Toxic Tweets

(The group formerly
known as BotsBotsBots)

Rose Gao, Andrea Koch, Ian Lyons

GOALS

Original BotsBotsBots Proposal:

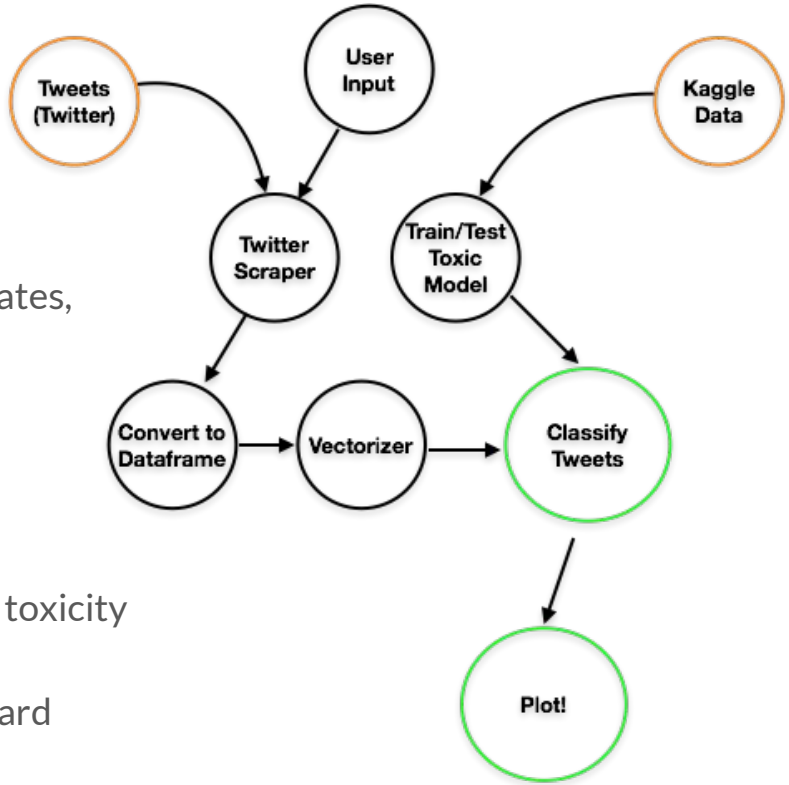
- Create a social bot on Twitter which would detect fake news bots
- Observe relationship (if any) between bot activity and political events

Toxic Tweets Project:

- Analyze several types of toxicity in tweets, focusing on the 2016 US election cycle
- Visualize trends in toxicity levels, major political topics, and the political timeline

APPROACH

1. Twitter Scraper:
    - User provides query (query term, search dates, tweet limit per day, output filename)
    - Query tweets and export as a csv
2. Classifier:
    - Read in the csv as a pandas dataframe
    - Vectorize the text in each tweet
    - Use the 6 toxic models to identify levels of toxicity
3. Visualization:
    - Create graphs and pull them into a dashboard
    - Visualize the most compelling insights



INSTRUMENTS

- NLP/Classification Libraries Used:
  - nltk:
    - Lemmatizer
    - Tokenizer
  - sklearn:
    - TF-IDF Vectorizer
    - Logistic Regression
    - Cross Validation Score
  - scipy:
    - hstack

- Other Libraries Used:
  - TwitterScraper
  - Datetime
  - Plotly (plotting and dashboarding)

- With more time, want to use:
  - Twitter API developer access
  - Networkx: to map Twitter accounts, tweets, and hashtags

