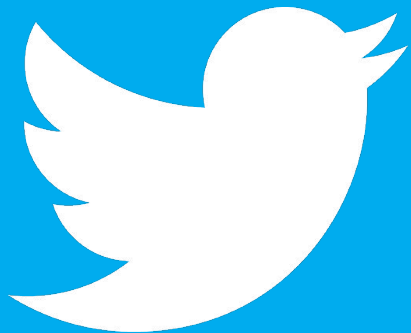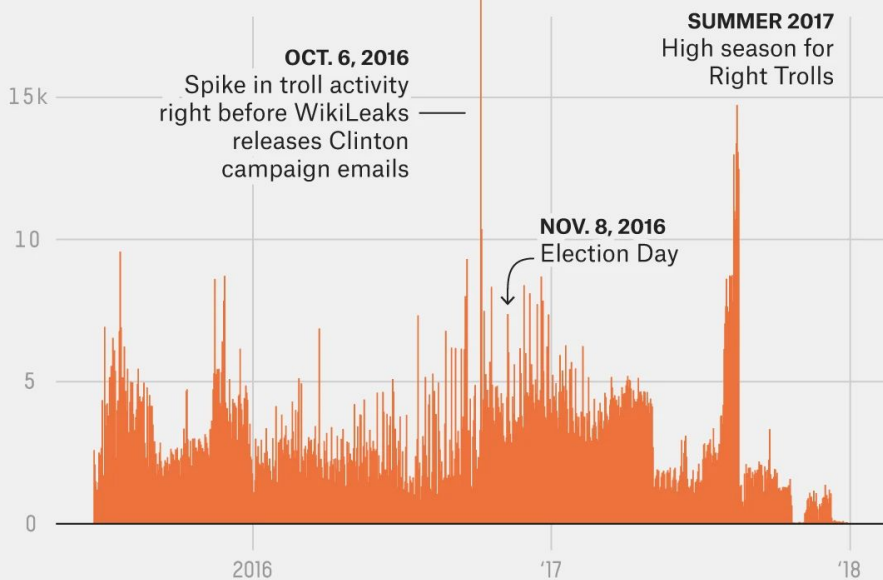# Bot Removal on Twitter

Training Lead By: Jake Moon
In Collaboration With FiveThirtyEight

FiveThirtyEight Uncovers Bot Data

Russian troll tweets by day

Nearly 3 million tweets sent by trolls associated with the Internet Research Agency

OCT. 6, 2016
Spike in troll activity right before WikiLeaks releases Clinton campaign emails

SUMMER 2017
High season for Right Trolls

NOV. 8, 2016
Election Day

15k

10

5

0

2016   '17   '18

FiveThirtyEight   SOURCE: DARREN LINVILL AND PATRICK WARREN

- FiveThirtyEight released 3 million tweets associated with the Internet Research Agency (IRA) in July 2018
- IRA deployed automated "troll" accounts that spread misinformation on Twitter
- Goal to misinform and divide the American populace

70 million accounts removed

But there are more accounts to be flagged and removed

Positive Cases

@covfefenationus · 7 Nov 2017

↻ Retweet

HYPOCRISY ALERT! Donna Brazile Reveals Top Hillary Aides Were
Sexist https://t.co/FnaJiCVEmZ https://t.co/I75EUH3qxg

