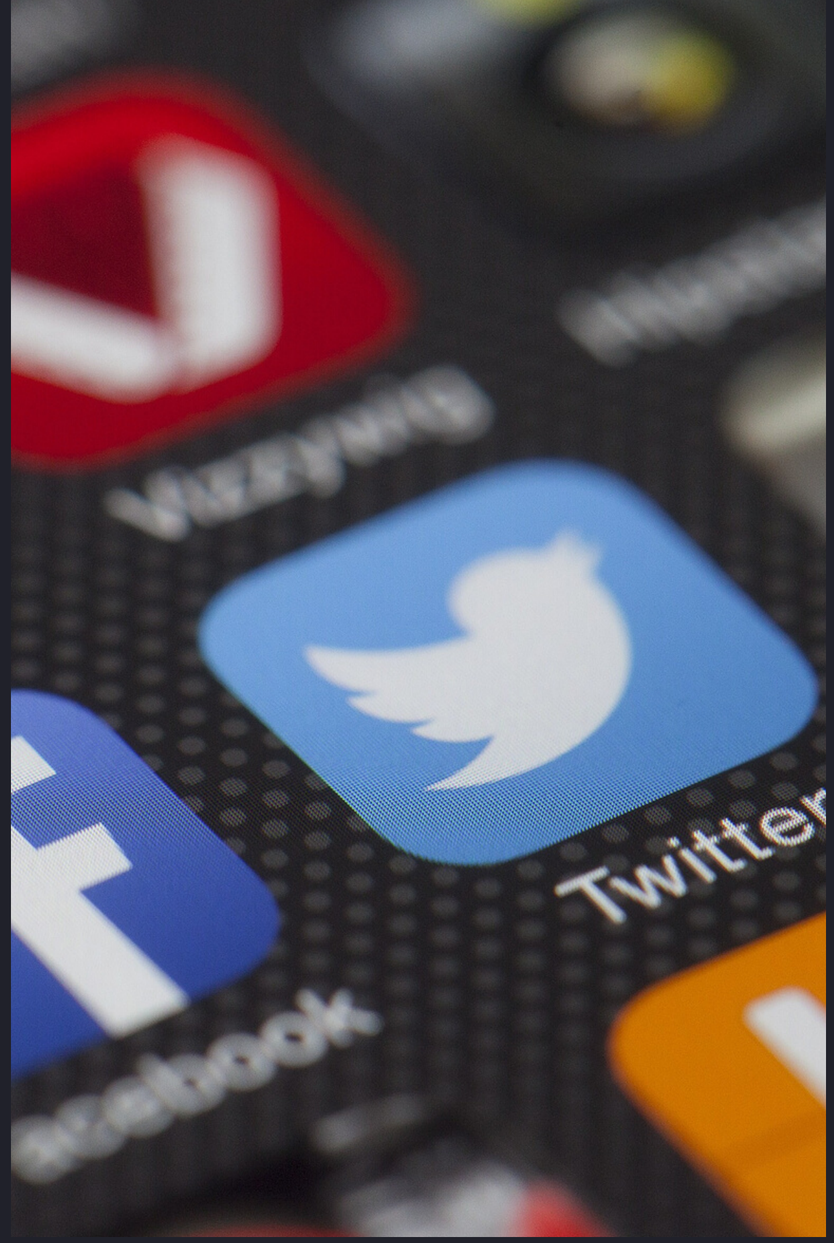
JANUARY 11, 2020

TWITTER BOT DETECTION

Sam Videlock



Agenda

Background
Project Scope
The Data & Model
Next Steps

What is a Bot?

- A bot is a piece of software that completes automated tasks over the internet
- **Good Bots** disclose their identity to the web servers they access while **Bad Bots** do not

1/2

OF ALL INTERNET TRAFFIC IS PERFORMED BY BOTS

56%

OF BOTS ARE CATEGORIZED AS BAD BOTS

~15%

OF TWITTER ACCOUNTS ARE ESTIMATED TO BE BOTS

Why Should We Care?

Studies show in the months leading up to the 2016 election 1/5 of all tweets that were election related came from a legion of bot accounts

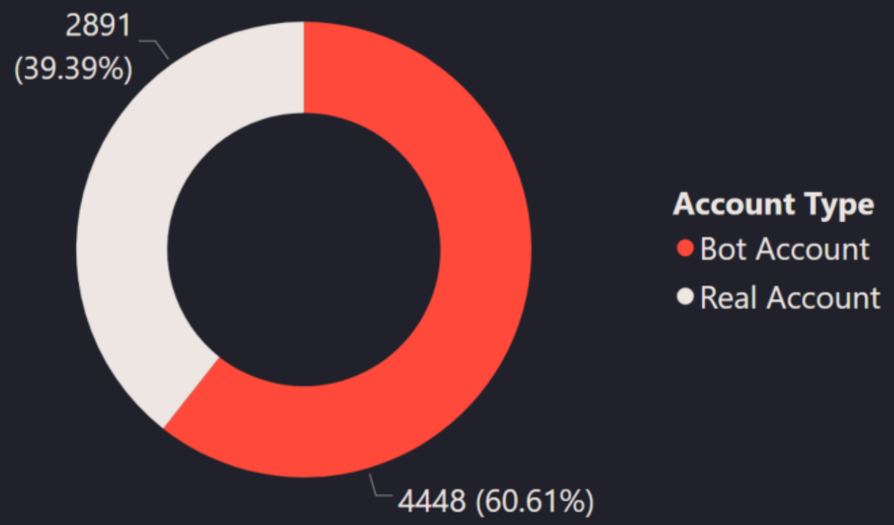
When working together in large cluster, bots have the ability to push narratives that could be false of misleading

DARPA (The Defense Advanced Research Projects Agency) wants to identify and remove bots whose goal aims to influence others and protect information exchange on the internt.

