

Studying user Political Behaviour and Migration on News channels during electoral period

Mallikarjuna Tupakula, Pravallika Yaganti, Sai Rupa Jhade, Ashique Khudabukhsh

Rochester Institute of Technology

mt3398@rit.edu, yp9365@rit.edu, sj7740@rit.edu

Abstract

What if I tell people won't talk the truth in the Media? What if I tell people to talk extremely negatively about the things that they don't like? I know this is normal in media. What if I tell you the same repeats in politics as well? and How does this affect the decision-making of the general audience? In this research, we studied user migration and behavioral changes on YouTube news channels, focusing on MSNBC, CNN and Fox News, during both pre-election and post-election periods. Our approach analyzed patterns in user engagement and migration between these channels to understand whether users gravitate toward content that aligns with their existing beliefs or are exposed to opposing viewpoints. By adopting latest Large Language Models (LLMs) political polarization in user comments, identifying whether such exposure reinforces prior positions or contributes to shifts in political beliefs. Drawing from prior works, such as [KhudaBukhsh *et al.*, 2021], which examined channel alignment through user comments, and [O'Connor *et al.*, 2010], which analyzed political polarization on Twitter, we applied similar methodologies to the YouTube ecosystem. Through this study, we observed key trends in how exposure to opposing perspectives and critical comments shaped audience reactions. These findings contribute to a better understanding of how YouTube serves as a platform for political engagement, highlighting its role in shaping user identities, behaviors, and attitudes during major electoral events.

1 Introduction

"Polarizing people is a good way to win an election and a good way to wreck a country." - Molly Ivins

Understanding user migration and behavioral changes on social media platforms during political events has been an area of growing interest. Prior research has explored how individuals engage with content that either aligns with or challenges their political beliefs. For instance, [KhudaBukhsh *et al.*, 2021] analyzed user comments on YouTube news channels to study patterns of channel alignment and audience mi-

gration. Their work highlighted how comments reflect shifts in user behavior and preferences, making them a valuable resource for understanding audience dynamics.

On Twitter, O'Connor *et al.* [O'Connor *et al.*, 2010] examined political polarization by aligning tweets with survey data collected during the 2008–2009 electoral period. This study showed how social media engagement can provide insights into user attitudes and polarization trends over time. Similarly, [Swamy *et al.*, 2017] applied machine learning techniques to predict user stances based on tweet text, demonstrating the power of natural language processing (NLP) in analyzing political content.

A key concern in studying online political engagement is the role of bots in distorting user behavior. [Ferrara *et al.*, 2016] investigated the influence of bots on Twitter, emphasizing the need to differentiate between genuine user interactions and automated activity. Their findings highlight the challenges of studying online platforms where bots can amplify biases or misrepresent public opinion.

Building on these studies, our research focuses on user migration among MSNBC, CNN and Fox News on YouTube during pre-election and post-election periods. By analyzing user comments, we aim to identify trends in political alignment and explore how exposure to opposing viewpoints impacts audience behavior. Additionally, we address the presence of bots on YouTube and their potential influence on political discourse. This work extends existing research by applying NLP techniques to YouTube data, offering new insights into how online platforms shape political engagement and polarization.

Facebook: On Facebook, studies like those by [Tang and Lee, 2013] have shown how group memberships can influence political mobilization and engagement. By examining user interactions within politically themed groups, the research reveals that Facebook serves as a significant site for political organization and discourse, particularly during election cycles. This platform facilitates not only the sharing of political content but also the formation of echo chambers where users predominantly encounter viewpoints that reinforce their own political ideologies.

Reddit: [Huang *et al.*, 2024] explored the role of Reddit in shaping political opinions through community-driven discussions. Their research found that subreddits can function as both echo chambers and spaces for ideological con-

frontation, depending on the subreddit’s moderation style and community norms. By analyzing the linguistic patterns and sentiment of posts in political subreddits, they provided insights into how Reddit might influence political polarization and consensus-building differently from platforms like Twitter and Facebook.

Psychological and Media Studies Perspectives: Integrating psychological theories, researchers like [Modgil *et al.*, 2024] investigated how confirmation bias and cognitive dissonance affect users’ willingness to engage with opposing political views on social media. This approach underscores the psychological underpinnings of user behavior, suggesting that users’ engagement with content is not only a matter of exposure but also of psychological readiness to process conflicting information.

