

Analyzing the Impact of Twitter Bots on the 2020 U.S. Presidential Election Discourse

#NLP

#Metis

#UnsupervisedLearning

Elliot Wilens

Metis Data Scientist

wilensel@gmail.com

363,000

Tweets Scraped

Oct 1 – Nov 2, 2020



#TrustTheProcess (Project Workflow)

- Preprocessing
 - 300+ custom n-grams
 - Lemmatization (i.e. ‘mice’ → ‘mouse’,)
- Word Vectorization
 - TF/IDF (Term-Frequency * Inverse Doc Frequency)
 - Significant improvement over count vectorizer
- Topic Modeling
 - NMF (Non-Negative Matrix Factorization)
 - 30 topics
- Bot / Not Bot

Bot/Not Bot

- Model came from Chris Doenlen's 'twitter-bot detection' project (thanks Chris!)
 - <https://medium.com/the-innovation/twitter-bot-or-not-cc2ad52b36ba>
- Bot_proba = a user's probability of being a bot (according to this model)

2020 Election Tweets were about...

anti-trump supporters	#hunterbiden	trump responsible for covid	covid stats	#blm	#economy	past candidates
	#votetrump					
#covidhoax	#vote	supreme court nom.	#factcheck	federal stimulus		truth and lies
		trump rally drama				
#polling	foreign affairs		healthcare for all	covid policy		voter
		#votebiden				
trump rally covid infections	vp candidates	trump covid diagnosis	#presidentialdebate			
			trump distrust science	town halls	#media	

negative campaign ads

Bot_proba 65.78%



ileiarez
@ileiarez



@JoeBiden Maybe you should have told the truth before because the election is tomorrow and people believe everything trump says even though it's all lies

Reply Retweet Favorite More

1:44 AM - 02 Nov 20 · Embed this Tweet

The model did well capturing some rather obscure topics...



quvester
@quvester



@JoeBiden Its time for you to run your ads. Did you see the video? He says he's good and if it's a lie, IT IS WHAT IT IS. We have a race to run.

Reply Retweet Favorite More

10:15 AM - 02 Nov 20 · Embed this Tweet

Bot_proba 94.26%



anthony76365973
@anthony76365973



@JoeBiden I need to see more of your adds attacking Trump. All I see is his lies about you. I want to vote for you. Please dont let him get away with these things.

Reply Retweet Favorite More

5:32 PM - 18 Oct 20 · Embed this Tweet

Bot_proba 31.41%



resistanceblue1
@resistanceblue1



@omarirob55 @Bill_Thompson We can do that. They've lied about 8 times Now, i subtract that from Trump's total... I'm looking at 35 lies by Kamala and Biden vs... 20,218 for Trump
<https://t.co/4AmhnpHxYb>

Reply Retweet Favorite More

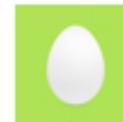
11:00 AM - 11 Oct 20 · Embed this Tweet

Bot_proba 46.14%

<https://twitter-bot-or-not.herokuapp.com/predict>

#covidhoax

Bot_proba 10.55%



gsmithnewport
@gsmithnewport

[User](#) [Follow](#)

@JasonMillerinDC @realDonaldTrump If Biden wins Americans win, the economy wins and the virus goes bye bye.

[Reply](#) [Retweet](#) [Favorite](#) [More](#)

9:52 PM - 13 Oct 20 · Embed this Tweet



shot_gangster
@shot_gangster

[User](#)

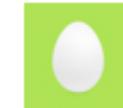
[Follow](#)

@davidaxelrod @realDonaldTrump Hard to believe he's recovering from a deadly virus. He looks more vigorous now than he did 4 years ago. Vote Trump!
<https://t.co/eTvxYkrOuy>

[Reply](#) [Retweet](#) [Favorite](#) [More](#)

2:01 PM - 12 Oct 20 · Embed this Tweet

Bot_proba 73.08%



nickprice72
@nickprice72

[User](#) [Follow](#)

@Ah114898 @realDonaldTrump So, if Biden wins and another 200,000 people die because, well it's a virus and we still don't know everything about it, do we get to politicize their deaths also??