Project Scope

Identify social spam-bots on Twitter whose goal is to influence users using specific tweets and user data

The Data



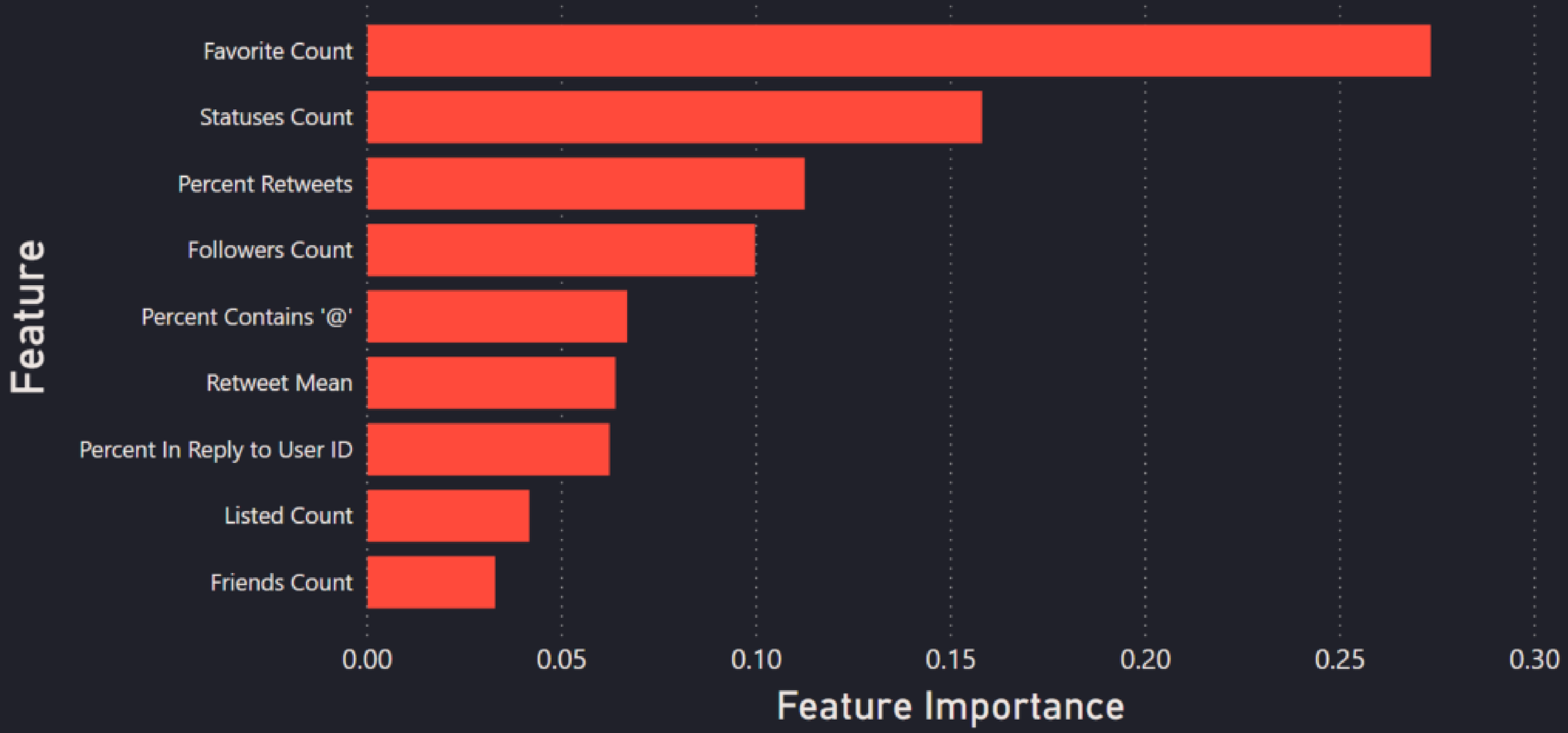
Number of Bot Accounts vs. Real Accounts in the Data Set

The gathered data comes from the Bot Repository and the Twitter API

Do to difficulties with finding public labeled data the project serves as a proof of concept. More data will be needed to create accurate models for more scenarios

5 MILLION
TWEETS
FROM
7300
USERS

Most Important Tweet Features



A combination of user data and actual tweet data performs best

The Model

identifies Social Spambots

97%

of the time compared with real human accounts

NEXT STEPS

MORE DATA

In order to ensure confidence in identifying social spambots, more data gathering is necessary to be representative of the larger Twitter world

CLUSTERING

Clustering users should be helpful to identify bots as they become more advanced in the future

DEEP LEARNING

Deep learning sentiment analysis on a users tweets could be a useful feature for these models across many scenarios

REAL USERS

Another large issue is outside influence in elections from real users

Couple this model with new features and data to identify real users who have ill intentions

QUESTIONS?