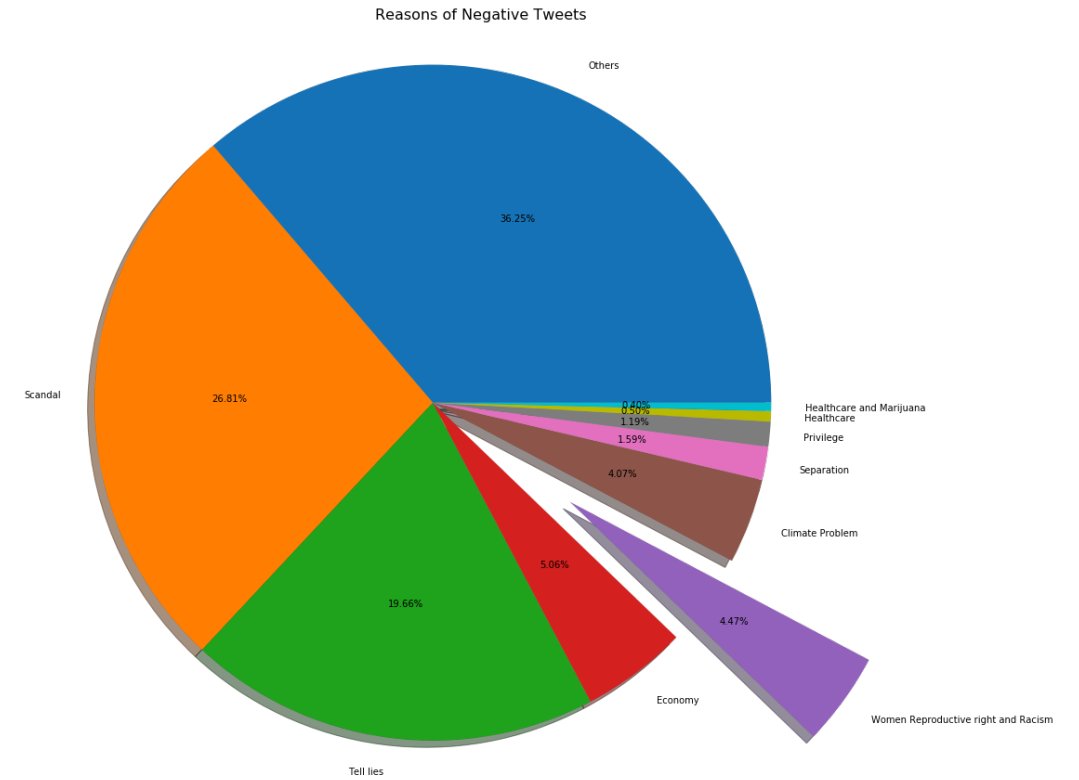
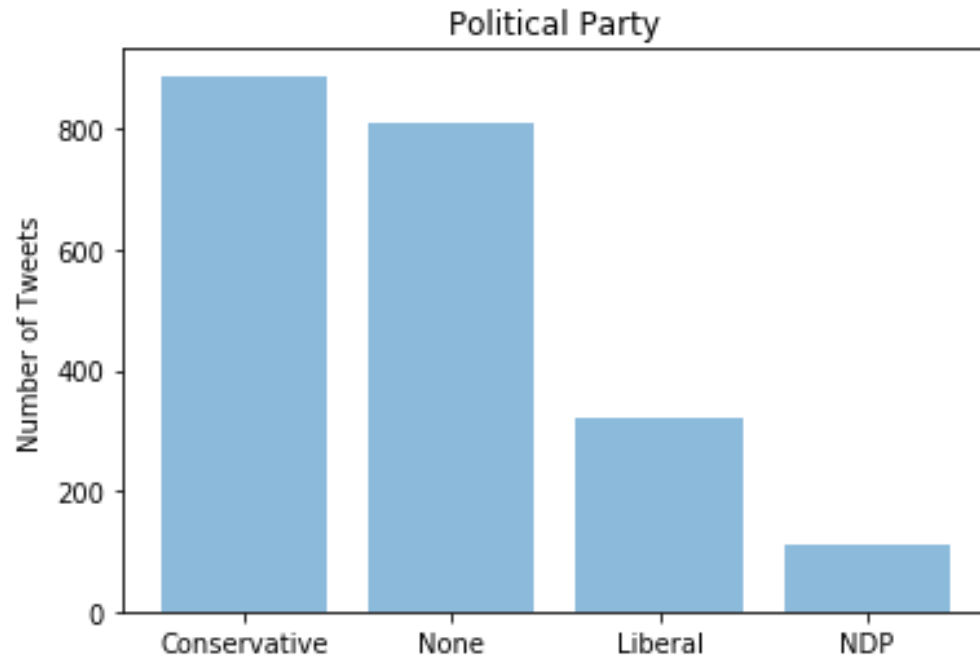


