

Analyzing Political Data on Reddit

Aniket Kulkarni
Erik Jonsson School of
Engineering and
Computer Science
The University of
Texas at Dallas

aniket.kulkarni@utdallas.edu

Rhugaved Narmade
Erik Jonsson School of
Engineering and
Computer Science
The University of
Texas at Dallas

rhugaved.narmade@utdallas.edu

Swetha Malaivaiyavur Elayavalli
Erik Jonsson School of
Engineering and
Computer Science
The University of
Texas at Dallas

sxm220052@utdallas.edu

Hridya Dhulipala
Erik Jonsson School of
Engineering and
Computer Science
The University of
Texas at Dallas

hridya.dhulipala@utdallas.edu

Abstract— This project introduces an innovative approach to analyzing political discourse on Reddit through sentiment analysis and summarization techniques. Utilizing the Reddit Stream API, posts, and comments relevant to politics are collected in real-time, forming the basis for sentiment analysis. By employing advanced sentiment analysis models, the project discerns nuanced sentiments towards various political figures or parties. Additionally, the project utilizes state-of-the-art large language model APIs to summarize key discussions, providing concise insights into prevailing public sentiments. This paper elucidates the methodology employed, discusses the implications of the findings, and underscores the significance of leveraging data-driven approaches in understanding contemporary political discourse on social media platforms.

Keywords— *Reddit, political sentiments, sentiment analysis, streaming, Kafka, summarization, SVM, large language models, APIs*

I. INTRODUCTION

In the digital age, social media platforms have emerged as pivotal arenas where political discourse thrives, shaping public opinion, and influencing electoral outcomes. Among these platforms, Reddit stands out as a dynamic hub where diverse perspectives converge, making it a rich source of political data ripe for analysis. This project, "Analyzing Political Data on Reddit: A Sentiment, Summarization, and Analytical Approach," delves into the realm of Reddit's political discussions, particularly focusing on the vibrant subreddit, *r/politics*, against the backdrop of the impending U.S. elections.

By leveraging real-time streaming data from Reddit, this project embarks on a meticulous journey through the digital landscape, honing in on discussions surrounding the two prominent presidential contenders, Trump and Biden. Through a systematic filtration process, posts pertinent to these candidates within the last hour are meticulously curated, serving as the foundation for a multifaceted analysis.

The project unfolds through three core analytical dimensions:

General Data Analysis: Through rigorous examination, this facet delves into the frequency of posts related to Trump and Biden, unveiling trends in engagement, discourse, and topic distribution. Utilizing advanced techniques, including word frequency analysis and user engagement metrics, the project uncovers the pulse of Reddit's political dialogue, shedding light on prevailing sentiments and prevailing topics.

Sentiment Analysis: Employing Support Vector Machines MapReduce-based Neural Networks in PySpark, this endeavor undertakes the intricate task of deciphering sentiment from Reddit posts. By discerning between positive, negative, and neutral sentiments, the project provides a nuanced understanding of the public's perception towards the respective candidates, bolstered by insightful visualizations.

Summarization: Recognizing the vast expanse of Reddit's discourse, the project streamlines comprehension through sophisticated summarization techniques. Leveraging the power of the GPT API and innovative prompt engineering methodologies, the project distills the essence of discussions into concise summaries, offering campaign managers and stakeholders a panoramic view of the political landscape in real-time.

The significance of Reddit as a crucible of political dialogue cannot be overstated. As a breeding ground for diverse viewpoints and grassroots activism, Reddit serves as an invaluable platform for gauging public sentiment, identifying emerging trends, and informing strategic decision-making, particularly in the context of electoral campaigns.

Through the amalgamation of comprehensive analytics, sentiment insights, and streamlined summarization, this project presents a pioneering approach to political data analysis. By automating the extraction of actionable insights from Reddit's dynamic discourse, this endeavor equips campaign managers with a potent tool to navigate the complex terrain of public opinion, empowering them to make informed decisions in the run-up to the elections.

Methodologically, the project harnesses a diverse array of tools and technologies, ranging from PRAW for Reddit data retrieval to PySpark for sentiment analysis, and GPT API for summarization. By synergizing these cutting-edge

methodologies, the project exemplifies the fusion of advanced analytics and computational prowess, epitomizing the forefront of data-driven political research.