[Reply](#) [Retweet](#) [Favorite](#) [More](#)

12:19 AM - 02 Nov 20 · Embed this Tweet

Bot_proba 97.93%

#1 Tweet from each topic (in terms of NMF coeff.)...

#votetrump



dianes01997758
@dianes01997758



[Follow](#)

@realDonaldTrump TRUMP'S not only
bring PEACE here, But he's bringing it
AROUND the WORLD. GREAT MAN and
LEADER. USA needs TRUMP!

[Reply](#) [Retweet](#) [Favorite](#) [More](#)

10:59 AM - 02 Nov 20 · Embed this Tweet

Bot_prob 68.35%

#votebiden



cargal
@cargal



[Follow](#)

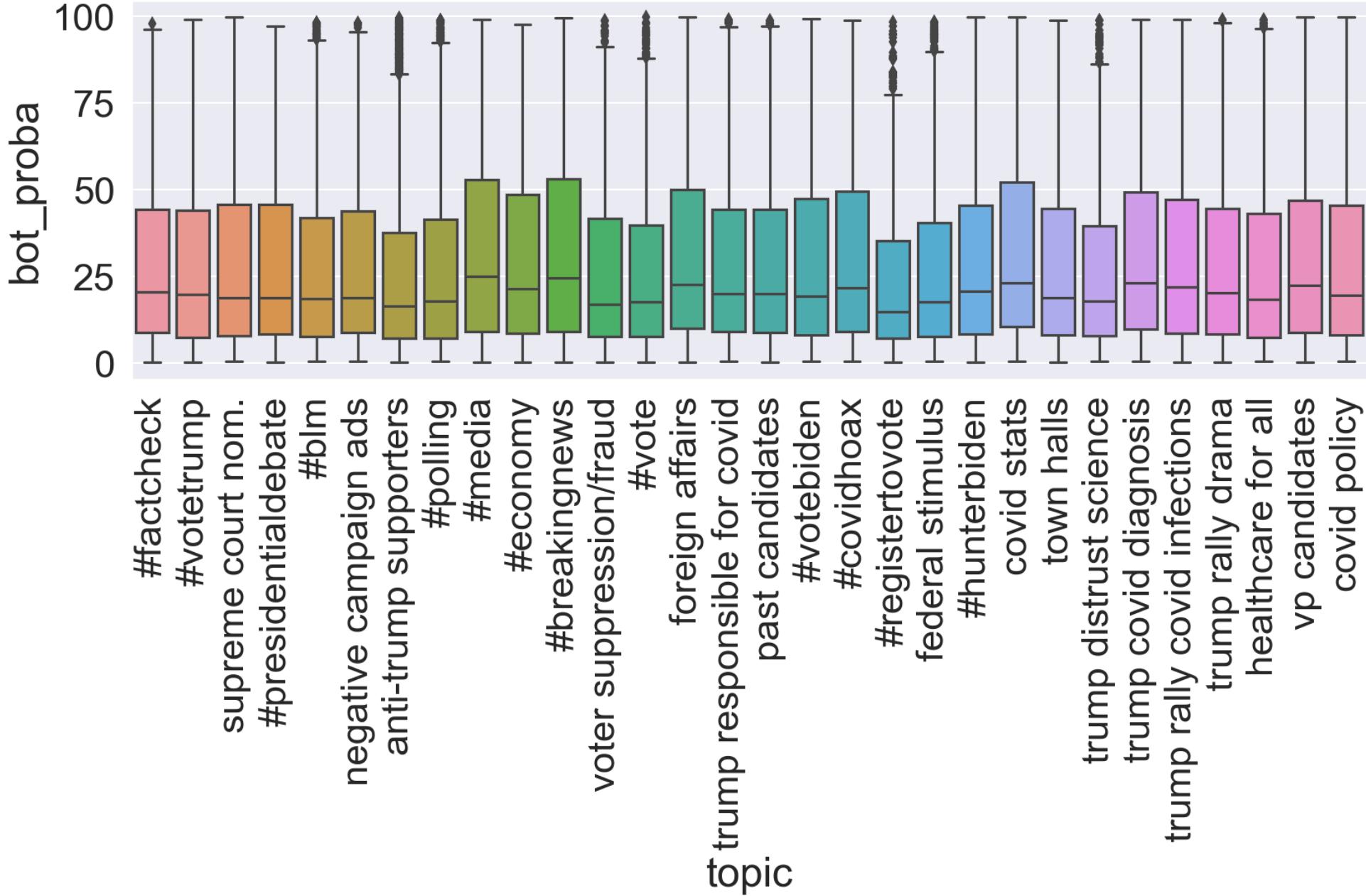
Trump cares about Trump He doesn't care
about the American people He doesn't care
about all those that have died He wants to
be president for life and will do anything to
get there God help us if he wins
#VoteBlueLikeYourLifeDependsOnIt
Because it does

