

“PROJECT PROPOSAL – SENTIMENT ANALYSIS: TRUMP TWEETS POSITIVE OR NEGATIVE?”

Problem Statement –

This project aims to propose the sentiment analysis of political views of Twitter users during the 2020 American Presidential Election campaign. This is an important study as public opinion for a particular candidate will impact the potential leader of the country. We are relying on Twitter platform as it acquires a large diverse data set representing the current public opinion about President Trump.

The business objective of this project is to answer following questions:

- What is the public sentiment distribution across the various locations/ states in U.S.?
- Which states require more strategic campaigning/ awareness by Trump Government?
- How the sentiments of public changed over time in year 2020?
- Was there a statistically significant difference in sentiments expressed in first quarter compared to the following ones?
- Does the positivity of a tweet influence the quantity of retweets/favorites?

To address all these general questions, we perform the analysis to study the current state of public opinions.

Data Collection-

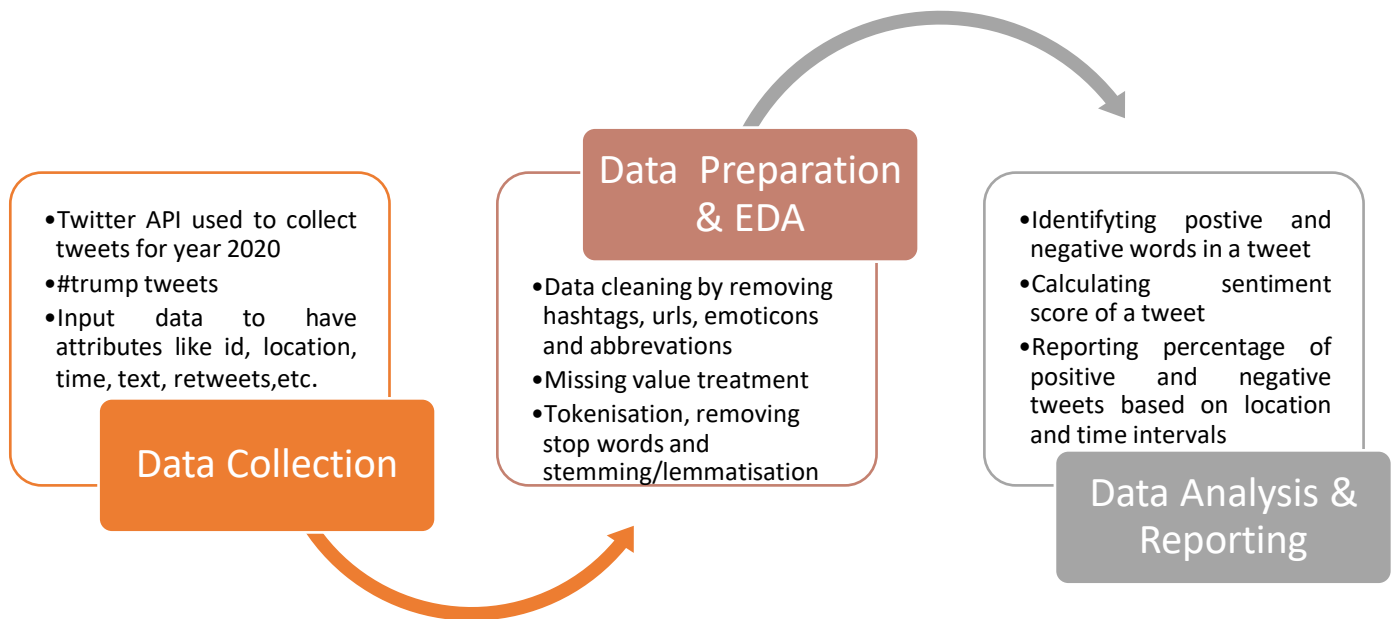
We plan to collect data from Twitter social network from January 1, 2020 to October 31, 2020 using the official Twitter API to get tweets, user profiles, and their contact networks. This dataset will contain all major attributes like actual tweet text, retweets, user id, location, time of tweets, no. of likes, etc to analyze emotions and draw conclusions based on locations and quarters.

It is noteworthy that the American Presidential Election is going to happen on November 3, 2020 and hence, it becomes useful to understand public attitudes before the elections.

Tools involved-

- Python language and Google Colab notebook for coding
- Twitter API for input data i.e. Raw tweets
- Building Apache Spark Streaming Application on Twitter and collect tweets with hashtag **#trump**
- Performing data cleaning and calculating sentiment score by using Hadoop and Spark (pyspark)
- Interactive dashboards by Tableau for presenting public opinion

Steps involved in process-



Note – More tools and models can be used during data analysis depending on nature of data

Output –

The output of this project aims to draw insightful conclusions on change of public sentiments towards Trump Government and their trend analysis across various locations in U.S.

Team 15 Members: Aditya Gajula - 660252623

Ritu Gangwal – 670646774

Viharika Bharti - 655974244