In essence, this project not only elucidates the quantitative landscape of political discourse on Reddit but also offers qualitative insights into the sentiments, themes, and narratives that shape the digital agora. By harnessing the power of data analytics, sentiment analysis, and summarization, this endeavor endeavors to unravel the complexities of contemporary political discourse, illuminating the path forward for stakeholders and researchers alike.

II. METHODOLOGY

A. Overall Architecture

1. Data Acquisition

- Reddit data is retrieved using PRAW (Python Reddit API Wrapper).
- Producer1 collects post data (title, text, comments, upvotes, URLs, etc.) from the r/politics subreddit at regular intervals and stores it in a Kafka queue in JSON format.

2. Data Filtering

- Consumer1 retrieves data from the Kafka queue and applies filtering using MapReduce to isolate posts related to Trump or Biden. This is achieved through Named Entity Recognition (NER) using PySpark. Filtered data is then written to another Kafka queue.

3. Sentiment Analysis

- Consumer2 reads the filtered data from the Kafka queue.
- Utilizes a trained sentiment analysis model (based on Support Vector Machines, SVMs) to classify sentiments (positive, negative, neutral) for posts about Trump and Biden.
- Generates a graph illustrating the distribution of sentiments for each candidate.

4. Summarization

- The summarization model processes all posts from the past hour using the GPT API.
- Returns a concise summary of the key discussions and topics, which is displayed as output.

5. General Analytics

- After data filtering, Consumer2 initiates the process of general analytics.
- This involves various analyses such as post frequency, user engagement, and URL analysis.

B. Methodology of General Analytics

1. Post Frequency Analysis

The first component of our analysis focuses on post frequency. We employ the MapReduce framework to analyze the frequency of posts related to both Biden and Trump. To achieve this, we group posts based on specific time intervals (e.g., hourly) to track changes over time. The results are then presented as a Bar graph, with time on the x-axis and post frequency on the y-axis.

2. User Engagement Analysis

Next, we delve into user engagement metrics. By analyzing factors such as upvotes, comments, and shares for each post, we gauge user interaction and engagement. Our approach combines PySpark SQL and MapReduce for efficient data processing. The resulting insights are visualized through line graphs, allowing us to observe trends in user engagement over time.

3. URL Analysis

The third aspect of our methodology involves examining URLs shared within posts. We aim to identify patterns or trends in the sources referenced by users. Regular expressions (regex) are employed to extract URLs from post content. Leveraging PySpark, we analyze and aggregate this data. Additionally, we generate a bar graph illustrating the frequency of different news organizations mentioned in posts. This analysis aids in understanding the sources influencing discussions.

By incorporating these detailed analyses into the overall architecture, the project provides a comprehensive understanding of political discourse on Reddit, enabling stakeholders to glean valuable insights into public sentiment, engagement, and trends surrounding the Trump vs. Biden narrative.

III. TEXT SUMMARIZATION

Text summarization is vital for understanding sentiment in political news as it condenses lengthy articles into concise summaries, allowing readers to grasp key information efficiently. These summaries highlight prevailing sentiments, aiding in the analysis of media bias or public opinion. Additionally, summarization facilitates comparative analysis between different sources and enables quicker decision-making for policymakers, journalists, and the public alike. Overall, it serves as a crucial tool in navigating the complex landscape of political news, extracting key sentiments, and making informed interpretations.

The text summarization tool we used was the Generative Pre-trained Transformer architecture which is built on the traditional Transformer model. At its core, the model consists of a stack of Transformer encoder-decoder layers. However, unlike in models like BERT, it only uses the decoder part. Each layer in the decoder is composed of multi-head self-attention mechanisms and position-wise feedforward neural networks. The text summarization model heavily relies on the attention mechanism, which allows it to focus on different parts of the input text when generating outputs. This mechanism

enables the model to capture long-range dependencies and context more effectively compared to earlier NLP architectures.

The input to the Generative Pre-trained Transformer (GPT) model is a stream of political opinions and headlines collected in real-time using the Kafka producer algorithm. This means that as opinions and headlines related to various political topics, including Joe Biden and Donald Trump, are posted on Reddit or other platforms, they are continuously gathered and fed into the model. The Kafka producer algorithm facilitates this real-time collection, ensuring that the model receives the latest data as it becomes available.

The input stream is dynamic and constantly evolving, reflecting the dynamic nature of online discussions and news coverage surrounding political figures like Biden and Trump. The data collected encompasses a wide range of opinions, from supportive to critical, and headlines covering various aspects of their actions, policies, and controversies.

