Developing Political Personas Through Generative Language Modeling

Tonoya Ahmed
ttahmed@ucsd.edu

Sruthi Papanasa
spapanas@ucsd.edu

Vivian Lin
vlin@ucsd.edu

Anna Liu
a2liu@ucsd.edu

Haojian Jin
h7jin@ucsd.edu

Abstract

The modern world divides people based on political ideals and beliefs, so this model of learning the different personas of different people's opinions answers political questions that help us efficiently gain a better understanding of people's responses. This addresses the limitation of surveys and cold-calling. In past approaches, others have not been able to solve the problem at hand in order to help political candidates locate their target audience and forecast the success of policies. We are specifically evaluating people's opinions on presidential candidates, so by taking the sentiments of people's opinions in Twitter data and using LangChain agents to create "personas," we believe this will take more factors into account that have previously been neglected as well as provide benefits that have not been focused on before. A LangChain agent will use our language model to interact with other tools to interact with the GPT-3.5 Turbo-Instruct API developed by OpenAI and also to answer these questions regarding political topics (Sahota 2023). As a result of this model, we are expecting fitting responses to questions regarding political ideals, current political issues, and presidential candidates.

Website: https://sruthipapanasa.github.io/capstone_website/
Code: https://github.com/sruthipapanasa/B14_Q2_Submission

1 Introduction

1.1 Project Description

The main goal of politicians is to cater to their supporters and understand their wants. The traditional method of finding out what people of different parties want is traditionally done through surveys and phone calls to gather information. The downside of this method is that it can be extremely time-consuming and expensive to send out surveys. Our project pertains to helping this problem. Twitter data contains a wealth of information from the public about their thoughts on certain policies and politicians. We want to take Tweets pertaining to politics from different political groups to perform sentiment analysis on whether a tweet is a Trump or Biden supporter. This will result in 5 clusters (strong supporter of Biden, weak supporter of Biden, neutral, weak supporter of Trump, and strong supporter of Trump), which will then each be trained on its own LangChain Agent. We will then be able to gauge different responses and reasoning based on various political questions. Surveying these created personas is meant to act as a cheaper and quicker substitute to surveying real people, while still being able to provide accurate political opinions for each of these groups.

1.2 Literature Review

Research has been done in the past pertaining to the realm of conversational language models and generative agents in order to replicate specific human behaviors. Large language models and their ability to accurately represent societal issues have also been studied. For example, (Durmus et al. 2023) built a large language model to simulate human responses, analyzing its ability to represent multiple perspectives from various countries. They discovered that when it comes to evaluating multiple perspectives, large language models don't take into consideration cultural stereotypes. When looking at our own project, our end goal would be to create models that are able to properly take into account these stereotypes and biases, especially since our project centers around political data. Another research that has been done that aligns closely to our domain deals with generative agents. Park et al. (2023) explored the realm of generative agents and their applications to various environments. Not only are these agents able to simulate human behavior, but by combining them with large language models, the research illustrates their ability to replicate personas and everyday social behaviors. We aim to fit an LLM-powered generative model to create a variety of different user personas, emulating the sentiments of political groups.

1.3 Data Description

Our data consists of two separate CSV files. The first dataset contains all of the tweets from the 2020 election that included the hashtag #biden in it. The second dataset contained all the tweets that included the hashtag #trump. Neither dataset was exclusively partisan to either candidate. Rather, they consisted of opinions and sentiments from voters who held various perspectives on the two candidates, irrespective of their stance. It included any tweet that had the corresponding hashtag, regardless of whether that hashtag was used to express support or opposition for the tagged candidate.

2 Methods

2.1 Data Collection

To start off, we first did some preliminary research into our topic and selected a usable and appropriate dataset from Twitter. This dataset contains tweets from the 2020 Presidential Election with the hashtags #biden and #trump, which allows us to extract various political personas from the given users. The dataset is also formatted in a way that makes it easy to extract the data for data pre-processing as well as feed into large language models. After choosing our dataset, we proceeded to clean and process the given tweets. We did some basic data cleaning by removing unnecessary columns (i.e. number of likes), turning all words into lowercase, and also filtered out non-English tweets and all the users with less than 20 tweets. We also removed all of the unnecessary punctuation, emojis, hashtags, and stop words. We took out all the tweets that mentioned both Biden and Trump, because the sentiment analysis would not accurately work. We also made everything lowercase in order for Biden or Trump sentiments to be easier to detect. The result are two separate datasets with one about Biden and one about Trump.

2.2 Sentiment Analysis and Clusters

After cleaning the two datasets about Joe Biden and Donald Trump, we then performed a sentiment analysis on both datasets to evaluate how each tweet's user felt about the candidate. The resulting sentiment scores were then converted on a scale of positive, neutral, or negative. This was also then converted to a scale of 1 to -1, 1 being positive about the candidate, 0 being neutral, and -1 being negative about the candidate. This was done for each dataset.