🇺🇸 United States / English

+ more details

@4mysquad · 11 Aug 2015

An american cop as he is. #StayWoke #Cops #Police #Racism
#BlackLivesMatter http://t.co/K1YXyjYTEN http://t.co/WJjldI2DEA
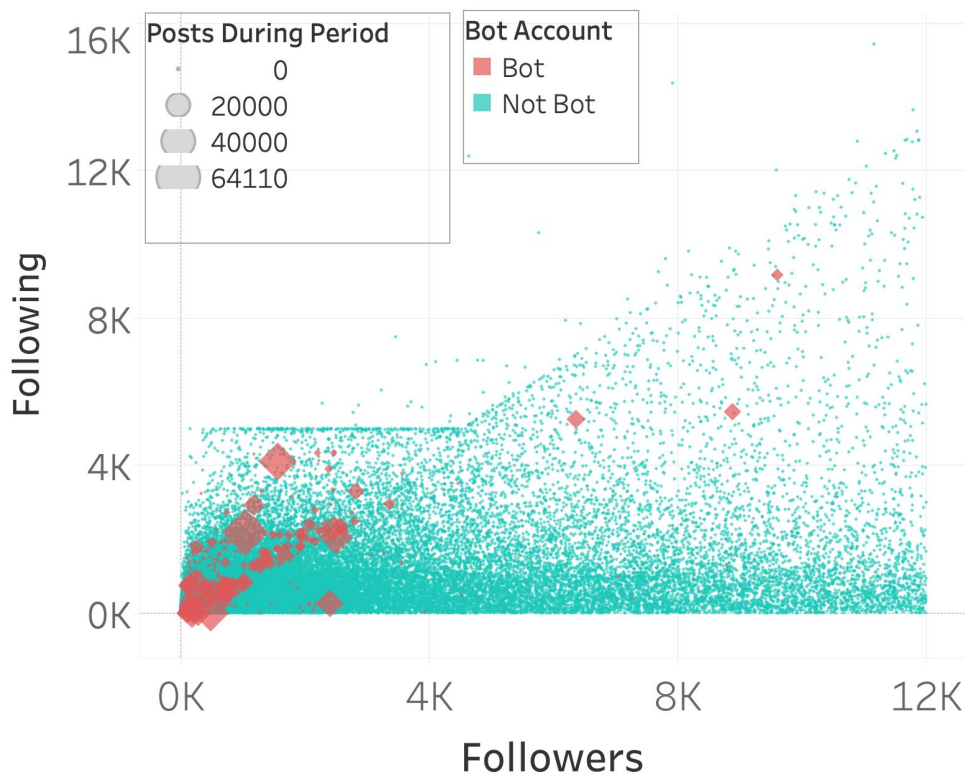
🇺🇸 United States / English

+ more details

- Left and right-wing content
- Ability to retweet and quote other tweets
- Link articles
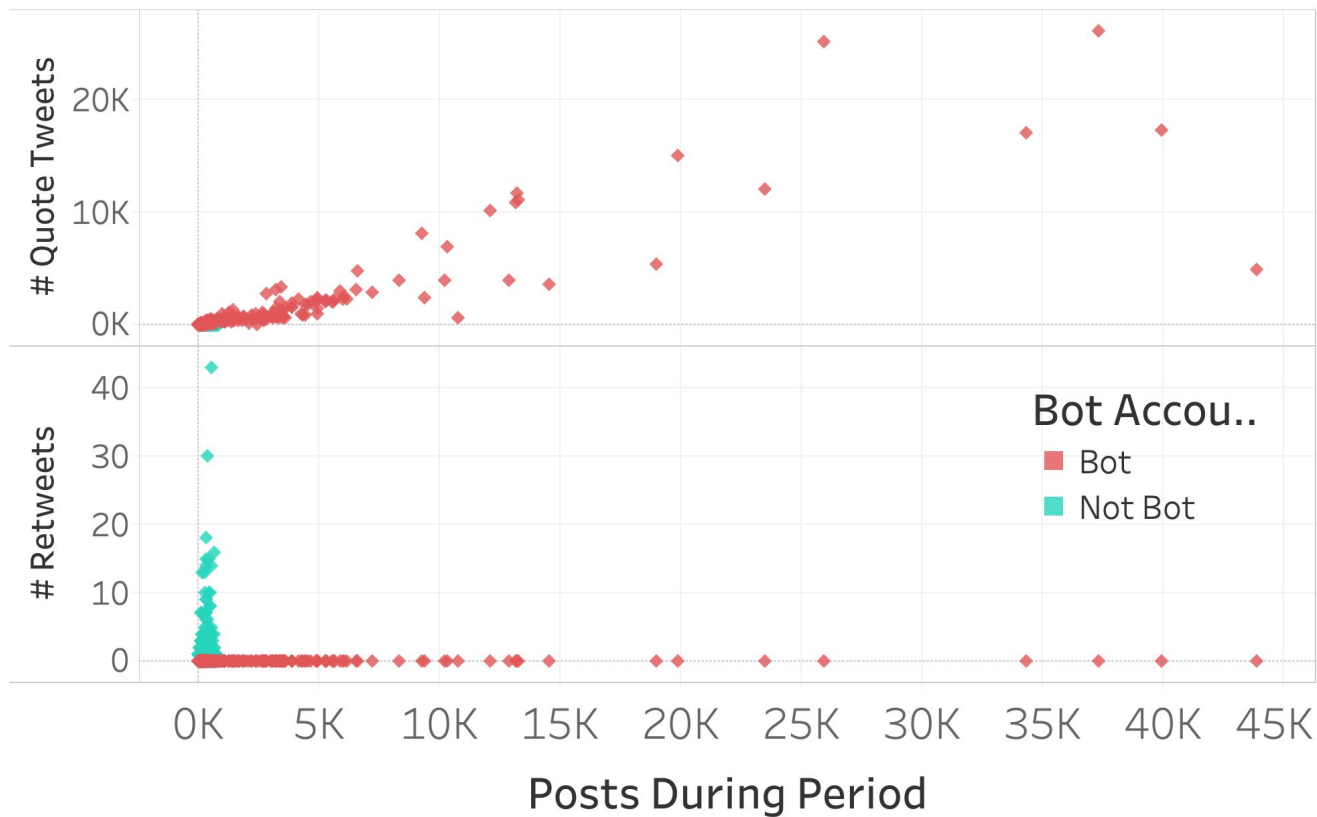- Hashtag
- Need to discern many quickly

