

Project 2 (Research) – Sentiment Classification

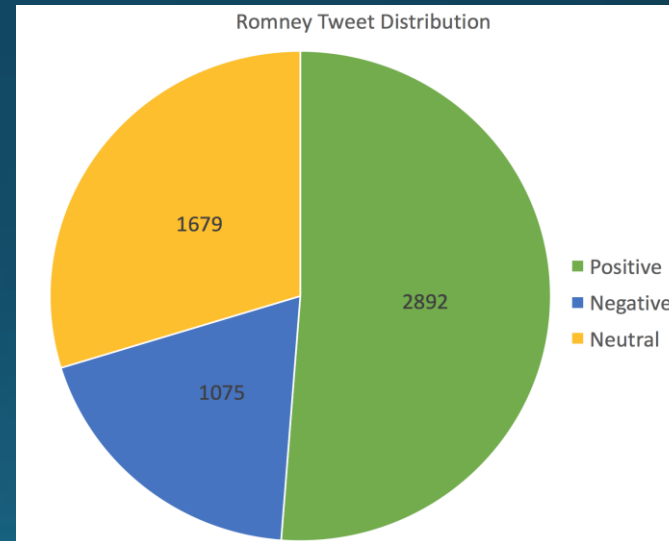
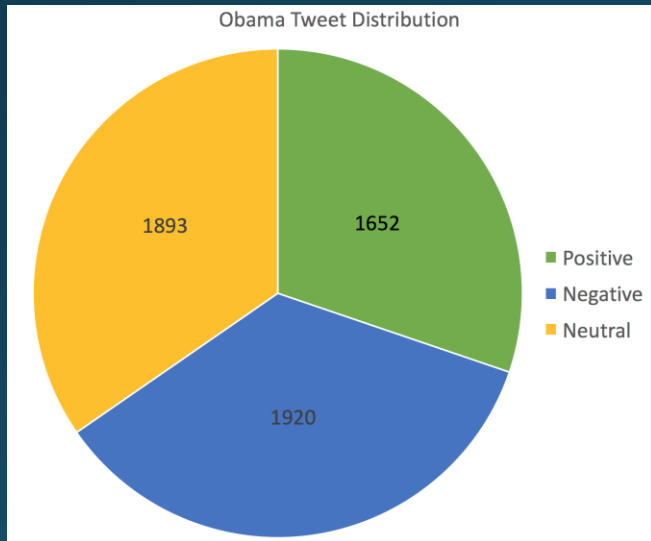
Team Members:

Sreejith Menon (smenon8@uic.edu)

Rajan Bhandari (rbhand4@uic.edu)

Data Preprocessing & Data Distribution

- Data Cleaning
 - Removed html tags, hyperlinks, symbols (all non-alphanumeric characters)
 - Converted all texts to lowercase, removed tweets from mixed class
- Data Distribution (post cleaning)