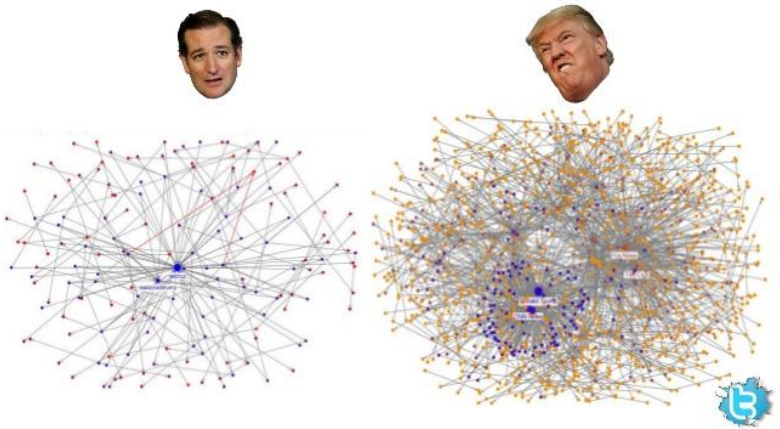
DATA

- Kaggle Toxic Comments Dataset
  - Wikipedia comments rated by human raters for several types of toxic behavior
  - 6 categories: Toxic, Severe Toxic, Obscene, Threat, Insult, Identity Hate
  - Training data: ~500,000 comments
- Tweets: Queried 5 topics from Jan. 1, 2016 - Feb. 28, 2018:
  - Hillary Clinton
  - Donald Trump
  - Gun control
  - Wikileaks
  - Abortion (abortion/pro-choice/pro-life)
- Vectorized tweet text and classified to get toxicity levels per toxicity category

LESSONS LEARNED

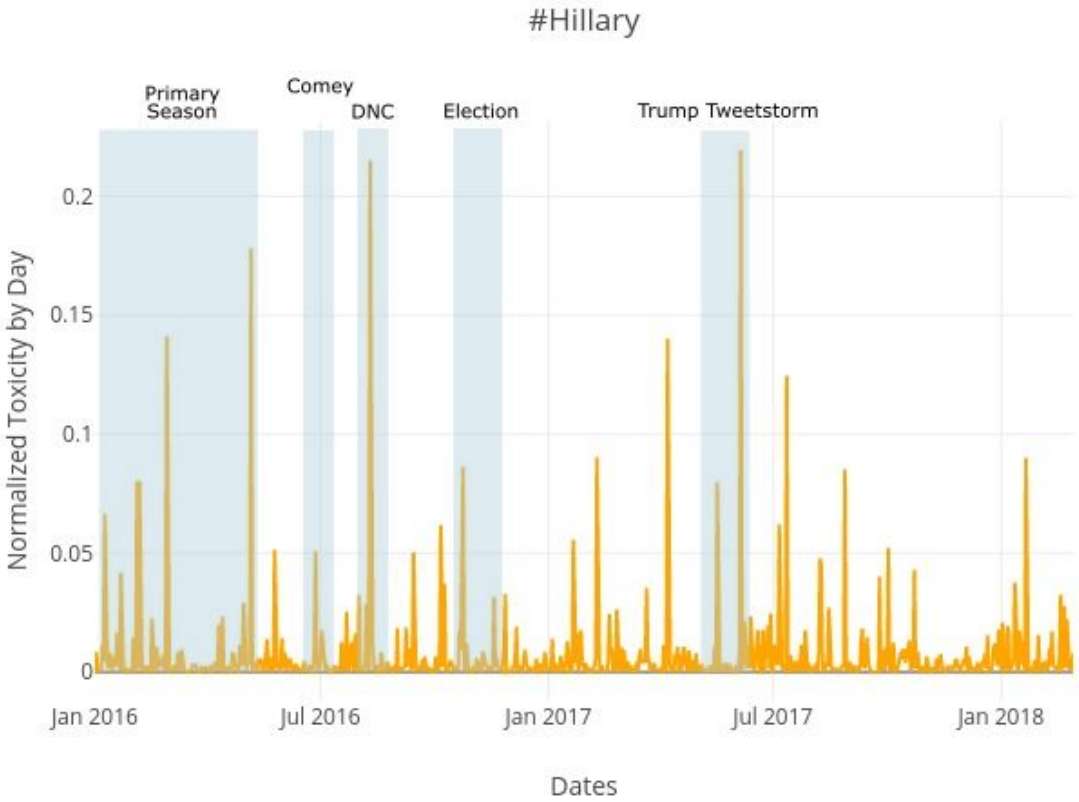
- Bot Detection: need to start with identified accounts, or identify accounts based on behavioral pattern, and expand on the network
- Fake News: need to use methods (e.g. neural nets) that can understand and ingest the *syntax* of words, not just individual words
- Syntax is **key** for the uses of NLP
- TwitterScraper is not your best friend, but it is the only one you have

