Business Plan

Introduction

Subreddit: #PoliticalCompassMemes

Politics, the art and science of activities related to governing nations, groups, or organizations, is one of those terms we all claim to avoid, yet invariably, it impacts everyone. It is the application of human ideals into a lived reality, which is complicated. It becomes visible one might say as we shift to achieve ideals and encounter the fundamental reality that while we may share many commonalities (physically, mentally/emotionally, etc.) as people, we are fundamentally and entirely dynamic and unique individuals.

The importance of politics includes that firstly, government and politics impact nearly every aspect of our lives. Besides, having knowledge of politics helps make you an informed voter and politics is entertaining. When we find the ability to recognize that politics is the necessary process of managing those differences while focusing on the shared ideal of a particular party to achieve the desired ideal, we can learn to recognize politics is a natural occurrence for finding unity amid complex ambiguity of the people, by the people, for the people.

Political compass measures ones' opinions towards economy and politics. The x axis measures one's opinion of how the economy should be run: "left" is defined as the desire for the economy to be run by a cooperative collective agency, while "right" is defined as the desire for the economy to be left to the devices of competing individuals and organizations. The y axis measures one's political opinions in a social sense. Libertarians believe that personal freedom should be maximized, while authoritarians believe that authority should be obeyed.

The goal for this project is to conduct a segment analysis to classify the posts into different moods. We will explore the posts and first classify them into different politic topics, like Ukraine and Russia, Brexit, etc. And then analyze the mood for all the posts under each topic, trying to tell the altitude of these posts.

Project Plan

1. **What is the distribution of the length of each post?**

   - **Technical Proposal:** Collect all the posts and get the length of each post and make a distribution plot based on that.

2. **Define the topic of each post.**

   - **Technical Proposal:** Use NLP methods to extract the hashtag or the key words of each post.

3. **Extract the key works in each post.**

   - **Technical Proposal:** Remove all the stop words, non-characters, and only keep the English contents.

4. **What is the distribution of the number of posts for each topic in the timeline?**

- **Technical Proposal:** For each topic, get the number of posts in each day and conduct EDA based on it.

5. **Analyze the moods of each word.**

   - **Technical Proposal:** Use NLP methods to transfer each word in the post to a mood score according to a mood questionnaire.

6. **Determine the mood score of each post.**

   - **Technical Proposal:** Based on the score of each word in that post, the average mood score of that post will be calculated.

7. **Determine the importance of each post according to the number of votes, comments, etc.**

   - **Technical Proposal:** Rank the post in the order of importance, in line with the hotness of that post, namely votes and comments.

8. **What is the trend of mood in the timeline of each topic?**

   - **Technical Proposal:** Make a plot to visualize the trend of mood score in each day for each topic.

9. **Will the mood have a drastic change when a new politic event comes out?**

   - **Technical Proposal:** Use NLP to analyze what happened when a new politic event comes out.

10. **Will the mood of the most important posts have an impact on the mood of the following posts?**

    - **Technical Proposal:** Use NLP to analyze the impact those more important posts have on the mood of the following posts.