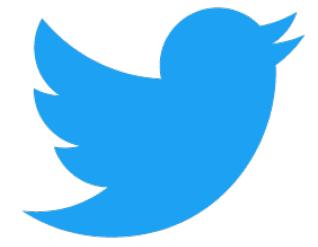


JON LINDENAUER

RUSSIAN TWITTER TROLLS

Natural Language Processing

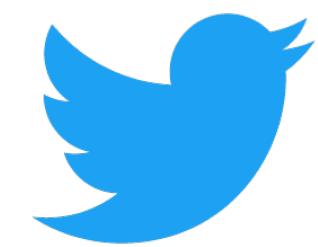
MOVITATION AND GOALS



- The Russian Government sponsors the Internet Research Agency to inflame and divide the public via Twitter, Facebook and other social media.
- Can we use NLP to characterize what the Russian Troll Tweets contain and classify new tweets?



DATA



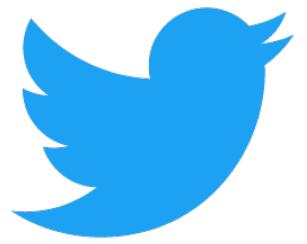
- Nearly 3 millions tweets in many languages made available by FiveThirtyEight and Clemson University.
- Used tweets from March 2017 to August 2017.
 - 191,000+ English language Right Troll tweets (no retweets) used for modeling.
 - Included only Trolls with 20 or more tweets.



 FiveThirtyEight



TROLL TWEETS



- “Marshawn Lynch arrives to game in anti-Trump shirt. Judging by his sagging pants the shirt should say Lynch vs. belt”
- “It took Hillary abt 5 minutes to blame NRA for madman's rampage, but 5 days to sorta-kind'a blame Harvey Weinstein 4 his sexually assaults.”
- “Police chief kills lost puppy for target practice”





■ Word Cloud of unfiltered data.

