

Sports!

Shomik Jain

2018 NCAA DIVISION I MEN'S BASKETBALL CHAMPIONSHIP BRACKET

First Round
MARCH 15-16

Second Round
MARCH 17-18

Regional
Semifinals
MARCH 22-23

Regional
Finals
MARCH 24-25

National
Semifinals
MARCH 31

National
Semifinals
MARCH 31

Regional
Finals
MARCH 24-25

Regional
Semifinals
MARCH 22-23

Second Round
MARCH 17-18

First Round
MARCH 15-16

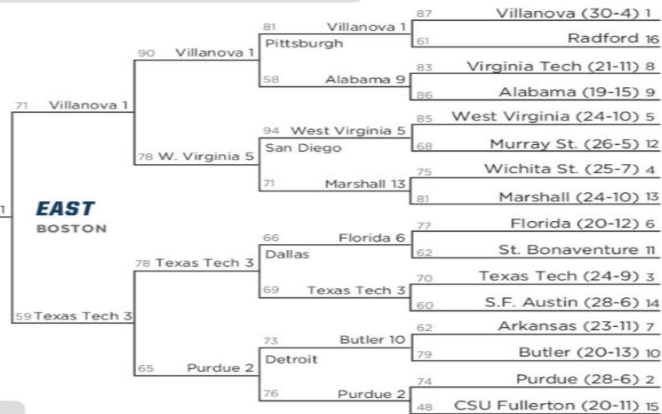
FIRST FOUR

16 LIU Brooklyn (18-16) 61	11 St. Bona. (25-7) 65	DAYTON MARCH 13-14	46 NC Central (19-15) 16	56 Arizona St. (20-11) 11
16 Radford (22-12) 71	11 UCLA (21-11) 58	Watch On tru®	W 64 Tex. Southern (15-19) 16	MW 60 Syracuse (20-13) 11



FINAL FOUR
SAN ANTONIO
MARCH 31 AND APRIL 2

NATIONAL CHAMPIONSHIP
APRIL 2



#MarchMadness

Watch the tournament on these networks
or online at NCAA.COM/MARCHMADNESS



March 15 and 17 First-/second-round sites: Pittsburgh, Wichita, Dallas, Boise; March 16 and 18 First-/second-round sites: Charlotte, Detroit, Nashville, San Diego; March 22 and 24 regional sites: Los Angeles, Atlanta; March 23 and 25 regional sites: Omaha, Boston

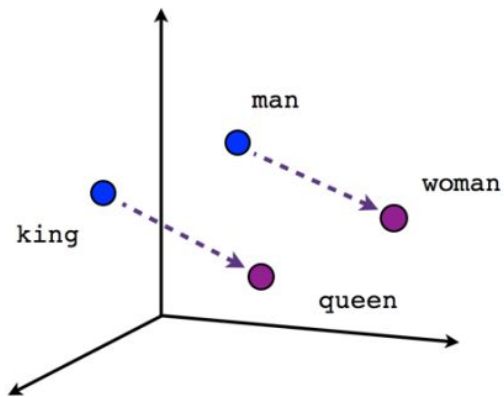


The NCAA opposes all forms of sports wagering

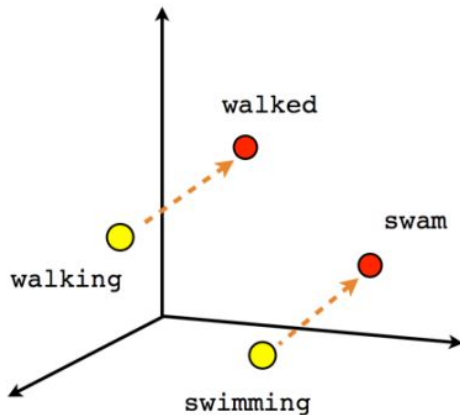
What is the relationship between...

- Whether a Team Wins or Loses
- Average Sentiment of Tweets About the Team
 - Sentiment: computational measure of whether a tweet is positive, negative, or neutral
- Basketball Power Index (BPI) Score
 - ESPN's game prediction measure
 - Simulates the game 10,000 times

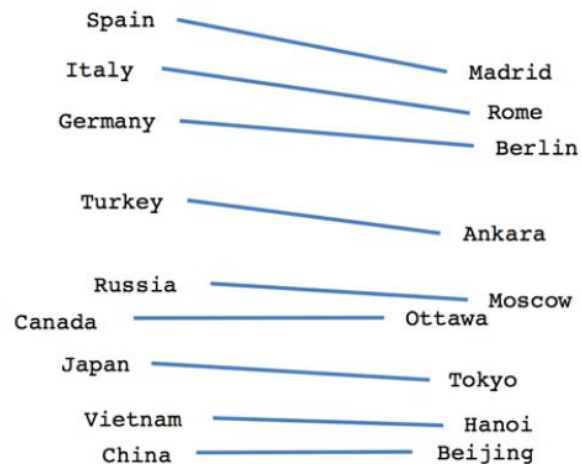
Sentiment Analysis: RNN & Word Embeddings



Male-Female



Verb tense



Country-Capital

Data

game_id	round	team	seed	win	win_margin	bpi	win_bpi	num_tweets	avg_pos_conf	win_pos_conf
1	1	Virginia	1	0	-20	98.5	1	488	0.63386	0
1	1	UMBC	16	1	20	1.5	0	97	0.68306	1
2	1	Georgia St	15	0	-15	5.5	0	52	0.56677	0
2	1	Cincinnati	2	1	15	94.5	1	395	0.59270	1
3	1	Wright St	14	0	-26	8.5	0	91	0.61571	0
3	1	Tennessee	3	1	26	91.5	1	778	0.61864	1
4	1	Arizona	4	0	-21	79.9	1	609	0.63861	0
4	1	Buffalo	13	1	21	20.1	0	197	0.72137	1
5	1	Davidson	12	0	-5	38.8	0	489	0.62061	0
5	1	Kentucky	5	1	5	61.2	1	686	0.66027	1
6	1	Miami	6	0	-2	57.3	1	180	0.50855	0
6	1	Loyola Chicago	11	1	2	42.7	0	2563	0.63720	1
7	1	Texas	10	0	-4	47.6	0	1125	0.56403	0
7	1	Nevada	7	1	4	52.4	1	1817	0.61897	1
8	1	Creighton	8	0	-10	35.2	0	261	0.65344	0
8	1	Kansas St	9	1	10	35.2	1	30	0.74587	1

Nonparametric Statistics

- Nonparametric Statistics
 - fairly general assumptions about probability distributions
- Matched Pairs Experiment
 - n pairs of random and independent observations in form (X_i, Y_i)
- Sign Test
 - Is the distribution of the X_i 's the same as the Y_i 's?
 - Test Statistic: $\sum 1(X_i > Y_i)$ = count # of times $X_i > Y_i$

Results

1. Winning Team BPI > Losing Team BPI
 - BPI is a reliable predictor of whether a team is likely to win/lose
2. Winning Team Sentiment \approx Losing Team Sentiment
 - Sentiment: not accurate for whether a team is likely to win/lose
3. Winning Sentiment > Losing Sentiment (when BPI Fails)
 - Examine games where BPI failed to accurately predict winner
 - "Upset" => higher positive sentiment for winning team

Bottom Line...

- Social Media & Sentiment
 - meaningful insights, trends/potential signals in unstructured data
- Applications Beyond Sports
 - Politics, Business (Market Trends), ...