Once all the data relevant to the problem statement – opinions associated with Joe Biden or Donald Trump – is streamed and collected, it undergoes a filtering process. This filtering ensures that only the items pertinent to Biden and Trump are retained, forming a text corpus specifically focused on these two political figures.

The GPT model then processes this filtered text corpus to generate a summarized version of the input. The output is segregated into Joe Biden and Donald Trump-related data, if applicable, to provide a clear distinction between the opinions and headlines associated with each figure.

The summarized output presents a condensed and coherent representation of the original input data, capturing the key sentiments, arguments, and events related to Biden and Trump. This enables users to quickly grasp the prevailing opinions and discussions surrounding these political figures without having to sift through a large volume of unstructured data.

Below is an illustration of the sample input (Figure 1) fed into the model to be summarized and the corresponding summarized output (Figure 1b).

```
User
I'm working on a project to summarize a stream of data from Reddit posts regarding Biden and Trump. Below are a number of posts that I filtered from Reddit. Using them, give me a paraphrased version of the opinion of people on Reddit or from titles along with a summary, don't give a list of what is present in the data below, but use it to form a good summary and compare and contrast them:
[
  ("ID": "Biden", "line": "US solar manufacturers target China-linked imports in new plea to Biden "),
  ("ID": "Trump", "line": "Tax Rates for Big Corporations Fell by Nearly Half After Trump Cuts | And this is the policy Republicans and their presumptive presidential nominee are vowing to extend if they win in November. "),
  ("ID": "Trump", "line": "I wish trump was president for life. "),
  ("ID": "Trump", "line": "'Weirdest moments': Experts call out Trump lawyers' 'planted evidence' defense over damning recording "),
  ("ID": "Biden", "line": "These illegal immigrants are eligible for Obamacare after Biden rule change "),
  ("ID": "Trump", "line": "[FS] - Central New York - 100$ 20gallon tank, will take the Best offer i just had to put a price in this picture is when my tank was thriving, the plants in the middle have died off i believe and so have the plants on the right. if you would like pictures of the tank in its current state, i will send them.\n\nI used to keep this tank in pristine condition but over time my love for it died down, the tank has had algae in it but it is all salvageable. in the tank is Black sand, Dragonstone, spiderwood, jungle valisneria, subwassertang, heater and Hob filter included along with a light. currently only has malaysian trumpet snails living in it."),
  ("ID": "Trump", "line": "Donald Trump u2018would force Nato members to spend 3% on defenceu2019 "),
  ("ID": "Biden", "line": "Holy Shit! In 4 years of the Biden Regime this is the only yard sign I have seen. I had to pull over to take a snap. Biden Supporters do exist. "),
  ("ID": "Biden", "line": "US diplomat who quit over Biden administrationu2019s Gaza policy speaks out "),
  ("ID": "Trump", "line": "Donald Trump u2018should force Nato members to spend 3% on defenceu2019 "),
  ("ID": "Biden", "line": "US diplomat who quit over Biden administrationu2019s Gaza policy speaks out "),
  ("ID": "Trump", "line": "Trump https://x.com/iLA_NewsX/status/1786730276477165606"),
  ("ID": "Trump", "line": "Donald Trump supporters wear nappies | US News "),
  ("ID": "Biden", "line": "Bernie Breaks His Silence: Biden Is 'Alienating' Young People (The Rational National) "),
  ("ID": "Trump", "line": "POLITICS SEC charges Trump Media auditor with u2018massive fraudu2019 on hundreds of companies, imposes lifetime ban "),
  ("ID": "Biden", "line": "Maher: Bidenu2019s Loan Plan Is u20182018Tax Dollarsu2019 u2018Supporting This Jew-Hatingu2019 on Campus ")
]
```

Figure 1 (a) : A sample user prompt to be summarized

Based on the provided Reddit posts, it seems like opinions regarding Biden and Trump are varied and sometimes polarized.

For Biden, there's a mix of support and criticism. Some praise his policies, such as making certain immigrants eligible for Obamacare, while others express disillusionment or disappointment, as seen in comments about a perceived lack of visible support for Biden. There are also mentions of diplomatic issues, with a diplomat resigning over the administration's Gaza policy.

Trump-related posts include discussions about his past policies, such as tax cuts for corporations, which are both praised and criticized. Some express strong support for Trump, even suggesting a desire for him to be president indefinitely, while others highlight controversies surrounding his legal defenses or his proposed NATO spending requirements. Additionally, there are mentions of SEC charges against entities associated with Trump and unusual behaviors of his supporters.