We then wanted to divide the users into 5 clusters total. The first group will consists of users who strongly support Biden. The second group will be for those who weakly support Biden. The third group will be for those who are neutral. The fourth group will be for those who weakly support Trump, and the fifth group will be for those who strongly support Trump. To do this, we started off with finding the average of each users' sentiment scores about each candidate. For example, if a user had 100 tweets written, we will find the average sentiment of all 100 tweets. From there, we evaluated the histogram of these averages and concluded that we should use a threshold of 0.22 to decide whether the user is for Trump, Biden or neutral. If the average scores were between -0.22 and 0.22, they were in the neutral cluster. If they were above 0.22, they were considered Biden supporters. If they were below -0.22, they were Trump supporters. To then divide into the five clusters, we used a threshold of 0.5. For the users already decided as Biden supporters, if their average score was above 0.5, they were considered strong Biden supporters. If they were below 0.5 (assumed also above 0.22), they were weak Biden supporters. For the users already decided as Trump supporters, if their average score was below -0.5, they were considered strong Trump supporters. If they were above -0.5 (assumed also below -0.22), they were weak Trump supporters.

2.3 The LangChain Framework

As a general overview, we are using LangChain as our model integration framework and the GPT-3.5 Turbo-Instruct API developed by OpenAI to create a Large Language Model-powered generative model that can use trained clusters to produce the political affiliation, policy related opinions, opinion polarity, and emotional tone relating to common political discourses of 5 select voter populations.

To create political "personas" after performing sentiment analysis, we opted to use LangChain agents, which is a framework for developing applications powered by large language models. A LangChain agent will use our large language model to interact with other tools to interact with an API and also to answer these questions regarding political topics (Sahota 2023). Like mentioned above, we used the OpenAI language model to build a pipeline that was able to process our tweets and create personas from them. The LangChain agents can parse through a dataset and answer any prompts given. We created a LangChain agent for each of the 5 different clusters to create 5 different "personas" representing the average voter in that cluster. We then asked each of these personas a set of political questions in Likert format (5 being strongly agree, 4 being agree, 3 being neutral, 2 being disagree, and 1 being strongly disagree), which would give us an accurate representation for the political opinions of each of these five clusters.

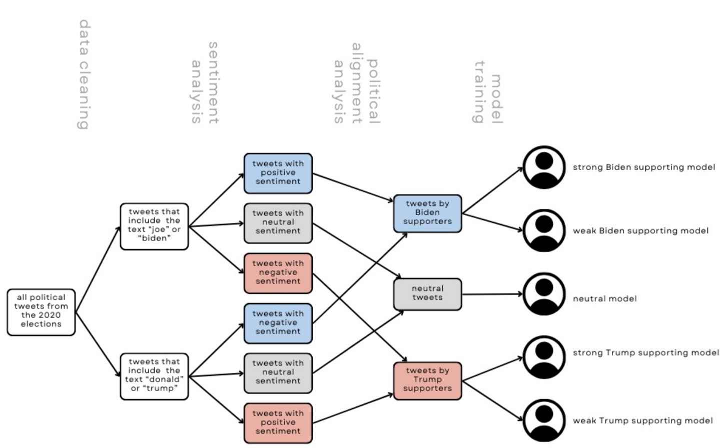

Figure 1.1: Flowchart mapping project structure from data to trained models

3 Results

3.1 Sentiment Analysis and the Five Clusters

To summarize our results after conducting sentiment analysis, we saw that for the histogram regarding tweets that only spoke about Trump, there was a heavy right skew. For the histogram about Biden, it is more a trimodal distribution, with the left most peak being the tallest and the middle and right peaks around the same height.

We then discovered that out of a total of 1,260 users, 645 support Biden, 461 support Trump, and 154 are neutral. In percentage terms, about 51% users are Biden supporters, about 37% users are Trump supporters, and 12% are neutral.

After dividing into the five clusters mentioned before, we found that 533 users are strong Biden supporters, 112 users are weak Biden supporters, 154 are still neutral, 362 are strong Trump supporters, and 99 are weak Trump supporters. Out of the total Biden supporters, about 83% are strong supporters and 17% are weak supporters. Out of the total Trump supporters, about 79% are strong supporters and 21% are weak supporters.

