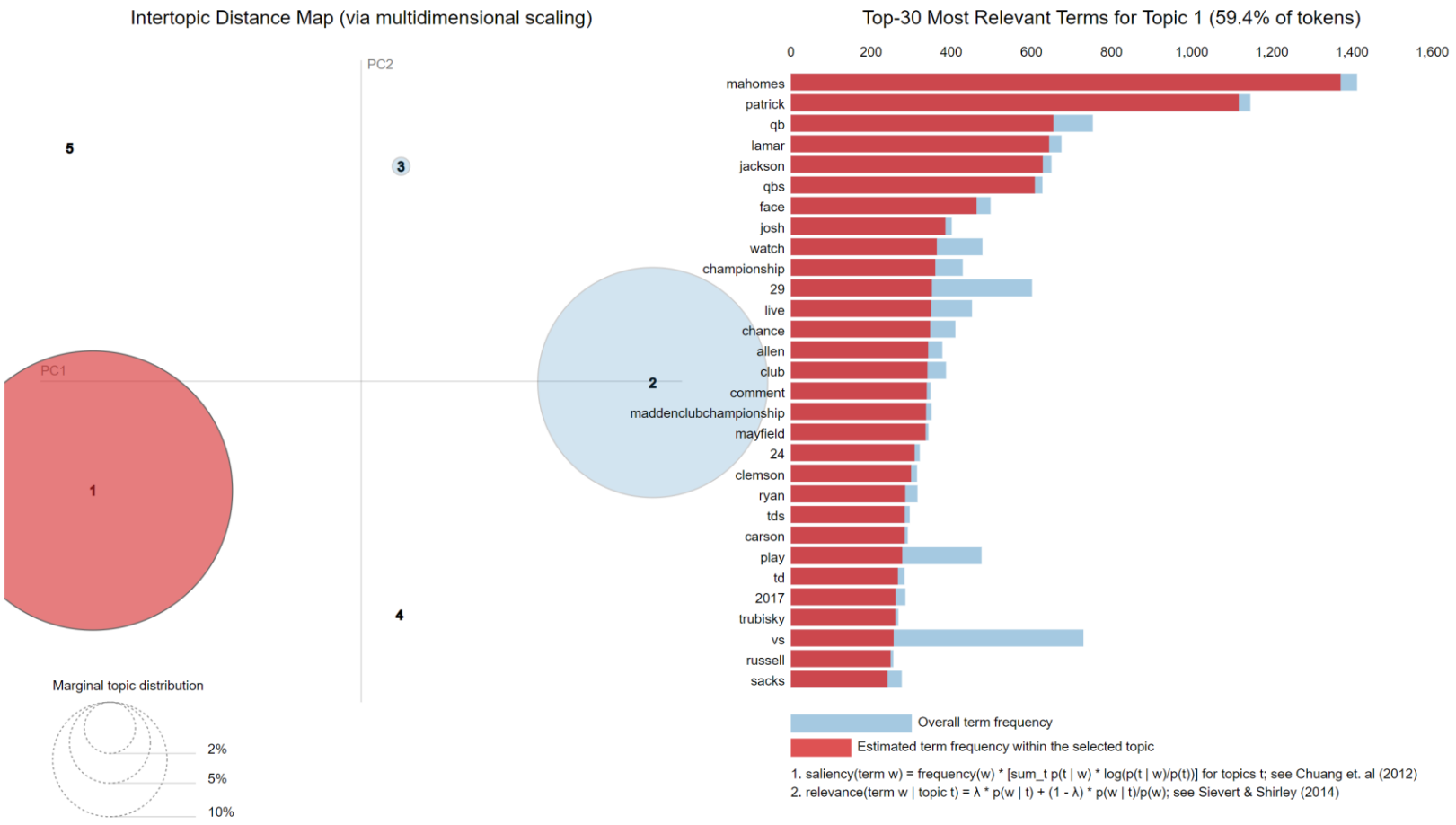
Bottom Features



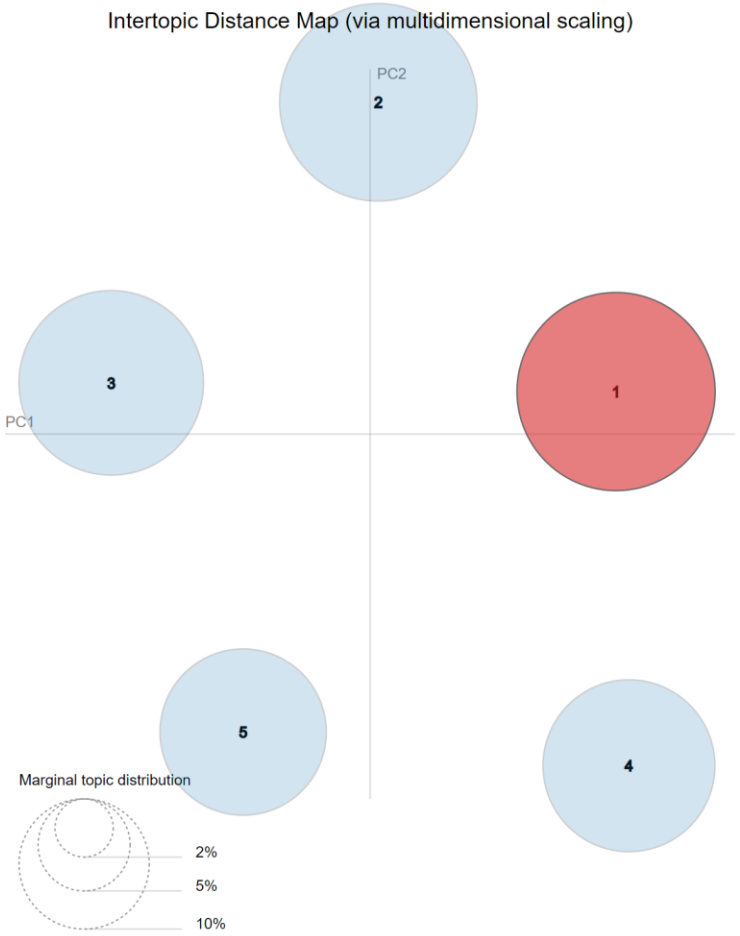
Topic Modeling

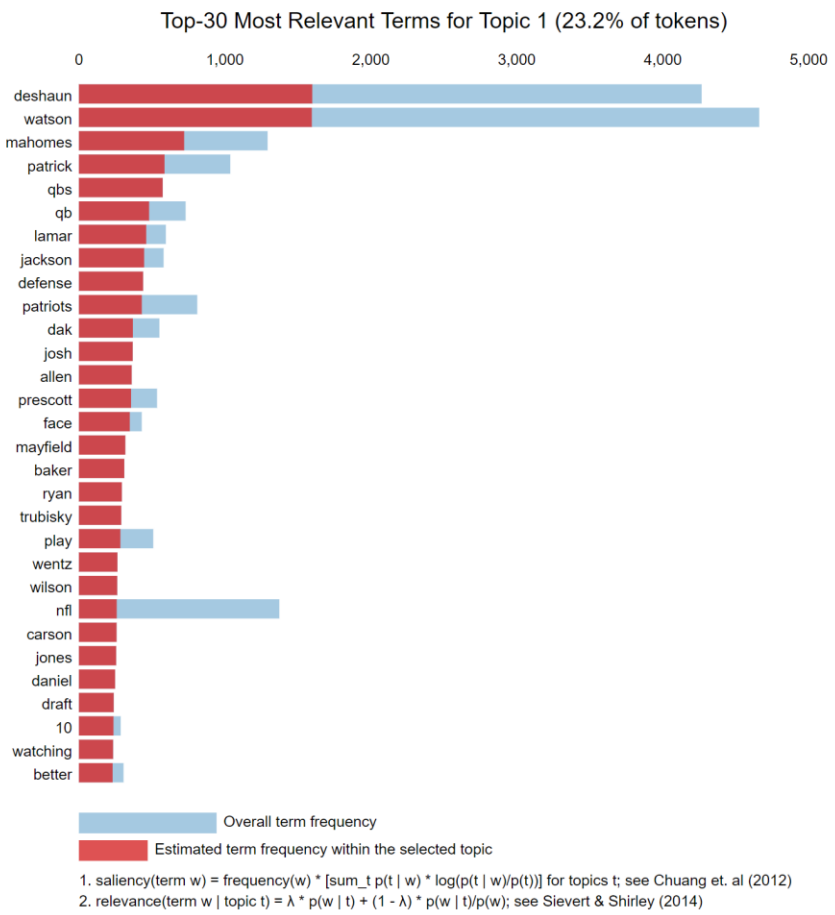
- Tweets
  - Player, Team, Coaches, Approaches
    - As files
    - As content
- Significance of emojis
  - Results
    - Included
    - Filtered

Topic Modeling - Files



Topic Modeling - Content





Question and Problem to Solve

Business Driver

- NFL revenue grew an estimated $900 million to $14 billion in 2017, in 2018 it generated about $15 billion.
- Fantasy football and the spread of legalized sports betting across the U.S. promises to lock in fans and keep them focused on the game.
- As a Fantasy football player, how can Data Science help me make the most intelligence selections when deciding my weekly roster? “How do I win more?”

Problem to Solve

- Identify if real-time public popular opinion on NFL teams, players, and coaches is a predictor of their weekly fantasy football stats.

About the Data:

- Text data will be mined from three primary sources

1) Fantasy football websites like Yahoo Sports: https://sports.yahoo.com/fantasy/
2) Twitter Social Media API streams
3) Facebook Social Media API streams

Media Headlines

FAN FAVORITE

You might not have heard of Fanatics yet—but it's taking over sports apparel one league at a time

