**Domain Description and Motivation**

* 1. **- What is the data domain?**

The data domain for the project consists of political tweets in Canada specifically splitting tweets among the different political ridings of Canada. The data set includes tweet contents (what the user has tweeted), username, user ID, the time and date the user posted the tweet, the location (latitude and longitude and/or city) and hashtags used.

* 1. - **What is the goal of your project?**

Our goal over the course of this project is to conduct a sentiment analysis over Canadian twitter data relating to politics. Throughout our analysis, we aim to understand political sentiment across Canadian ridings and use the sentiment to identify what political party that specific region will be. We hope that our sentiment analysis model can be used as a platform by politicians to better understand what people value in Canada and to use the information to further improve their campaign.

* 1. - **What is the motivation for rigorous data analytics?**

The motivation for our data analysis stems from helping politicians better understand the public perception of their campaign and improve public sentiment about their campaign and themselves. A single misquote or negative news about a candidate can be the difference between him/her winning or losing the election. It becomes key to have a platform to guide and direct the candidate’s campaign. As a result, we aim to provide an analysis platform that provides as much insight as possible about people’s sentiment of the Canadian election. The sentiment analysis is narrowed down by province/city/riding and what each of those locations value allowing candidates to improve their campaign and public sentiment.

1.4 - **What are the questions that you want to answer?**

* What is the political sentiment distribution of a major geo-locations in Canada?
* Which riding/city/province have a positive/negative sentiment towards a political party?
* What do people value in each riding/swing state? (e.g. taxes, housing, education, etc.)
* What’s the percentage that a specific riding will have a political shift in the upcoming election?
* Time Analysis: How does the public sentiment about a candidate change over time?

1.5 - **Why is the analysis important?**

(Elections are a vital part of democracy allowing people to vote for the candidate, they think can best lead the country. Throughout the election, a lot of campaigning is done by each party. As campaigning is done, ….) [Long version]

The analysis is important because our data analytics help politicians understand and gain insights of people's sentiment and values across Canada. As a result, this helps politicians improve their campaign and public sentiment by guiding their campaign to prioritizing specific campaign activities. [Short and sweet version]

1.6 - **What are a few potential applications?**

* Provide a platform for politicians that analysis’s their campaign and public sentiment to help them better understand how people view their political campaign and allows them to make data driven decisions to further improve their public sentiment and campaign.
* Train different deep learning models to predict future election results using the data from our sentiment analysis as well as including other information such as past twitter sentiment analysis and voting behavior
* Creating a generalized twitter sentiment analysis model that can be used in different countries during their elections