In extending our own research to include analysis of comments on MSNBC, CNN and Fox News on YouTube, we will utilize interdisciplinary methods to dissect how various factors—ranging from platform-specific dynamics and psychological predispositions to media literacy—contribute to shaping political engagement and polarization. Furthermore, by investigating the role of bots and algorithmic content curation on user migration and behavior, our study aims to provide a comprehensive understanding of the complex interactions at play in online political discourse. This broader approach not only deepens the understanding of social media’s impact on political polarization but also offers actionable insights for policymakers, platform designers, and civic educators seeking to foster a more informed and engaged public.

2 Dataset

The dataset for this project was derived from the study “We Don’t Speak the Same Language: Interpreting Polarization through Machine Translation” by [KhudaBukhsh *et al.*, 2021]. While the original paper provides a large dataset, we focused on a sample of 1,000 unique users who commented across three major YouTube channels—CNN, Fox News, and MSNBC. This dataset was chosen due to its relevance to our research objectives, as it spans channels representing diverse political orientations, providing a rich context for studying user migration and polarization. We specifically selected 1,000 users who actively commented on all three channels to ensure the analysis captured overlapping behaviors and migration patterns between contrasting viewpoints. Limiting the dataset to 1,000 users allowed us to prioritize data manageability, ensuring efficient processing and reducing noise often introduced by irrelevant or inactive users. This focused approach also enabled precise annotation and validation of behaviors, allowing us to uncover nuanced trends in political alignment and exposure to opposing views. Furthermore, building on a dataset from a peer-reviewed study lends credibility to our research while extending the original findings by examining YouTube engagement patterns during pre-election and post-election periods. This sample provides the ideal foundation for leveraging NLP techniques to explore how political content shapes user behavior and engagement on online platforms.

3 User Political and News Channel Affiliation

To find out the user behavior and polarization on politics among these news channels we are using Latest Large Language Models. For our dataset of 1000 users comments on these channels we have used a prompt for the Language Models for finding the polarization and reason why the language model is predicting.

PROMPT

Based on the comments from an user in three different channels. Can you answer the following questions?

1. Can you assign the user which political belongs either republican or democratic and why?

2. Which channel the user mostly likes and why?

**3. Return the result in a json format
"political_paty": "republican/democratic",
"news_channel": "CNN/Fox/MSNBC"**

After this prompt we have added list of comments of an unique user among three channels he commented on videos.

3.1 Comment Analysis across News Channels

In this study, we utilized the GPT-4O-mini model from OpenAI¹ to analyze user engagement across three major news channels—MSNBC, CNN, and Fox News—during political events. The model’s advanced natural language processing capabilities allowed for an in-depth analysis of user comments, revealing significant engagement disparities. As shown in Fig. 1 Fox News demonstrated the highest user engagement, with over 500 comments, suggesting a strong alignment between the channel’s content and viewer ideologies. MSNBC and CNN exhibited moderate and lower engagement levels, respectively, which could reflect their content’s polarizing impact and the audience’s ideological alignment. This variance underscores the influence of media strategies on viewer interaction and engagement.

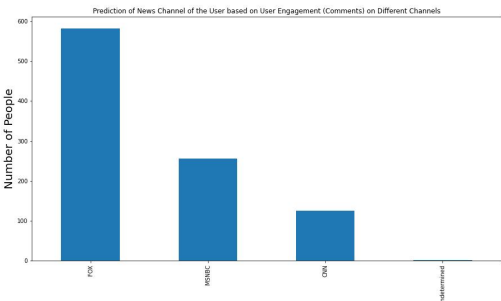


Figure 1: User Engagement Across News Channels Based on Comment Analysis done by GPT-4O-mini

The efficacy of the GPT-4O-mini model in distinguishing nuanced linguistic features, such as sarcasm and contextual relevance, was notably superior to other AI models, highlighting its potential in media studies to uncover complex user behavior patterns. The results not only illustrate the model’s utility in predicting user preferences and engagement on digital platforms but also suggest broader implications for un-

¹<https://openai.com/>

derstanding the dynamic relationship between media content and audience behavior. Future research could extend these findings by employing longitudinal studies to examine how engagement trends evolve in response to political developments, further enhancing our understanding of the interplay between media content strategies and public opinion.