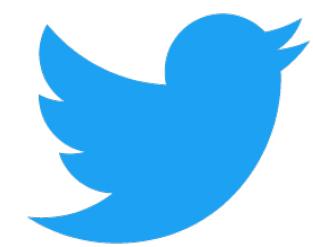
Left Trolls



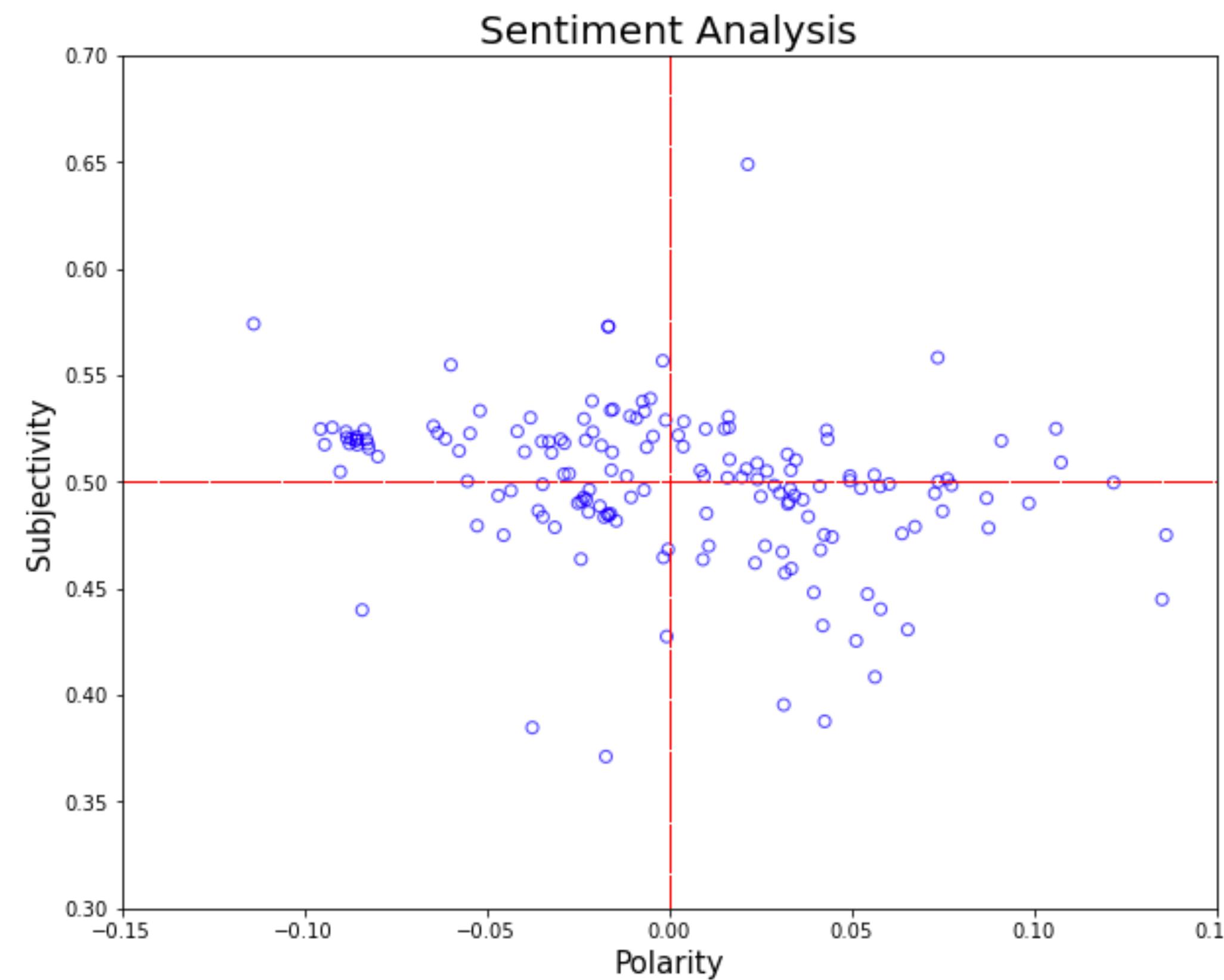
Right Trolls



SENTIMENT ANALYSIS OF RIGHT TROLLS



- Right Troll tweets are slightly more opinionated and slightly more negative.



TOPIC MODELING OF RIGHT TROLLS



— Trump vs. Obama Politics

trump
breaking
video
new
obama
house
president
donald
donald trump
liberal

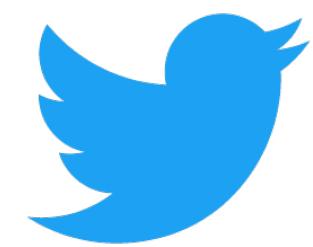


— Right Wing Militia

enlist
patriot
army
join
stand
america
freedom
people
read
truth



TOPIC MODELING OF RIGHT TROLLS



— Russia/Mueller Investigation

trump

congress

mueller

american

democrat

russia

obama

medium

fbi

gop



— Trump vs. Obama Media

trump

potus

realdonaldtrump

foxnews

president

people

liberal

obama

look

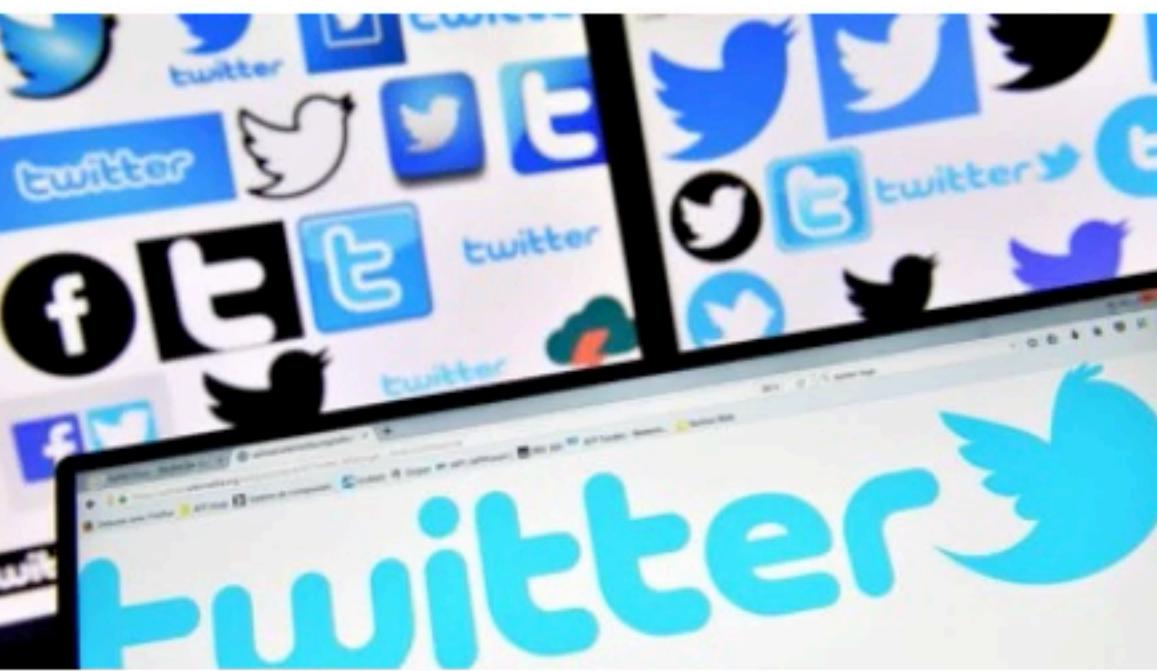
like



RANDOM FOREST TWEET CLASSIFIER



Right Troll Tweet Classifier

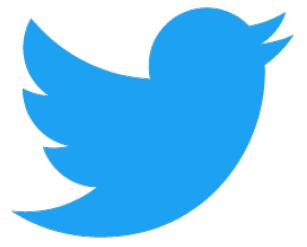


Enter Tweet:

put text here

Not Right Troll

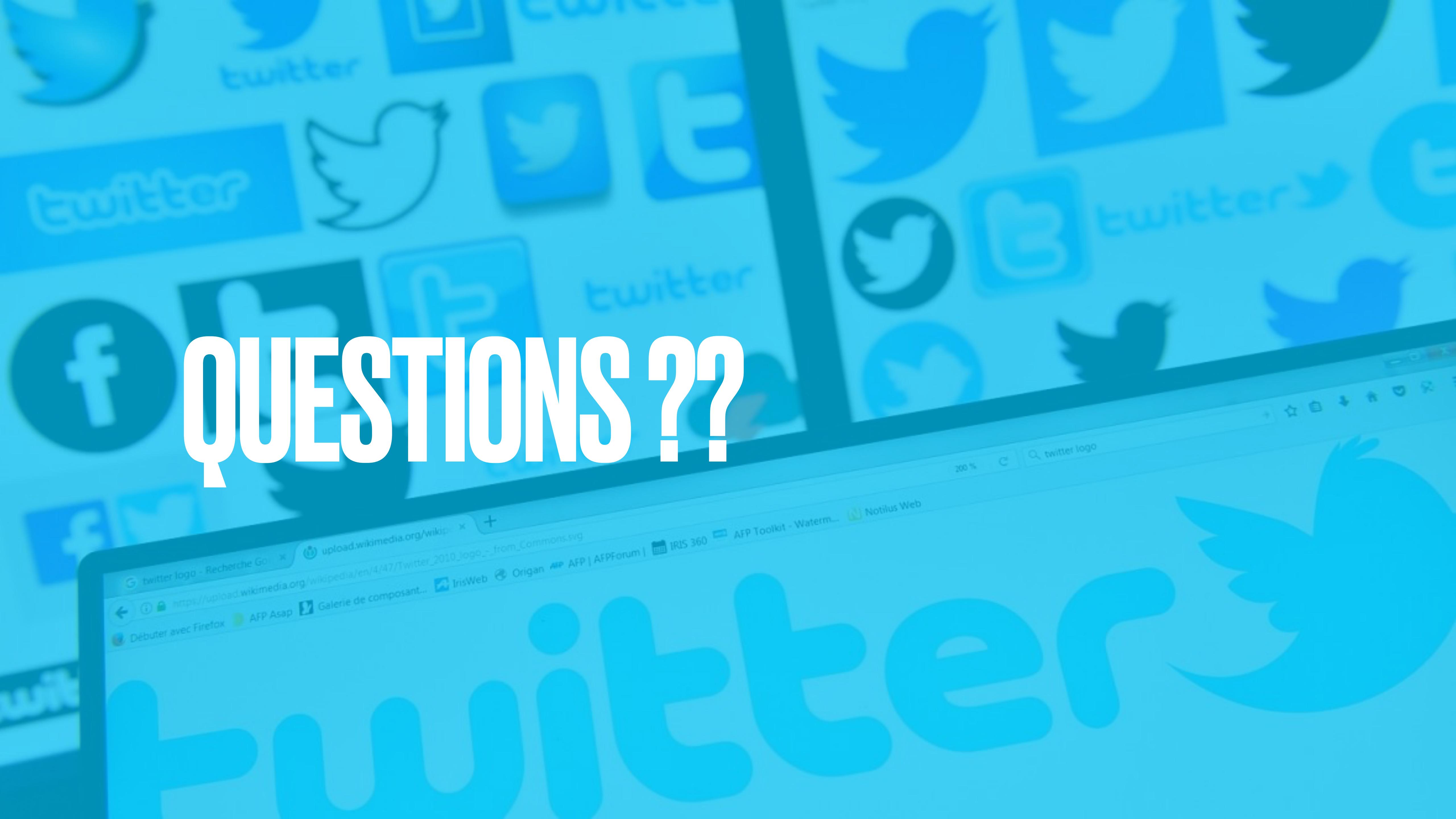
CONCLUSIONS



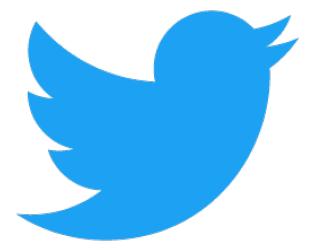
- 4 Topics does a good job of modeling the Right Troll tweets.
 - Trump vs. Obama Politics
 - Right-Wing Militia
 - Russia/Mueller Investigation
 - Trump vs. Obama Media
- Random Forest model is able to classify tweets as Right Trolls or not.