Overall, the Reddit posts indicate a diverse range of opinions on both Biden and Trump, reflecting the broader political divisions and debates present in society.

Figure 1 (b) Summarized version of sample input in Fig 1(a)

IV. SENTIMENT ANALYSIS

Sentiment analysis is crucial in understanding public opinion, especially in domains as dynamic as politics where public sentiment can influence decisions and shape electoral outcomes. By analyzing sentiments expressed on social media platforms like Reddit and Twitter, stakeholders can gauge public reactions in real-time and adjust their strategies accordingly. This is particularly relevant in the context of political campaigns where understanding voter sentiment towards different candidates can provide a competitive edge. Sentiment Analysis data is used to predict pre-election opinions of people and determine which candidate is most likely to win at that point in time.

Twitter data of the political landscape containing text and the sentiment score associated with it was used to train a SVM – Support Vector Machine Algorithm which in turn was used along with the stream data collection code to predict the sentiments of the text in real time. The training data is hosted on AWS S3 and can be found [here](#). The size of training data was over 163000 rows which was substantially large enough to train the model well. This data was loaded into a Spark RDD (Resilient Distributed Dataset) object to parallelize the computations to yield results faster than sequential data. This also makes it compatible to run the predictions on incoming streaming data easily. The sentiment analysis component includes a data preprocessing module which removes unnecessary symbols, punctuations, and numbers from the text data so that model understands the sentiments from words rather than these ambiguities. This cleaned RDD was further tokenized on its text feature using PySpark's ml.feature.Tokenizer function. This data was further vectorized using HashingTF function provided by the pyspark.ml.feature library. This HashingTF uses Murmurhash3 algorithm to vectorize data. Murmurhash3 algorithm is also popular in big data and distributed systems and it is also used in Cassandra technology. Finally, the label column contains string value for three classes of positive, neutral, and negative sentiment values, and was converted to 1.0, 0.0 and -1.0 values respectively.

This data was used to train and fit the SVMWithSGD model. The reason SVMWithSGD was chosen over simple SVMModel is due to efficiency and speed of the SGD (Stochastic Gradient Descent) algorithm which chooses one random instance of data to compute the gradient. Unlike normal Gradient Descent which goes sequentially until a minima is found, SGD is batched and random which is more likely to find the global minima too.

The Support Vector Machine (SVM) algorithm stands out as a robust tool for supervised machine learning tasks, adept at handling classification and regression challenges. Its core principle revolves around identifying a hyperplane within an N-dimensional space (where N represents the number of features) to effectively segregate data points into distinct classes. Within this space, numerous hyperplanes could potentially serve this purpose. However, the objective is to pinpoint a hyperplane with the widest possible margin, denoting the maximum separation between data points of different classes. By maximizing this margin, SVM reinforces the confidence in classifying future data points. It demonstrates efficacy across both linear and non-linear datasets.

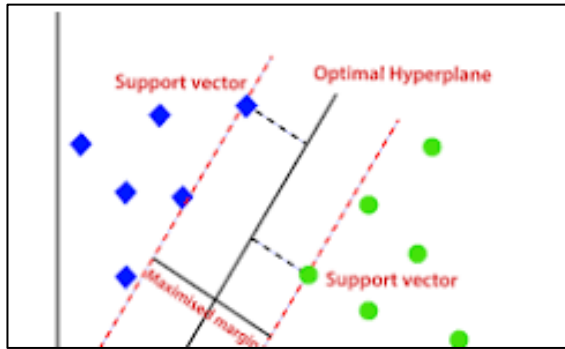


Fig. 2: Support Vector Machine Visual Representation

Hyperplane Equation:

$$w \cdot x + b = 0$$

where w is the normal vector to the hyperplane.