A word cloud background with various terms in different colors and sizes. The most prominent words include 'going', 'canadian', '1', 'e', 'n', 'p', 'o', 'i', 'government', 'reopen', 'gpc', 'daily', 'andrewscheer', 'let', 'cost', 'voting', 'real', 'scheer', 'debate', 'yup', 'right', 'canada', 'cityofto', 'justin', 'lpc', 'ask', 'next', 'protect', 'reminder', 'go', 'billion', 'body', 'brantford', 'callwomanspacewalk', 'trust', 'women', 'womanspacewalk', 'go', 'canadian', 'scheerhypocrisy', 'real', 'voting', 'cost', 'let', 'andrewscheer', 'daily', 'gover', 'reopen', 'gpc', 'suggestion', 'onpoli', 'instead', 'canada', 'right', 'yup', 'abortion', 'debate', 'yup', 'andrewscheer', 'daily', 'gover', 'reopen', 'gpc', 'suggestion', 'onpoli', 'instead'.

Sentiment Analysis of Canadian Elections 2019 Tweets

Tianyi Yu

Mar 26, 2020

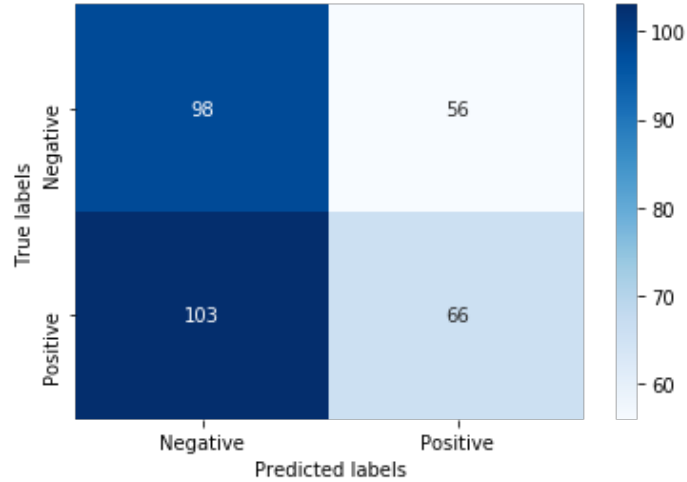


About the Canadian Elections 2019 Tweets

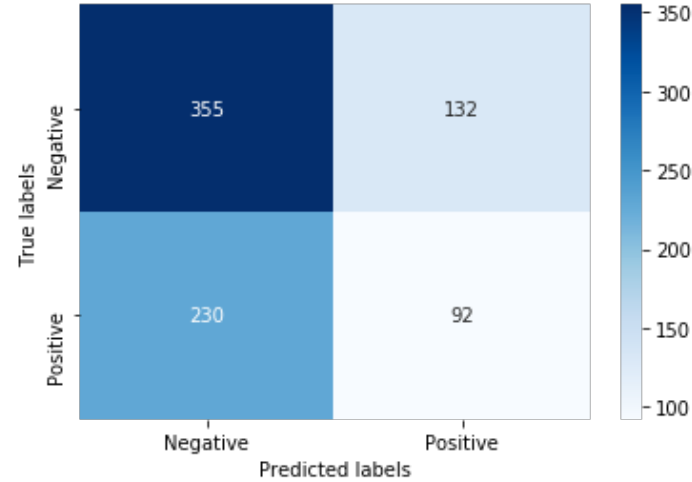
Model Selection

	Model	Word Frequency Features	TF-IDF Features
0	Logistic Regression	71.771	71.751
1	k-NN	66.371	66.182
2	Naive Bayes	67.483	67.483
3	SVM	72.628	72.695
4	Decision Tree	66.33	72.695
5	Random Forest	70.327	66.358
6	XGBoost	70.327	66.71