Bot Behavior

Following/Follower Comparison

- High # of posts

- High following/follower ratio

  - Steep curve

  - Unnatural # of follows in short amount of time

Separable Features

Account Level Features

Followers/Following Counts
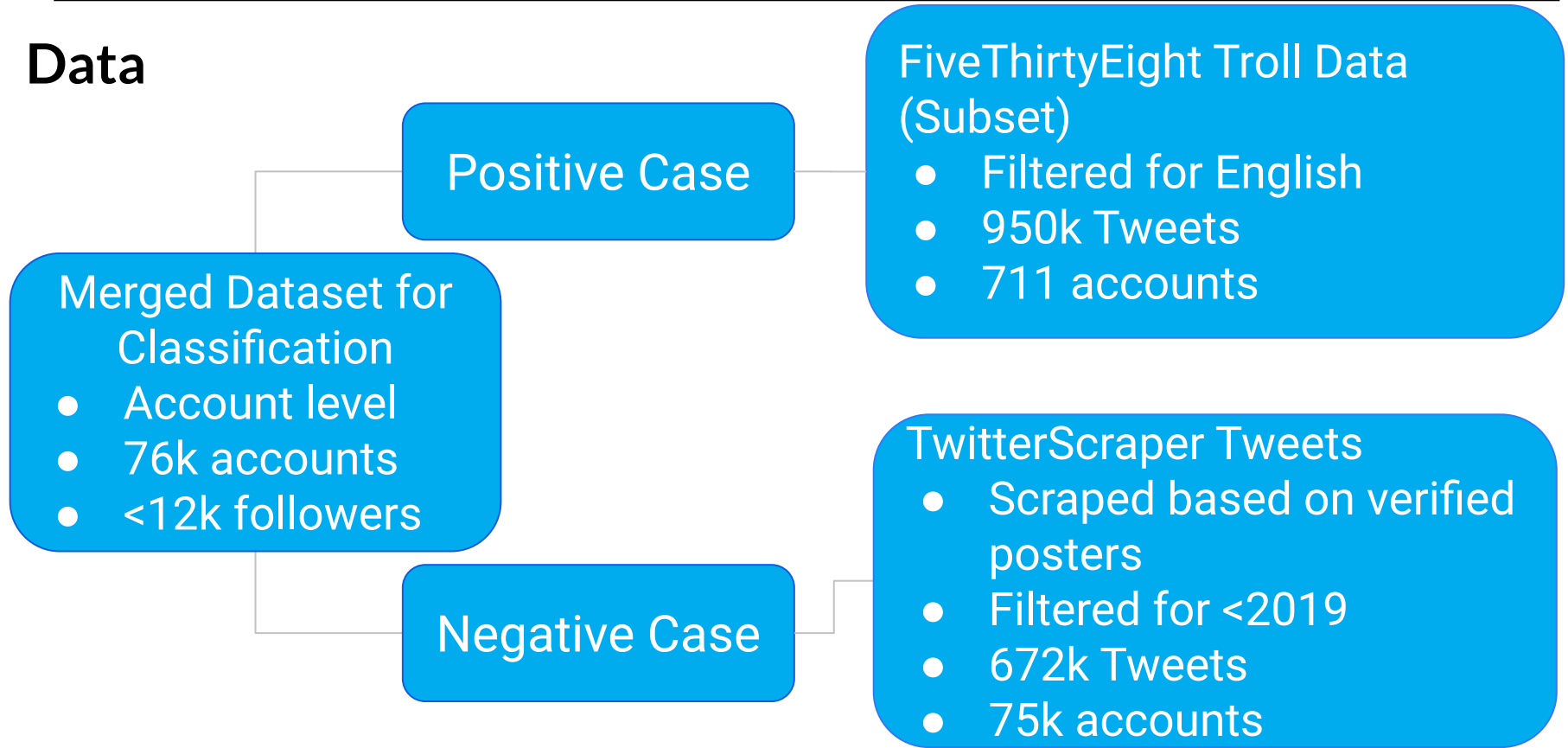
Total Posts Counts (During Period)