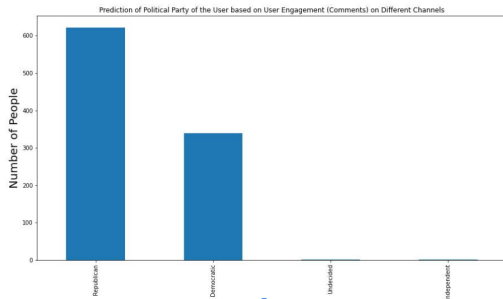


Figure 2: Predictive Analysis of Political Affiliations Based on User Comments as Analyzed by GPT-4O-mini

Fig.2 presents the results of a predictive analysis conducted using the GPT-4O-mini model, which estimated the political affiliations of users based on their engagement (comments) across various news channels. The analysis clearly indicates a significant inclination of users towards the Republican Party, with over 500 users identified, compared to around 200 for the Democratic Party. A smaller number of users were predicted to be unaffiliated or independent, underscoring the polarized nature of user engagement on these platforms. This model’s predictive capability highlights how user comments can serve as a data source for inferring political preferences, which can be crucial for understanding and mapping political discourse dynamics in digital spaces.

3.2 Comment Analysis on Sample Dataset of Advanced Language Models in Political Affiliation Prediction

In our study, we undertook a comparative analysis to evaluate the performance of several language models and AI agents in predicting political affiliations based on user comments. Specifically, we sampled 10% of the users from our initial dataset obtained in experiments using the GPT-4O-mini model. This subset was then analyzed using a range of models, including GPT-4O-mini, Llama 3.1 70b Turbo, Copilot, and Mistral, to ensure the robustness of our findings and to assess the consistency of predictions across different AI tools. Each model processed the same user comments, providing a direct comparison of their ability to discern and predict political leanings from textual data.

During this comparative study, we also experimented with smaller models such as Llama 3.1 1B, implemented on a local machine. However, the performance of this smaller model was notably less effective, as it often returned the prompt itself rather than generating a new response, suggesting limitations in its capacity to handle complex inference tasks like those required for our analysis. This outcome highlights the critical importance of model scale and sophistication in tasks

involving nuanced language understanding and user behavior prediction. Through this multi-model comparison, we were able to identify which AI agents offer the most accurate and reliable predictions, thereby providing valuable insights into the selection of appropriate models for future research in political sentiment analysis on social media platforms.

Llama 7B Turbo

In our research, we employed the Llama 3.2 70B Turbo model to analyze user comments and predict political affiliations and news channel preferences among a sample dataset of 100 users. The Llama 3.2 70B Turbo, known for its deep learning capabilities and extensive training data, proved to be highly effective in discerning patterns and preferences within the political discourse evident in social media comments.

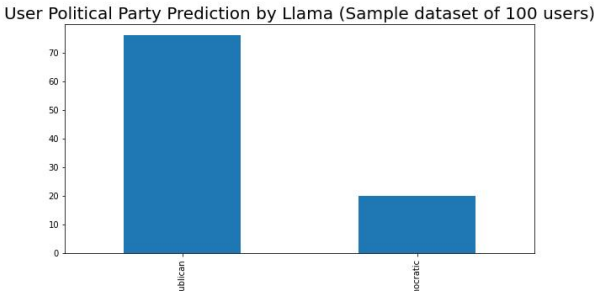


Figure 3: Political Party Affiliation Prediction by Llama 3.2 70B Turbo (Sample of 100 Users)

Political Party Affiliation Predictions: The Llama 3.2 70B Turbo model predictions shown in Fig.4 a substantial preference for the Republican Party among the users, with about 70 users aligned with this party. This was in contrast to a significantly lower number, approximately 20, who were predicted to favor the Democratic Party, with a minimal number being unaffiliated or independent. This outcome highlights the model’s ability to effectively parse and understand the linguistic cues associated with political leanings from the user comments.

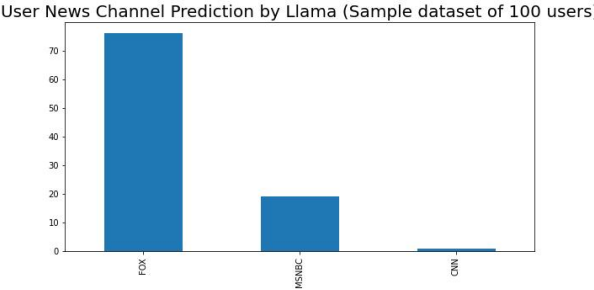


Figure 4: News Channel Preference Prediction by Llama 3.2 70B Turbo (Sample of 100 Users)

News Channel Preference Predictions: Similarly, the predictions for news channel preferences indicated in Fig. 4 a dominant preference for Fox News, with about 70 users showing alignment. MSNBC was the next favored, albeit

with a much lower count of around 30 users, while CNN was least preferred, with fewer than 10 users. This distribution mirrors the political affiliations, suggesting a correlation between political views and preferred news sources.