[Reply](#) [Retweet](#) [Favorite](#) [More](#)

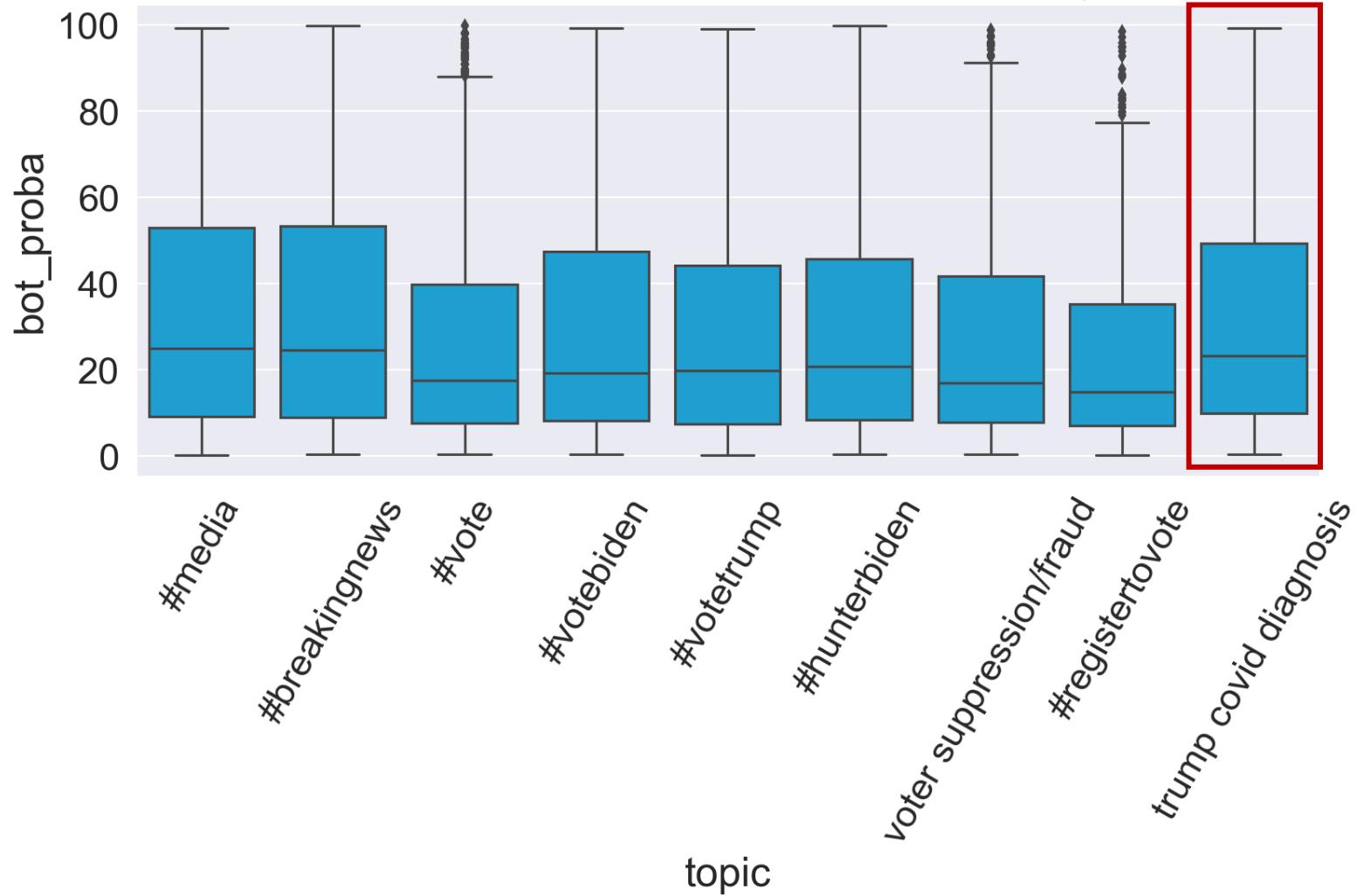
3:18 PM - 09 Oct 20 · Embed this Tweet

Bot_prob 10.97%

Bot probabilities for tweets in each topic



But...bots can be valid sources of information, too...



Bot_proba increases for topics with high media coverage



Twitter Bot or Not

Enter a Twitter handle

Submit

Prediction for @cnn

Bot

Probability of being a bot: 91.13%

Twitter Bot or Not by Chris Doenlen |

<https://twitter-bot-or-not.herokuapp.com/predict>

Model Pros

- Most topics are easily understandable thru topic words (see Appendix)
- Accurately captured a few surprising topics
 - Hunter Biden
 - Negative campaign ads
 - Hillary Clinton #VOTE mask

Model Cons

- Some topics present in all others (even with higher n. components/topics)
 - Lies
 - Race
- Needs more preprocessing for further research

#UpNext

- Further text preprocessing (more stop words, n_grams, etc.)
- More bot analysis
- Try more topics (50?)

Appendix

anti-trump-supporters	trumpsupporter	since	hate	call	party
#vote	vote	votejoebiden	ballot	votedonaldtrump	party
covid stats	covid	positive	test	tested	hospital
vp candidates	kamalaharris	mikepence	ballot	nancypelosi	votejoebiden
#presidentialdebate	debate	presidential	virtual	moderator	commission
trump responsible for covid	trump	fault	tweet	response	base