Example: Twitter User Network

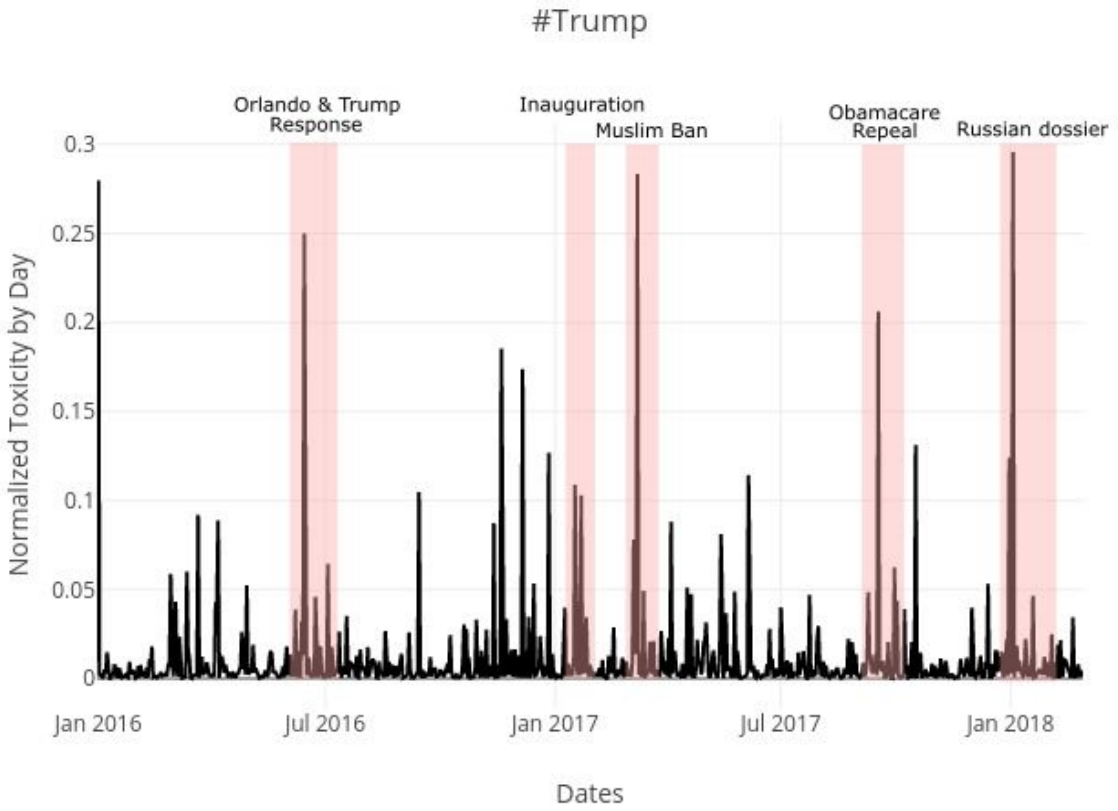


DEMO

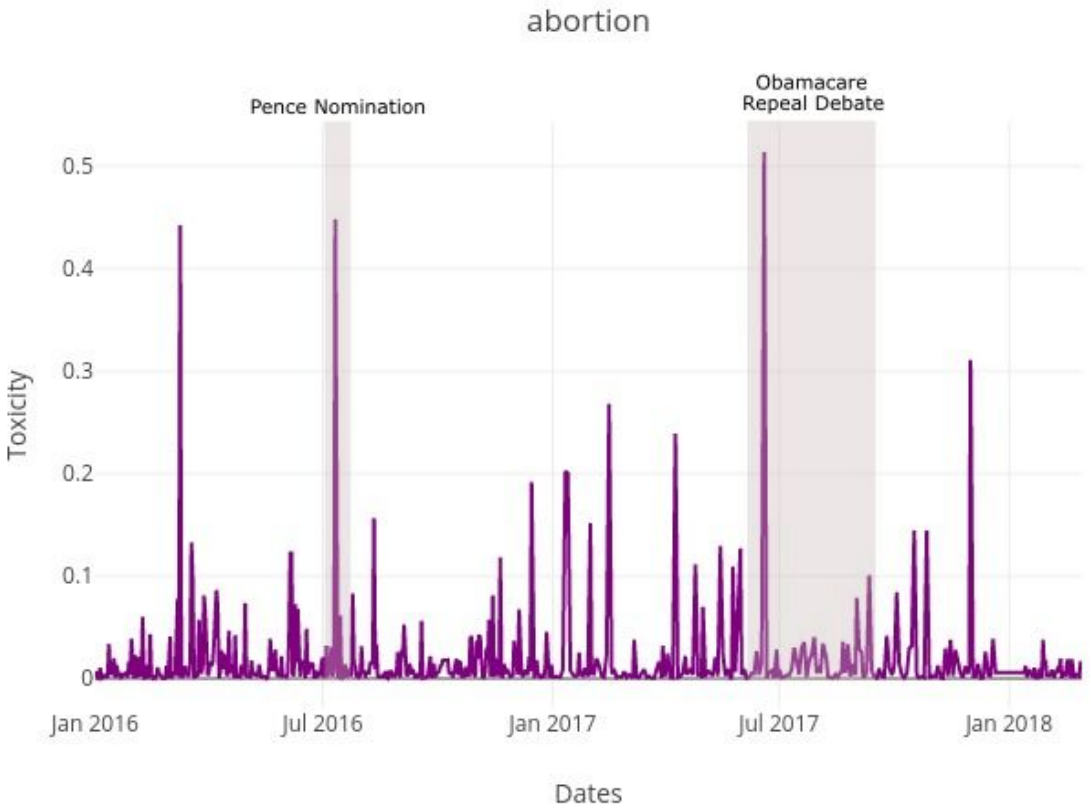
INSIGHTS: #Hillary