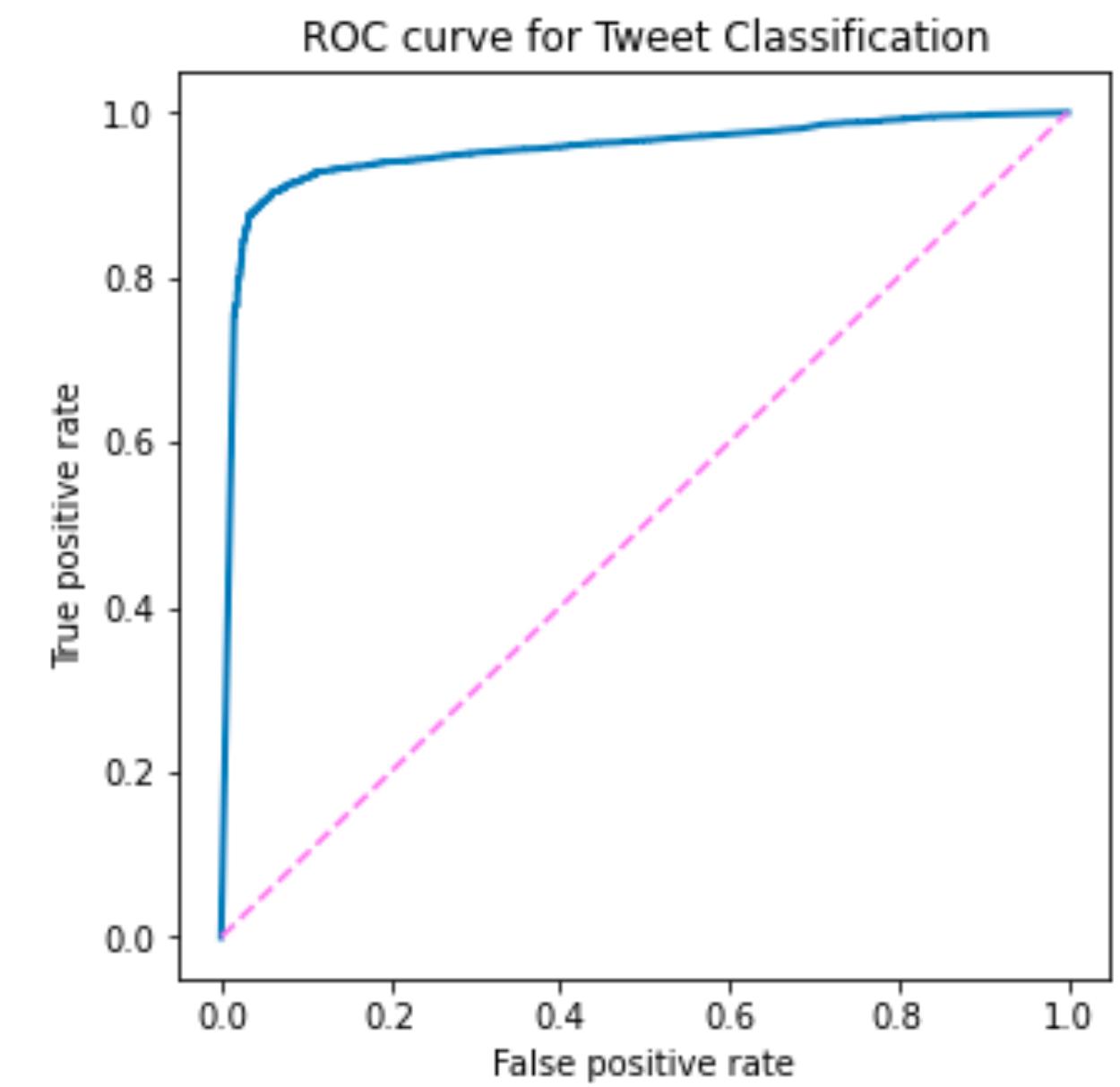
QUESTIONS ??



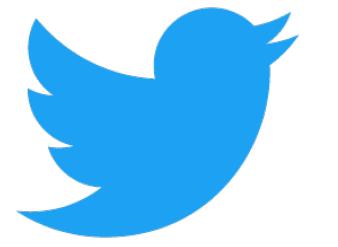
APPENDIX



- Random Forest Classifier ROC AUC Score = 0.95



FUTURE WORK



- **Random Forest Classifier with Left and Right Trolls.**
- **K-Means Clustering on all labeled data to classify unlabeled data.**