Politician Twitter Sentiment Analysis

Jeremy Chan[1], Gabriel Fajardo[1], Kenny Wang[1], Jason Wu[1]
[1]Boston College, Department of Computer Science

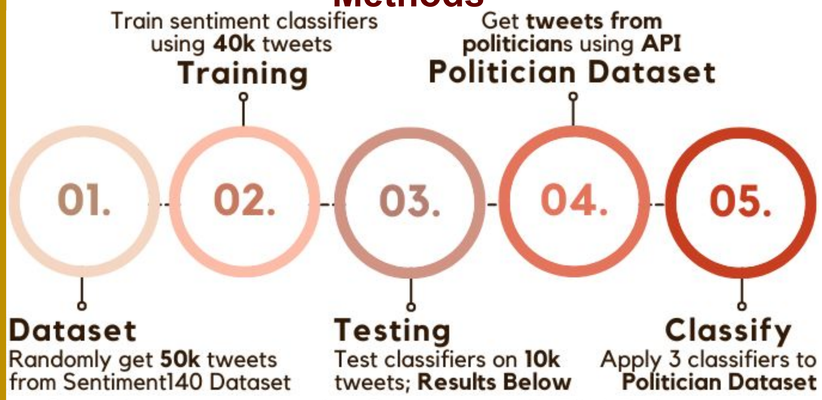
Introduction

- Twitter = “go-to” place to follow influential and affluent people including **politicians**
- Goal of this project: to gauge Twitter **sentiment** towards **Democrats** vs. **Republicans**
- Secondary goal: to use Twitter sentiment to analyze **sentiment after the election** results

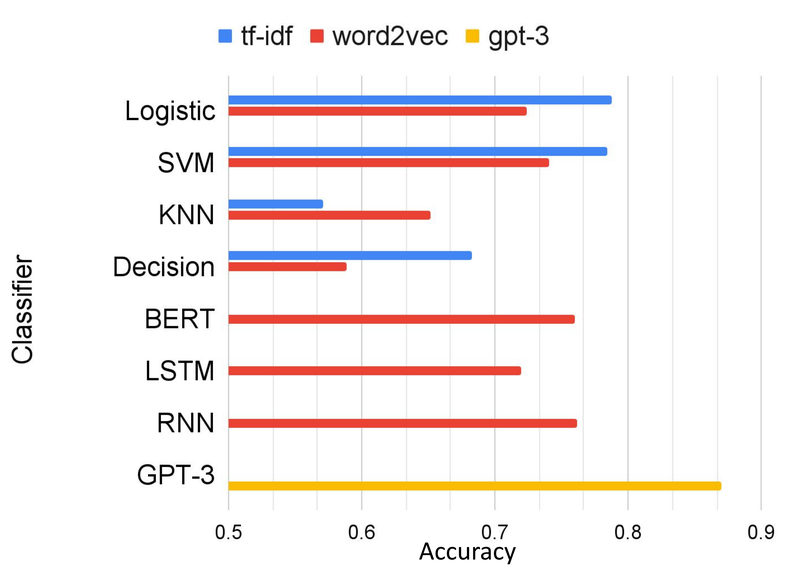
Data

- Randomly sampled **25k positive** and **25k negative** tweets from **Sentiment140** Dataset (1.6 million tweets total)
- Processed text using **tf-idf or word2vec**

Methods



Classifier Accuracies for Sentiment140 (50k) Data

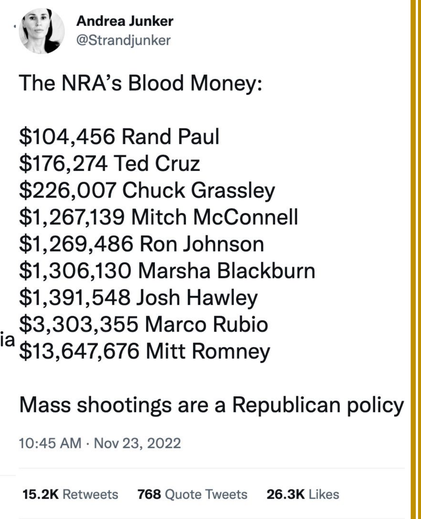


Politician Sentiment

- Chose **popular Democrat vs. Republican** opponents from 2022 US Senate election
- Extracted ~300 tweets about each politician using **Twitter Search API**; all tweets are from **November 23, 2022**

State	Republican	Democrat
PA	Mehmet Oz	John Fetterman
NV	Adam Laxalt	Catherine Cortez Masto
WI	Ron Johnson	Mandela Barnes
NH	Donald Bolduc	Maggie Hassan
NC	Ted Budd	Cheri Beasley
NY	Joe Pinion	Charles Schumer
OH	JD Vance	Tim Ryan