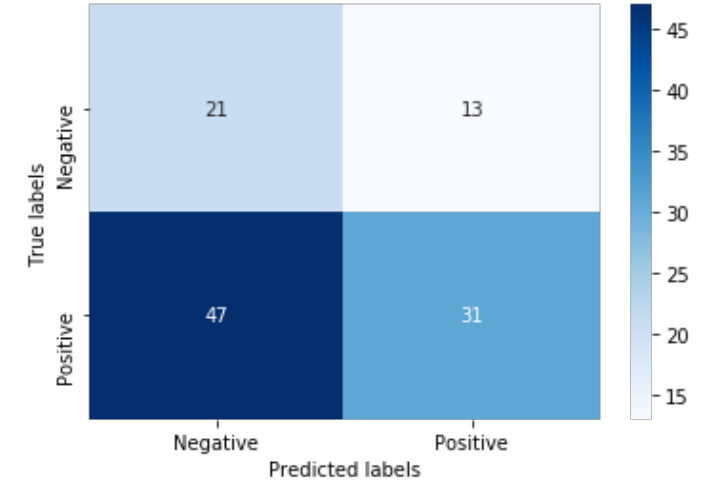
Canadian Election Tweets Sentiment Prediction for Liberal



Canadian Election Tweets Sentiment Prediction for Conservative



Canadian Election Tweets Sentiment Prediction for NDP



Results for 3 Parties
Overall Accuracy = 50.54%

Cross Validation Accuracy: 0.506 (+/- 0.115)
Prediction Accuracy Score: 50.7%

	precision	recall	f1-score	support
0	0.53	0.63	0.58	105
1	0.51	0.53	0.52	78
2	0.43	0.44	0.44	68
3	0.50	0.50	0.50	14
4	0.86	0.33	0.48	18
5	0.40	0.22	0.29	9
6	0.50	0.25	0.33	4
7	0.00	0.00	0.00	2
8	0.00	0.00	0.00	4
accuracy			0.51	302
macro avg	0.42	0.32	0.35	302
weighted avg	0.51	0.51	0.50	302

Best Model for Predicting Negative Reasons: Logistic Regression
Accuracy = 50.7%

How to improve the accuracy of the model?

1. Include more features when developing word frequency and TF-IDF dataset
2. Use larger and more relevant dataset (i.e. last year election data) to train the model
3. Include more keywords when determining political party
Implement more sophisticated models of NLP
4. Augment the dataset to include more tweets with negative reasons other than 'Other'

- What can public opinion on Twitter tell us about the Canadian political landscape in 2019?
 - Many people talk about the election but without mentioning any specific party. It is possible that these people do not have any inclination in which party they are going to vote for, or they might be talking about political parties other than the three classified in the analysis.
 - Among the three major political parties (Liberal, Conservative and NDP), the Conservative Party of Canada received the most tweets. This makes sense as the party has very popular votes. Liberal is the second as it stands on the opposite side to the Conservative Party.
 - Among the entire dataset, the positive tweets and negative tweets are about half to half, which results in a comparatively balanced dataset.
- How each party is viewed in the public eye?
 - Liberal Party: Positive tweets and negative tweets are about the similar amount, indicating that the overall sentiment towards the party is fairly neutral.
 - Conservative Party: There are more negative tweets mentioning the party than negative ones, meaning more dissatisfaction is discovered on Twitter towards the party.
 - NDP: Fewer tweets mentioned the party, indicating less attention from the public, however, the party received twice more positive tweets than negative ones.