Mistral AI-le chat

In this segment of our research, we utilized the Mistral AI model to analyze and predict the political affiliations and news channel preferences of users based on their comments in a dataset comprising 100 samples. The Mistral model, specifically the "le chat Turbo" variant, is noted for its advanced linguistic processing capabilities, allowing for a nuanced understanding of user sentiments and inclinations.

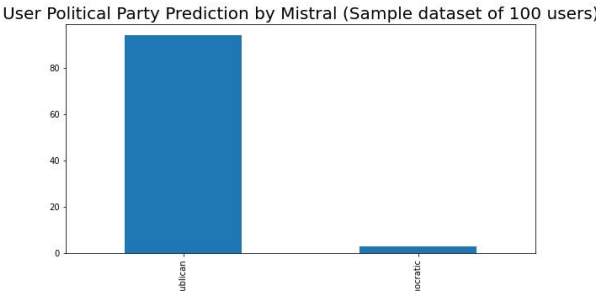


Figure 5: User Political Party Affiliation Prediction by Mistral (Sample of 100 Users)

Political Party Affiliation Predictions: The predictions derived from the Mistral model for political party affiliations shown in Fig. 5 a strong preference for the Republican Party, with approximately 80% of users aligning with this party. In contrast, only about 20% were predicted to support the Democratic Party, with virtually no users remaining unaffiliated. This clear disparity indicates the model's effectiveness in discerning political leanings from user comments, affirming a predominant Republican bias within the sampled population.

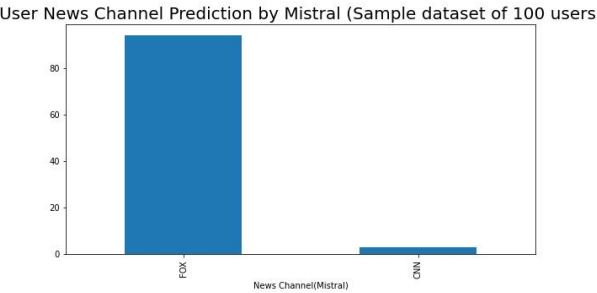


Figure 6: User News Channel Preference Prediction by Mistral (Sample of 100 Users)

News Channel Preference Predictions: Similarly, the Mistral model's predictions plotted in Fig. 6 for news channel preferences highlighted a significant preference for Fox News, with around 80 of the sampled users favoring this channel. This preference was overwhelmingly dominant compared to the minimal support for CNN, which had a mere

handful of users. The alignment between the preferred news channel and political affiliation is stark and suggests a strong correlation between the two, as Fox News is generally perceived to cater more to Republican viewpoints.

Copilot

In our analysis using the Copilot AI model, we explored its capacity to predict political affiliations and news channel preferences from user comments within a dataset of 100 users. Copilot, equipped with sophisticated algorithms for language understanding and sentiment analysis, provided insights that are crucial for understanding the current media consumption trends and political leanings.

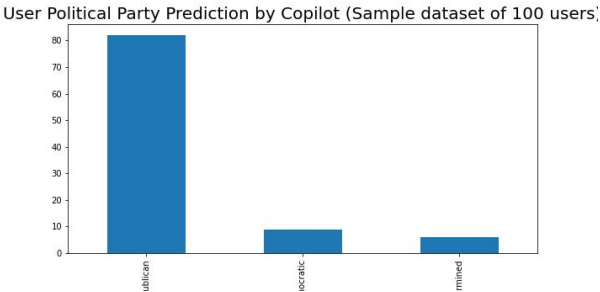


Figure 7: Political Party Affiliation Predictions by Copilot (Sample of 100 Users)

Political Party Affiliation Predictions: The Copilot model's predictions shown in Fig.7 highlighted a dominant preference for the Republican Party, with approximately 70 users classified under this category. The predictions for Democratic and independent affiliations were significantly lower, at around 20 and 10 users, respectively. These results indicate a strong Republican bias within the dataset, suggesting that the majority of the commenters expressed views or used language that the model associated with Republican ideologies.