Retweet Counts

Quote Counts

Character Sum and Aver Char. Per Tweet

Hashtag Counts

Uppercase Characters Counts

Data

Merged Dataset for Classification
- Account level
- 76k accounts
- <12k followers

Positive Case

FiveThirtyEight Troll Data (Subset)
- Filtered for English
- 950k Tweets
- 711 accounts

Negative Case

TwitterScraper Tweets
- Scraped based on verified posters
- Filtered for <2019
- 672k Tweets
- 75k accounts

Models

F1:71%

Logistic Model

L2 Penalty

- Recall: 59%
- Precis: 76%

F1:83%

Random Forest

5 Trees
Criterion: Gini
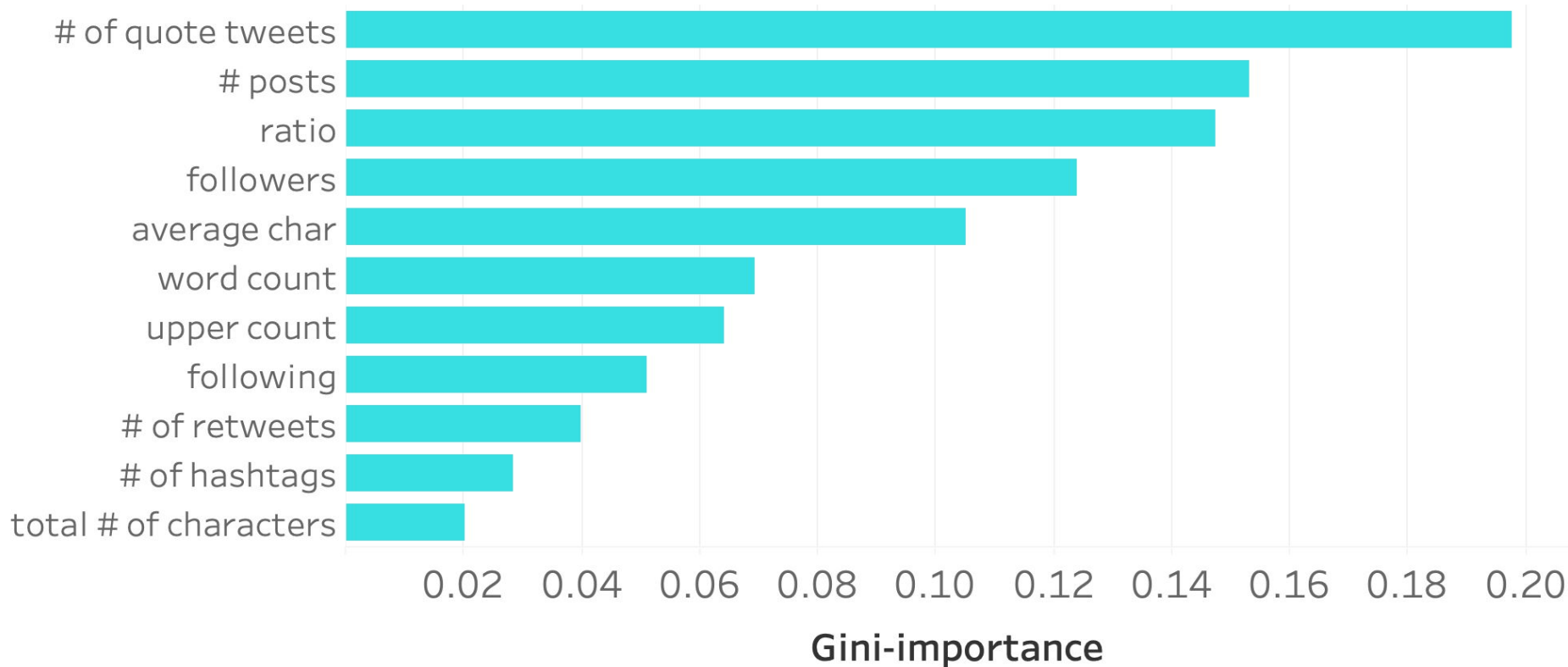
- Recall: 76%
- Precis: 91%

F1:82%

XGBoost

Max Depth: 3
Learning Rate: 0.011
1000 Trees

- Recall: 74%
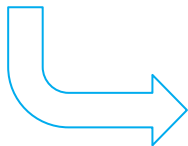- Precis: 93%

Feature Importances

Incorporating Text Data and NLP

Adding cleaned text information
- CountVectorizer
- Minimum #: 10
- Max %: 30%

F1:94%

Random Forest

5 Trees
Criterion: Gini

- Recall: 92%
- Acc: 99.8%

F1:100%

XGBoost

Max Depth: 3
Learning Rate: 0.011
1000 Trees

- Recall: 100%
- Acc: 100%

Final Classification Model and Results

- Ignore NLP features

- XGBoost Model

- Ban based on account activity as a whole

- Flag when certain behaviors detected and monitor

Acc:58%

100% Bot-Only Validation Set

Missed Bots Example

@allahomore · 7 Dec 2015

↻ Retweet

Art stolen from Frisian Museum in 2005 foun in hands of Ukraine's top extreme right politicians. Thanks @JefHorians https://t.co/pdYoVQGQxG

Unknown Region / English

created by the **Internet Research Agency**, a "troll factory" government. It was part of a campaign to sway politics and United States. For more information, read our **FAQ**