Quartz

FANTASY FIGURES

How the $7 billion US fantasy football industry makes its money in 2017

The bizarre, multibillion-dollar industry of American fantasy sports

December 13, 2014 · Quartz

CHOOSE WISELY

Fantasy football now matters more than official NFL teams

Quartz

There's finally a great use for IBM's Watson: making fantasy football picks

Quartz

REAL MONEY

The two charts that explain why Yahoo is pushing deeper into fantasy sports

Quartz

Why investors are pouring millions into fantasy sports

Quartz

Fantasy Football Industry

Q: How did fantasy go from weekender hobby to big-time business?

A: 2 words – DAILY FANTASY

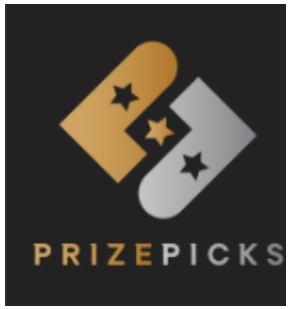
In 2006, Congress passed the Unlawful Internet Gambling Enforcement Act (UIEGA), which banned online poker – deemed a game of chance – but left room for fantasy sports wagers – deemed a game of skill.

DRAFT KINGS and FanDuel

$$ the two largest daily fantasy services – will each pay out more than **$2 billion** in winnings this year. That's a big chunk of change.

$$$ But with entry fees - from $2 to more than $5,000 - generating **$2.6 billion** this year, bearish projections continue.

Notable Fantasy Football Brands



DraftKings Promo Code and Review – Get 3 Free Months of RotoGrinders Premium for DraftKings