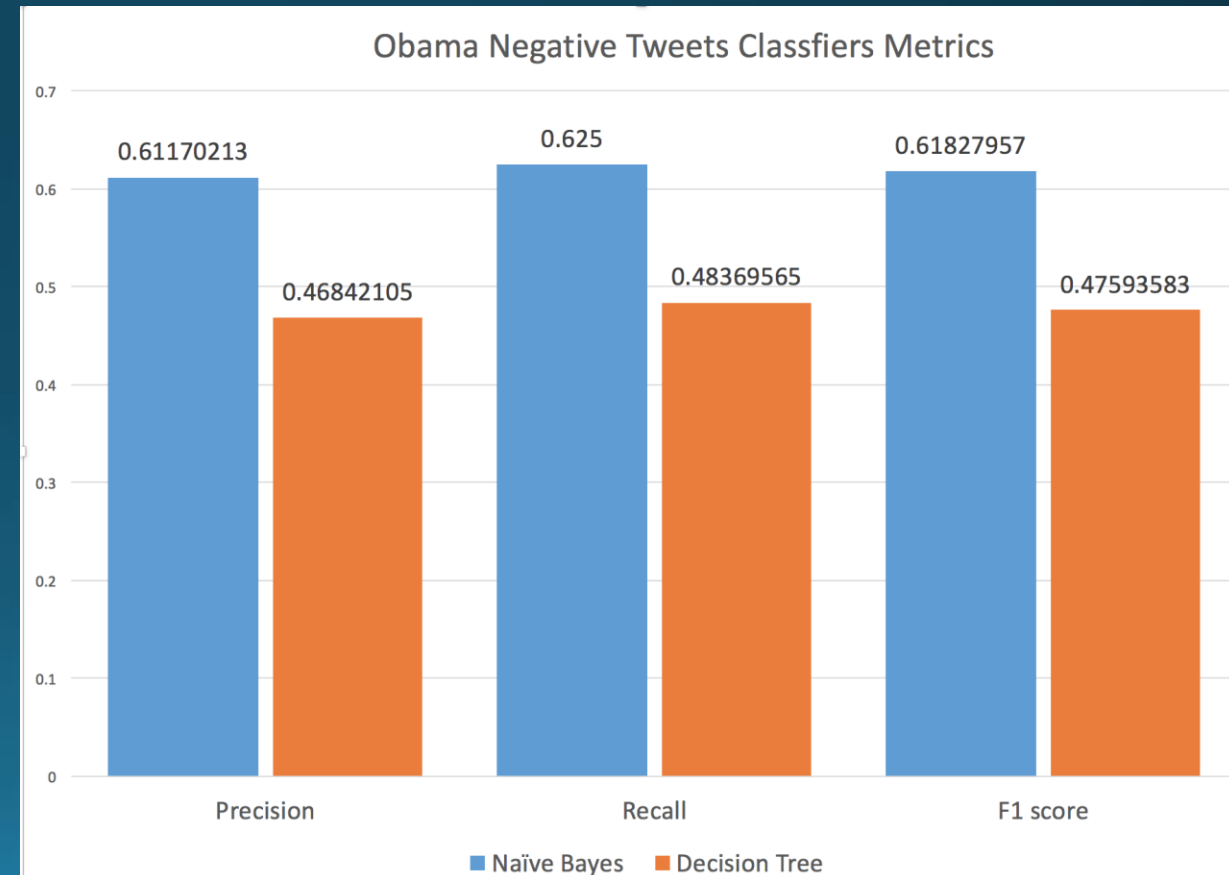
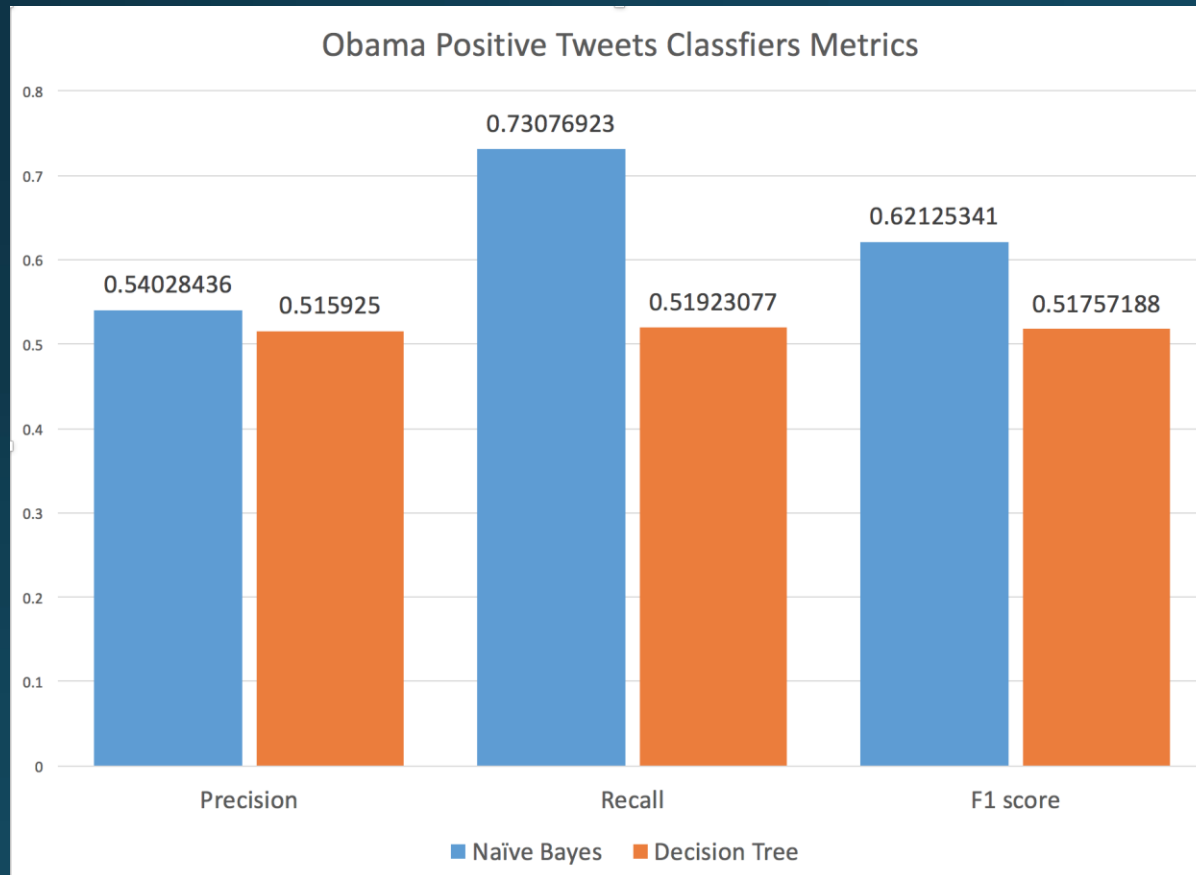


- Converted the tweets to bag of words model (binary sequence transformation for each tweet)

Classification Algorithms & Results - Obama

Naïve Bayes: Overall Accuracy = 0.581352833

Decision Tree: Overall Accuracy = 0.49360146



Classification Algorithms & Results - Romney

Naïve Bayes: Overall Accuracy = 0.569911504

Decision Tree: Overall Accuracy = 0.513274336