The classification decision function:

$$f(x) = \text{sign}(w \cdot x - b)$$

This model produced an accuracy rate of 85% which is good considering the vast dictionary of words used in the political landscape, most of whose sentiment cannot even be determined by an untrained eye. Sometimes, accuracy does not give the right representation of model performance, so here is the class distribution of correct and incorrect predictions:

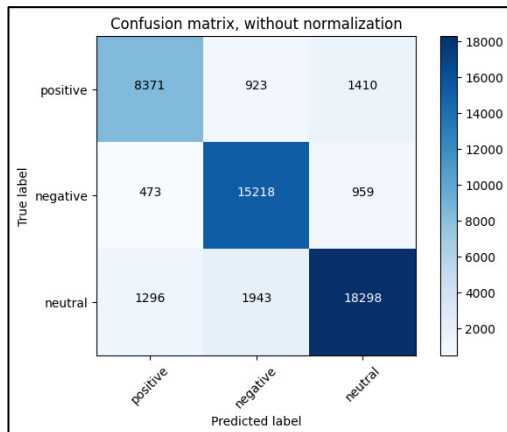


Fig. 3 Confusion Matrix of SVM Predictions

To understand better of how the model performed, here is the normalized confusion from which we can infer the percentages of correct and incorrect predictions:

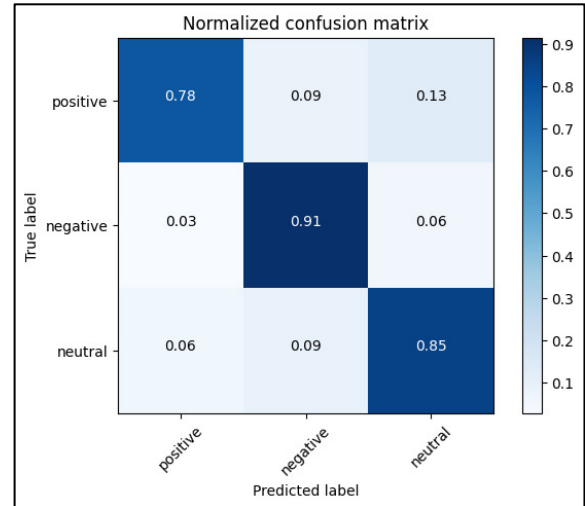


Fig. 4: Normalized Confusion Matrix

As we can see, even the precision and recall of the model is good as the true positive and true negatives rates are much higher than the false positives or the false negatives. This model was saved for further usage directly with the streaming data using the `model.save()` function.

V. RESULTS AND ANALYSIS

The general analytics section of the project yielded insightful findings regarding political discourse on Reddit, particularly surrounding discussions related to Trump and Biden. The results are presented in the form of graphical representations, providing a visual overview of key trends and patterns observed within the data.

A. Post Frequency Analysis

The analysis of post frequency reveals a notable discrepancy between discussions surrounding Trump and Biden on Reddit. Trump consistently garners significantly higher post frequency compared to Biden across the observation period. On day 1, Trump amassed over 100 posts, nearly three times the number attributed to Biden, which stood at approximately 40. This trend persists over subsequent days, with Trump maintaining a substantial lead. For instance, on day 1 and day 2 combined, Trump accumulated around 350 posts compared to Biden's approximately 100. Similarly, over the span of three days, Trump's post frequency exceeds 1100, whereas Biden's hovers around 330. These findings underscore the prominence and intensity of discussions pertaining to Trump on Reddit, suggesting a greater level of engagement and interest among users in comparison to discussions surrounding Biden. [Refer fig. 5, 6, 7]

B. User Engagement Analysis

The analysis of user engagement metrics highlights a stark contrast between discussions surrounding Trump and Biden on Reddit. Comments on posts related to Trump significantly outnumber those on Biden, a trend primarily attributed to the controversies surrounding Trump. Over a period of three days, Trump posts amassed over 25,000 comments, whereas Biden posts received only a few hundred. Similarly, Trump exhibits exponentially higher upvotes compared to Biden, with over 200,000 upvotes accrued over the same timeframe, contrasted with approximately 25,000 for Biden. These findings underscore the heightened level of engagement and interest among Reddit users towards discussions involving Trump, indicative of the intense scrutiny and fervor surrounding his political discourse. [Refer fig. 8, 9]

C. URL Analysis

The analysis of URLs shared in posts unveils the prominent sources referenced within political discussions on Reddit over a three-day period. The top four news sources identified during the analysis, in descending order of frequency, are thehill.com, salon.com, newrepublic.com, and nytimes.com. Additionally, nbcnews.com and cnn.com also emerged as significant contributors to the discourse. These findings underscore the diverse array of media outlets influencing political narratives on Reddit, with a notable presence from reputable news organizations. [Refer fig. 10]

D. Sentiment Analysis

The sentiment analysis encompassed three distinct time frames: single-day streaming data, weekly data, and bi-weekly data. The findings reveal intriguing patterns in sentiment distribution between discussions pertaining to Trump and Biden.

