

## EECS4415 – RoadMap

### Data Collection

- > Collect data through Twint
  - Split by provinces -> Ontario, Quebec etc...
  - Split by Main City's -> Toronto, Vancouver, Quebec City
  - Split by ridings -> North York, Downtown etc...
- > Split data for each party; Liberal, Conservative, NDP

### Cleaning Data

- > Remove Emoji's, URL links, Non-Char's, Translate French

### Analysis/Tasks

1. **What is the political sentiment distribution of major geo-locations in Canada?**
  - i. Run a sentiment analysis on the provinces, cities and ridings
  - ii. Run it for each Liberals, Conservatives, NDP
  - iii. **Goal: Create a chart of sentiment vs time (1 month intervals) that shows the distributions sentiment of each party's**
2. **Which riding/city/province have a positive/negative sentiment towards a political party?**
  - i. Using the sentiment analysis data, create a map that shows which party is the most favored (most positive) for each province then for each city and lastly each riding
  - ii. **Goal: Create a Map showing the most positive party**
3. **What do people value in each riding/swing state? (e.g. Taxes, housing, education, etc.)**
  - i. Using our data set of tweets, look for the occurrence of keywords in each tweet for each province/city/riding
  - ii. TOPICS: Crime, Education, Environment, Families, Foreign Policy, Government, HealthCare, Immigration, Jobs & Economy, Science & Tech, Taxes
  - iii. For each topic create a set of keywords. Use those keywords and total the sum of each of them to create a total value for each topic
  - iv. **Goal: Create a bar chart that shows the value for each topic of each location**
4. **What's the percentage that a specific riding will have a political shift in the upcoming election?**

- i. Figure out the ridings that change the most (ridings with 50/50 sentiment)
- ii. Ridings that are closer to 50/50 sentiment split have a higher percentage that it will have a political shift
- iii. **Goal:** Create a map that shows the percentage of each riding in terms of whether it is closer to 50/50 or not

## 5. How does the public sentiment about a candidate change over time?

- i. Collect new data that is only about the candidate for each party (Use twint and search only for Trudeau, Scheer, and Jagmeet)
- ii. Run a sentiment analysis on that data
- iii. **Goal:** Determine two things;
  1. The current sentiment for each political figure
    - a. Create a circle chart of positive, negative and neutral sentiment
  2. The sentiment over time of 1-month intervals
    - a. Create a sentiment vs time (1-month intervals) of each candidate

## 6. Streaming (Real Time Sentiment)

- i. Collect real time political tweets
- ii. As they are collected run a sentiment analysis on the tweets
- iii. **Goal:** Create a real time sentiment vs time chart for each party (interval every 1 min?)

**End Goal:** Create a visual and interactive analysis platform