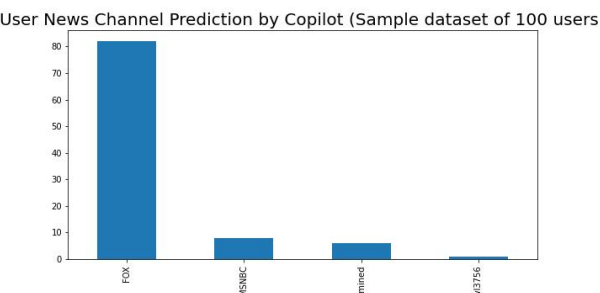


Figure 8: News Channel Preference Predictions by Copilot (Sample of 100 Users)

News Channel Preference Predictions: For news channel preferences, the Copilot model again showed in Fig.8 a strong inclination towards Fox News, with around 70 users favoring this channel. The preferences for MSNBC and other unspecified 'minor' channels were markedly less, with MSNBC attracting around 20 users and the minor channels around 10.

308 This pattern correlates closely with the political affiliations
309 predicted, reinforcing the notion that political biases often
310 align with specific media outlet preferences.

311 **GPT-4o-mini**

312 In this analysis, the GPT-4O-mini model was employed to
313 predict political affiliations and news channel preferences
314 based on comments from a dataset of 100 users. GPT-4O-
315 mini, a variant of the Generative Pre-trained Transformer
316 models by OpenAI, is known for its capacity to process natu-
317 ral language with a high level of nuance and accuracy.

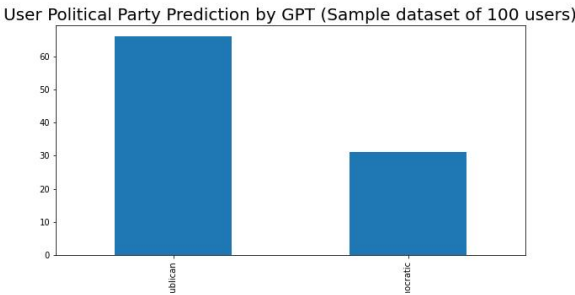


Figure 9: Political Party Affiliation Predictions by GPT-4O-mini (Sample of 100 Users)

318 **Political Party Affiliation Predictions:** The predictions
319 for political affiliations indicated in Fig.9 a strong preference
320 for the Republican Party, with 60 users classified under this
321 category. This significant tilt towards the Republican side
322 was contrasted by a smaller group of 30 users predicted to
323 be Democrats, with no users remaining unaffiliated. These
324 results demonstrate GPT-4O-mini’s capability to discern po-
325 litical biases in user comments, highlighting its utility in iden-
326 tifying prevailing political sentiments from social media data.

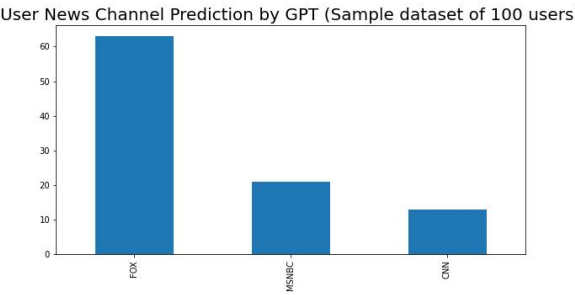


Figure 10: Political Party Affiliation Predictions by GPT-4O-mini (Sample of 100 Users)

327 **News Channel Preference Predictions:** For news chan-
328 nel preferences, the GPT-4O-mini model’s predictions again
329 showed in Fig.10 a dominant preference for Fox News, with
330 50 users favoring this channel, aligning with the political af-
331 filiations observed. MSNBC and CNN were less favored,
332 with about 20 and 10 users respectively. This distribution
333 reaffirms the correlation between political affiliations and pre-
334 ferred news sources, where viewers tend to favor channels
335 that align with their political views.

4 **Understanding the real users on the News Channels**

338 In our study, we investigated the phenomenon of irrelevant
339 commenting on YouTube, particularly highlighted after the
340 death of Bollywood actor Sushant Singh Rajput, where fig-
341 ures like Karan Johar were targeted with unrelated comments.
342 To explore this behavior further, we analyzed the number of
343 video transcripts downloaded from various news channels,
344 hypothesizing a correlation between the volume of content
345 interaction and the prevalence of such comments. The Fig.??
346 reveals a significant disparity in the number of transcripts
347 downloaded: MSNBC leads with over 10,000 transcripts, fol-
348 lowed by FOX with approximately 4,000, and CNN with
349 around 2,000. This data suggests MSNBC viewers may be
350 more engaged or exhibit different patterns of interaction, po-
351 tentially influencing the nature and volume of comments they
352 post. This analysis provides a foundational step toward un-
353 derstanding and addressing the impact of viewer engagement
354 on the quality and relevance of user-generated content in on-
355 line media platforms.