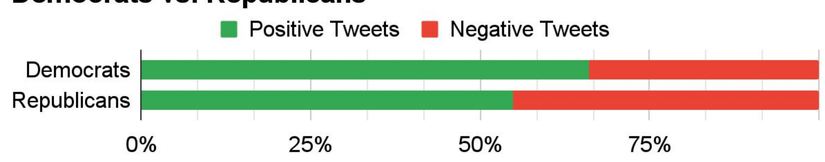
Green = won election Red = lost election



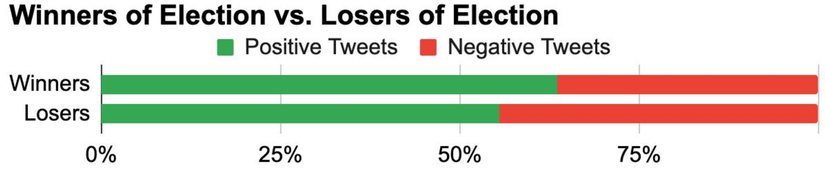
Results

Used average of LSTM, BERT, LogReg, and GPT-3 to predict politician tweets

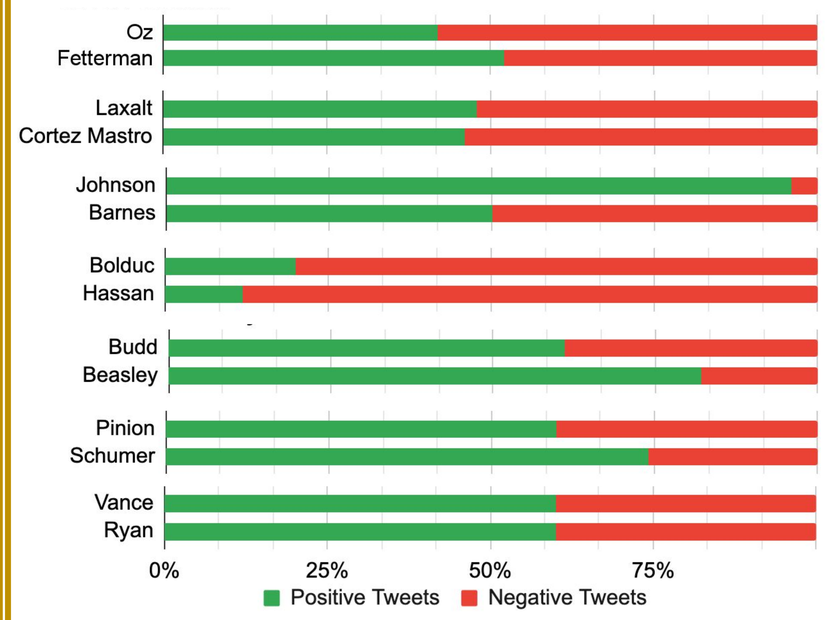
Democrats vs. Republicans



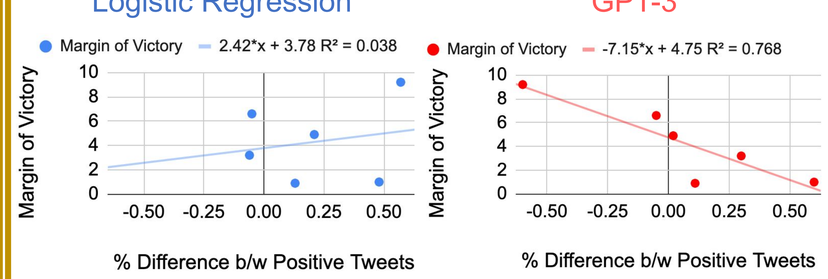
Winners of Election vs. Losers of Election



Positive vs. Negative Twitter Sentiment for Each Political Race



Margin of Victory vs. Percent Difference in Positive Tweets Using Different Classifiers



Conclusions

- **Findings**:
  - 1) **Democrats** had **more** positive tweets than **Republicans**
  - 2) **Winners** had **more** positive tweets than the **Losers**
  - 3) **Negative correlation** between **Positivity** and **Margin of Victory** for GPT-3 model; no correlation for LogReg
- **Limitations**:
  - 1) Twitter users do **not** make up a **representative sample**
  - 2) **Twitter** *expressions of emotion* are far **different** from **interpersonal** *expressions of emotion*
  - 3) **Full API**: see how sentiment changes over time, get Tweets from **November 8th**
  - 4) Train more **complex Neural Network** classifiers
- **Future**: dataset of political tweets, non-binary classification, and factor in the number of retweets/likes