past candidates	barackobama	hillaryclinton	administration	bush	economy
#hunterbiden	hunterbiden	story	twitter	post	report
trump rally drama	campaign	negative	event	bus	running
#registertovote	voting	party	ballot	bidenharris	mail
#blm	black	racist	woman	white	community
covid policy	mask	wear	wearing	mandate	without
negative campaign ads	lie	truth	lying	liar	corruption
voter suppression/fraud	voter	fraud	suppression	undecided	ballot
supreme court nom.	question	answer	court	real	tough
#economy	tax	pay	money	cut	return
foreign affairs	china	russia	business	deal	money
#media	cnn	foxnews	msnbc	abc	fox
trump distrust science	believe	hard	word	god	stupid
#votebiden	life	save	lost	saved	risk
#polling	poll	lead	ahead	winning	hillaryclinton
healthcare for all	plan	pandemic	health	healthcare	administration
#factcheck	fact	check	claim	TRUE	truth
town halls	town	hall	nbc	abc	rating
federal stimulus	help	god	stimulus	russia	bill
#votetrump	world	leader	real	free	rest
trump covid diagnosis	whitehouse	event	walter	reed	return
#covidhoax	virus	death	spread	deadly	pandemic
#breakingnews	news	real	breaking	report	abc
trump rally covid infections	rally	hold	maga	supporter	superspreaders

Tech Stack

- Python3
 - Twint
 - Scraping Tweets
 - Chris Doenlen’s “Twitter-bot Detection”
 - <https://github.com/scrapfishies/twitter-bot-detection>
 - Requires Tweepy library
- Sci-Kit Learn
- NLTK
- Tableau