Over a single day, Trump exhibited a higher proportion of positive posts compared to Biden. However, upon extending the analysis to a two-week period, Trump's proportion of negative posts surpassed that of Biden. Notably, most posts across all three timeframes were categorized as neutral, indicating a tendency towards impartiality in Reddit discussions regarding the two political figures. [Refer fig. 11, 12, 13]

These observations underscore the dynamic nature of sentiment within political discourse on Reddit, with fluctuations observed over varying time frames.

VI. CONCLUSION

The project presents an innovative approach to analyzing political discourse on Reddit, employing sentiment analysis and summarization techniques to glean insights from real-time data. Through meticulous data collection, advanced analytics, and state-of-the-art models, the project offers a comprehensive understanding of public sentiments towards political figures, particularly Trump and Biden. The findings underscore the dynamic nature of political discourse on Reddit, revealing trends in post frequency, user engagement, and sentiment distribution. Despite the prevalence of neutral sentiments, fluctuations are observed over different time frames, highlighting the evolving nature of public opinion. Overall,

the project demonstrates the significance of leveraging data-driven approaches in deciphering contemporary political dialogue, providing stakeholders with actionable insights to inform strategic decision-making.

VII. FUTURE WORK

In future research, enhancing sentiment analysis with advanced models like RNNs or transformer models could improve accuracy. Exploring abstractive summarization methods and expanding analysis to include more political topics and subreddits would offer a broader perspective. Additionally, integrating real-time analysis into user-friendly applications and conducting longitudinal studies for extended insights are avenues for further exploration.