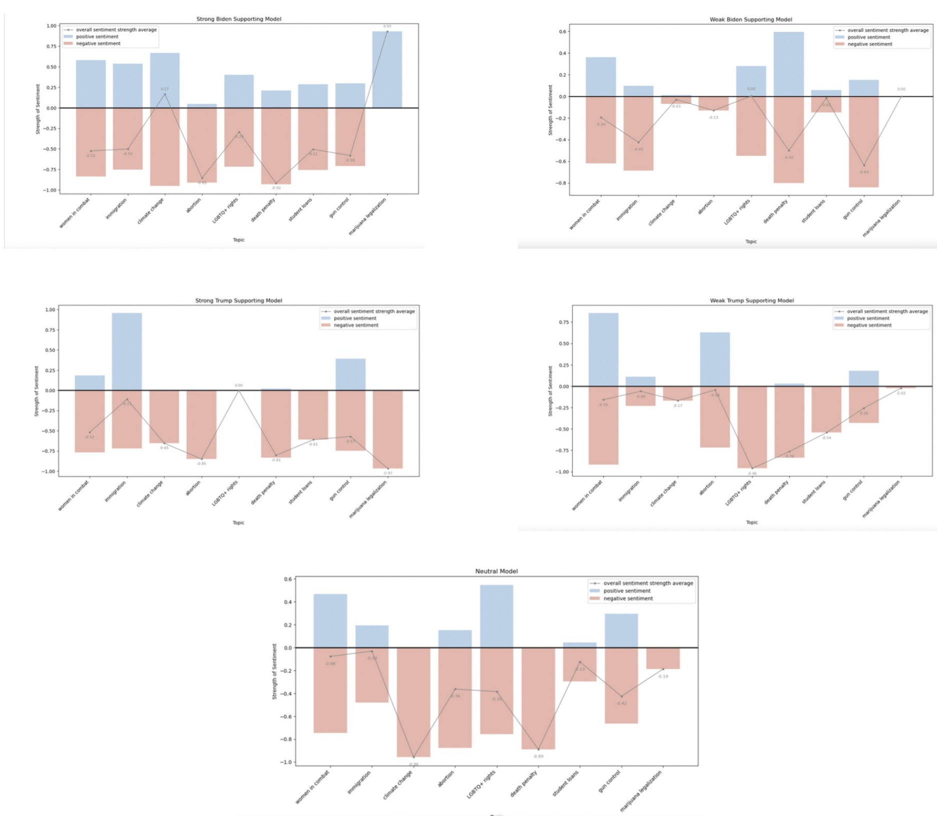

Figure 1.2: Displays sentiment analysis by topic and candidate affiliation

3.2 NLP Findings

Most of the results show strongly negative sentiment scores, which means it is more common for users to express support for a certain candidate by expressing strong negative views about the competing candidate, as opposed to strong positive views about their own. In Figure 1.2 above, you can see the distribution of both the strength of positive and negative sentiments regarding controversial topics from different parties.

3.3 LLM Findings

The outputs for the models were very emotionally charged, polarizing, and made use of strongly opinionated language. Even the neutral model expressed its moderate views with strong conviction. This result aligns with our initial expectations that Twitter contains the most intense opinions, and this is reflected by our model.

3.4 Agents and Bias

According to an article "ChatGPT Leans Liberal, Research Shows" by the Washington Post, the ChatGPT chatbots are ingrained with inherent political biases when dealing with political data, and because our model uses the ChatGPT API, this could be another cause for the results we found from our model (Vynck (2023)). Despite the designers of ChatGPT wanting to make the model neutral, it still leans liberal due to the data from the internet it was trained on.

4 Discussion

4.1 Limitations

A huge limitation we faced is the number of tweets we can sample from the five different clusters when creating LangChain agents. We originally wanted to sample 500 tweets; however, we ran into issues about having too many tokens. Therefore, we can only sample 75 tweets right now. We tried shortening the length of tweets as that was a potential solution per our research, but this did not fix the errors. The datasets we used also have limited data and are very skewed towards tweets with negative sentiments, making it a bit difficult to perform analysis. Another limitation we ran into was how datasets were only available for Trump and Biden, limiting the insights we can gain from other political parties. For this project specifically, only data from 2020 was used, and thus the current political climate cannot be the most accurately reflected.

5 Next Steps

As for our next steps, we want to be able to feed our models the latest Twitter data to get a more accurate picture of the current political climate. We also want to find a way to sample more Tweets than we can with the API limitation of our current model. Lastly, we want to have our model output more streamlined answers through prompt engineering so that further analysis can be performed on the model responses.

References

Durmus, Esin, Karina Nguyen, Thomas I. Liao, Nicholas Schiefer, Amanda Askell, Anton Bakhtin, Carol Chen, Zac Hatfield-Dodds, Danny Hernandez, Nicholas Joseph, Liane Lovitt, Sam McCandlish, Orowa Sikder, Alex Tamkin, Janel Thamkul, Jared Kaplan, Jack Clark, and Deep Ganguli. 2023. In *Towards Measuring the Representation of Subjective Global Opinions in Language Models*. [Link]

Park, Joon Sung, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. In *Generative Agents: Interactive Simulacra of Human Behavior*. [Link]

Sahota, Harpreet. 2023. "Implementing Agents in LangChain." [Link]

Vynck, Gerrit De. 2023. "ChatGPT leans liberal, research shows." [Link]