- Only 2 English posts in dataset

- Reasonable # of followers and following

TWEETS ❓	See all 465 tweets by @allahomore
HASHTAGS ❓	See 35 hashtags used by @allahomore
FOLLOWERS ❓	198
MAX FOLLOWERS ❓	200
FOLLOWING ❓	94
RETWEETS ❓	462 (99%)
CATEGORY	NonEnglish
TYPE	Russian
TIME ACTIVE	Jun 20, 2015 - Dec 28, 2015 (6 months)
LANGUAGES ❓	Bulgarian, English, Macedonian, Russian, Serbian, Ukrainian, Uzbek
REGIONS ❓	Unknown

Improvements

- Scrape more political tweets for negative cases
  - Political terms and spam words dominated NLP
  - Shift predictive power to numeric features
  - More robust model that can generalize
- Verify and add other, more recent bot data for positive case
  - Iranian bot data
- More feature engineering to capture relationships

Questions?

"[...] a botnet promoting anti-vaccine content alongside the #DemDebates hashtag. Some 46 percent of the tweets pushing anti-vaccine disinformation came from bot accounts; 19 percent of these users average more than 100 tweets per day."

- Washington Post, November 20, 2019

Appendix

https://fivethirtyeight.com/features/why-were-sharing-3-million-russian-troll-tweets/

https://github.com/fivethirtyeight/russian-troll-tweets/

https://www.washingtonpost.com/technology/2018/07/06/twitter-is-sweeping-out-fake-accounts-like-never-before-putting-user-growth-risk/

https://github.com/taspinar/twitterscraper

https://www.washingtonpost.com/politics/2019/11/20/are-bots-manipulating-conversation-heres-whats-changed-since/

FiveThirtyEight