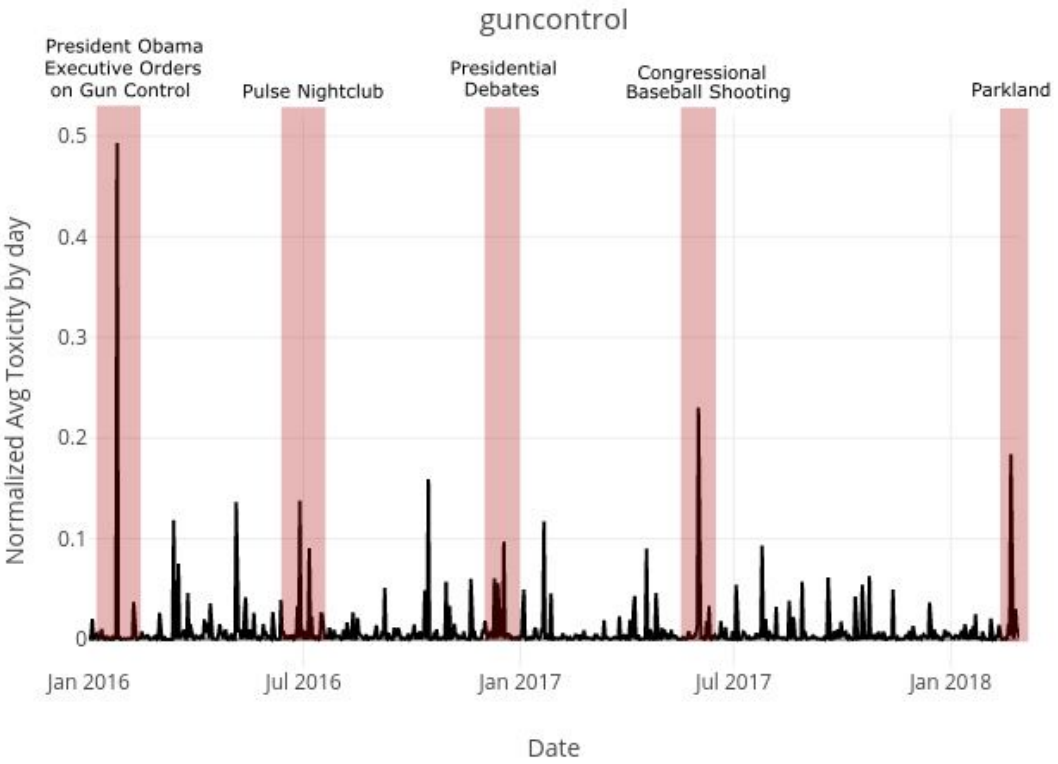
INSIGHTS: #Trump



INSIGHTS:
#abortion
#prochoice
#prolife



INSIGHTS: #guncontrol



INSIGHTS: #wikileaks