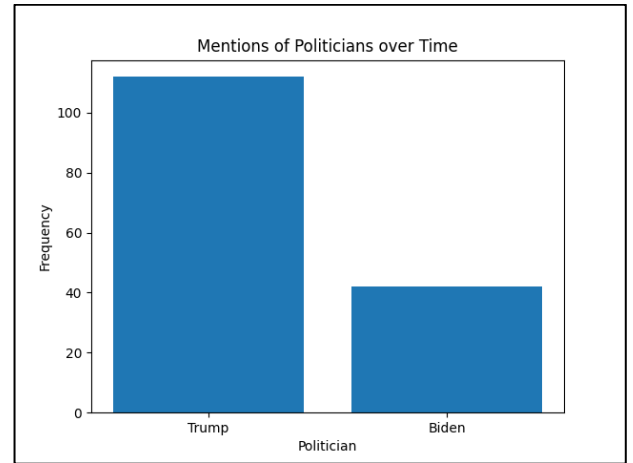


Fig. 5: Trump vs Biden Post Frequency over 1 day

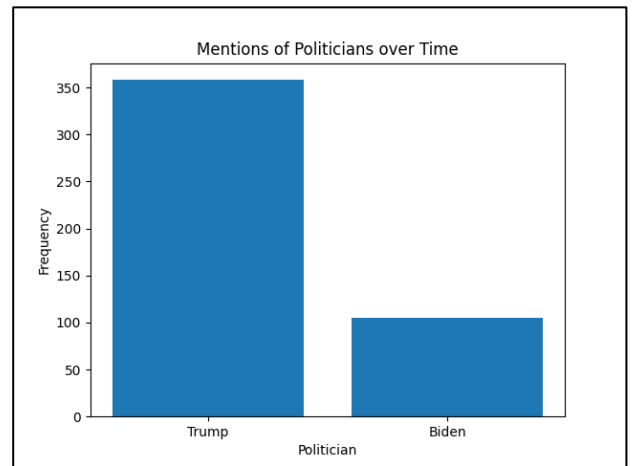


Fig. 6: Trump vs Biden Post Frequency over 2 days

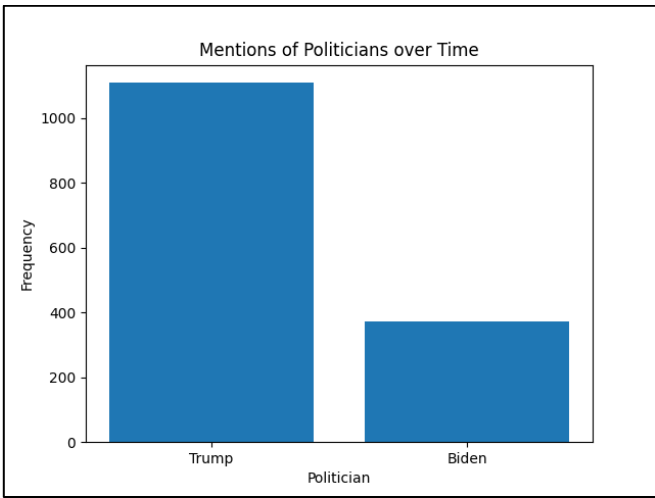


Fig. 7: Trump vs Biden Post Frequency over 3 days

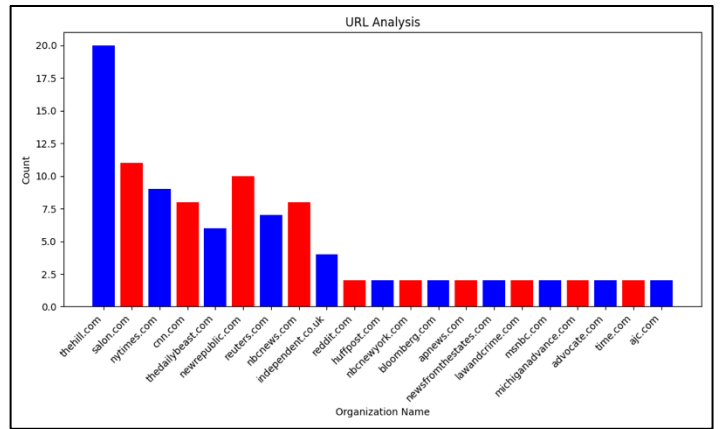


Fig. 10: Frequency of URL Sources over 3 days

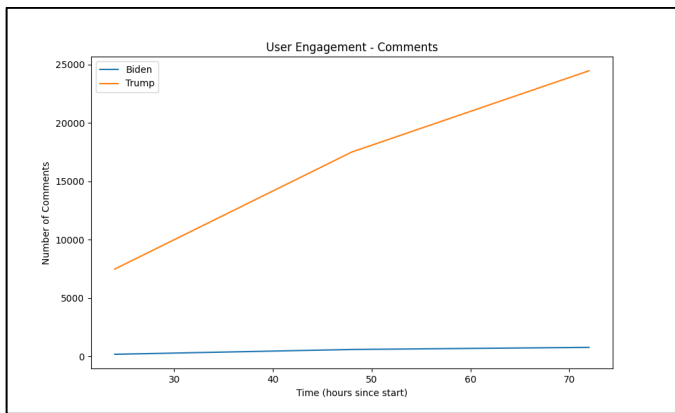


Fig. 8: Trump vs Biden Comment Count over 3 days

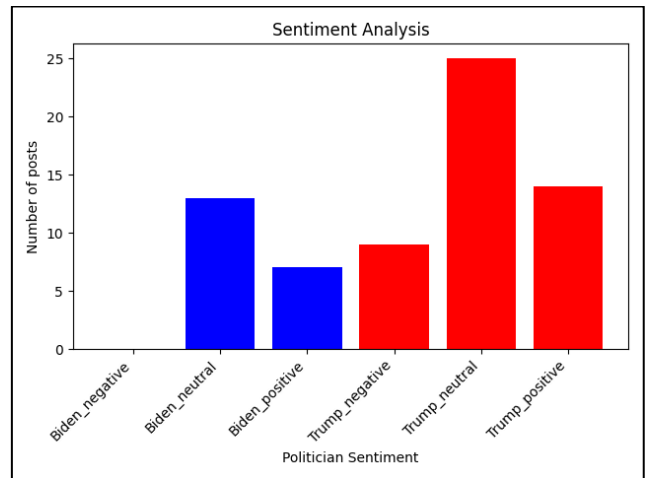


Fig. 11: Biden vs Trump Sentiment Count over 1 day

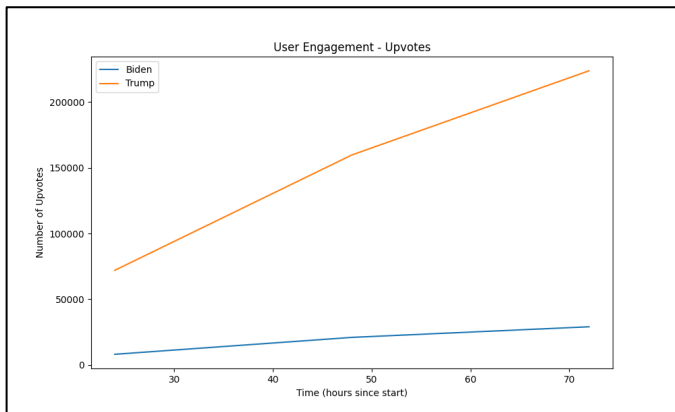


Fig. 9: Trump vs Biden Upvote Count over 3 days

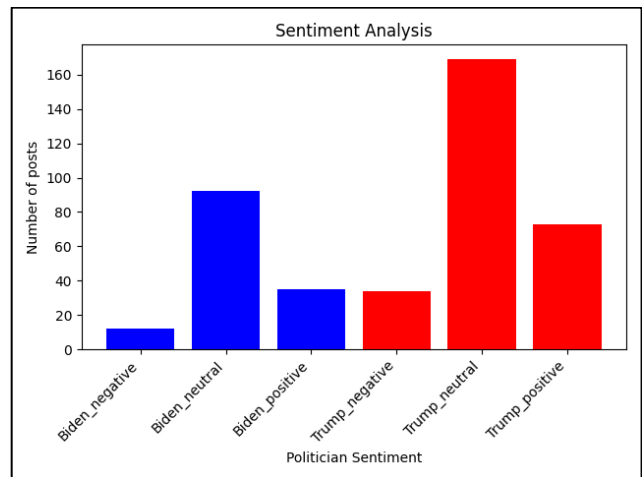


Fig. 12: Biden vs Trump Sentiment Count over 1 week

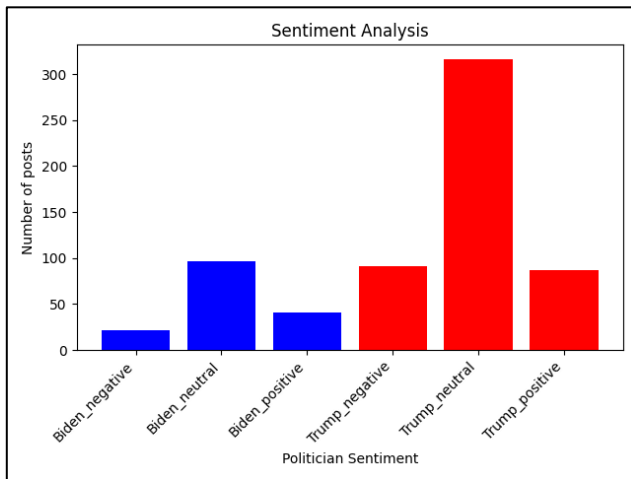


Fig.13: Biden vs Trump Sentiment Count over 2 weeks

VIII. REFERENCES

1. A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Wepel, "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment," ICWSM, 2010.
2. R. K. Garrett and P. A. Resnick, "Resisting Political Fragmentation on the Internet," *Daedalus*, vol. 142, no. 3, pp. 92–114, Summer 2013.
3. L. H. Nelson, "Theoretical Foundations of Sentiment Analysis," in *Handbook of Sentiment Analysis in Finance*, Springer, 2016, pp. 3–12.
4. R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Ng, and C. Potts, "Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2013, pp. 1631–1642.
5. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention Is All You Need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017, pp. 5998–6008.
6. J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1532–1543.
7. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
8. A. Karpathy, "The Unreasonable Effectiveness of Recurrent Neural Networks," blog post, May 2015. [Online]. Accessed: May 7, 2024.
9. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, 2019, pp. 4171–4186.
10. J. Kreps, N. Narkhede, and J. Rao, "Kafka: A Distributed Messaging System for Log Processing," in *Proceedings of the NetDB*, vol. 11, no. 2011, pp. 1–7, June 2011.
11. X. Meng, J. Bradley, B. Yavuz, E. Sparks, S. Venkataraman, D. Liu, et al., "Mllib: Machine Learning in Apache Spark," *Journal of Machine Learning Research*, vol. 17, no. 34, pp. 1–7, 2016.
12. M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support Vector Machines," *IEEE Intelligent Systems and Their Applications*, vol. 13, no. 4, pp. 18–28, 1998.
13. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
14. B. Boe, "PRAW: The Python Reddit API Wrapper," 2012–, [Online]. Available: <https://github.com/praw-dev/praw/>. Accessed: September 29, 2017.
15. M. Thelwall, "Social Network Sites and Politics: A Review of Research," *Journal of Computer-Mediated Communication*, vol. 17, no. 4, pp. 357–374, July 2012.