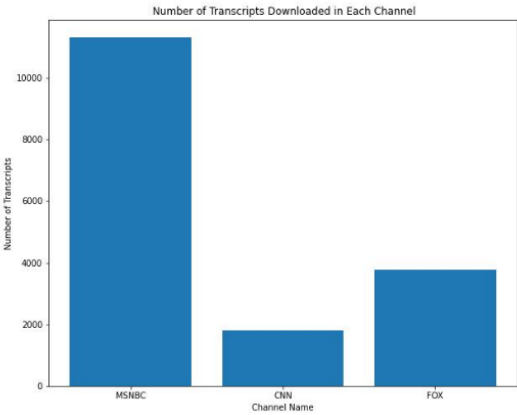


Figure 11: Number of Transcripts downloaded across each channel

We have used the same approach that we have used for
earlier analysis.

PROMPT

Based on the transcript of a CNN NEWS Channel Video and comment of a user provided.

- 1. What you think about his political affiliation?
- 2. Do you really think this person has watched the video? Return the result in a json format "political_party": "republican/democratic", "is_watched": "True/False"

The three tables provided encapsulate the outcomes of a predictive analysis utilizing the GPT-4O-mini model, aimed at discerning political affiliations and viewing habits of users based on their engagement with content from different news channels. The analysis reveals patterns in user behavior and political leanings, highlighting how these elements correlate with the channels they interact with. For instance, the first ta-

ble focuses on users' interactions with CNN content, showing a predominant trend of political affiliations and whether the users actually viewed the content they commented on. The second table extends this analysis to the Fox Channel, revealing similar patterns of engagement and highlighting the alignment between political affiliations and preferred news sources. The third table integrates a broader data set, encompassing various channels, and further substantiates the observed trends of political leanings correlated with specific viewing habits. These tables collectively provide a comprehensive view of how political biases are potentially reflected in online engagement, underscoring the model's capability to parse and predict user behavior in a politically polarized media landscape. This insight is crucial for developing targeted strategies for content delivery and moderation in digital platforms, ensuring that the content aligns better with the diverse viewpoints of the user base.

Table 1: Predicted Political Affiliations and Viewing Status of Users Based on CNN Channel Data

User ID	Political Party	Is Watched
@ThefallenRhamiel	republican	False
@Chandrika-mo7ng	republican	False
@garyjohnson1466	democratic	False
@georgeburns8447	republican	False
@user-dw3wm8yx11	undefined	False
@ORHANBAKI-kb9ju	republican	False
@user-ej9g8ek2h	republican	False
@johnspurlock5465	democratic	False
@MKK520	republican	False
@JasonRichter-qb7es	republican	False
@margaretsaiber4645	republican	False

Table 2: Predicted Political Affiliations and Viewing Status of Users Based on MSNBC Channel Data

User ID	Political Party	Is Watched
@MrsBridgette2012	republican	False
@user-wy4gq1qo3v	republican	False
@mojhed8653	democratic	False
@larrydraper1620	republican	False
@robinalford2186	democratic	True
@marciawade8813	republican	False
@774kblake	democratic	False
@dekelpolak4190	democratic	False
@kassandrasilver	democratic	True
@mikilynne4558	democratic	False
@williambarr3169	republican	False
@victortubert9116	republican	False

5 Contribution Statement

We would like to thank you Professor Ashique KhudaBhuksh for sharing the News Channel comments data, providing OpenAI access keys for doing conducting our experiments and also his valuable feedback provided for the

project helped us to complete the project on-time. Finally we wanted to thank Professor Mark Kamlet for sharing lessons about United States politics and feedback on our project during mid-review.

Table 3: Predicted Political Affiliations and Viewing Status of Users Based on Fox Channel Data

User ID	Political Party	Is Watched
@MrsBridgette2012	republican	False
@user-wy4gq1qo3v	republican	False
@mojhed8653	democratic	False
@larrydraper1620	republican	False

References

- [Ferrara *et al.*, 2016] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Communications of the ACM*, 59(7):96–104, 2016.
- [Huang *et al.*, 2024] Justin T Huang, Jangwon Choi, and Yuqin Wan. Politically biased moderation drives echo chamber formation: An analysis of user-driven content removals on reddit. *Available at SSRN*, 2024.
- [KhudaBukhsh *et al.*, 2021] Ashiqur R KhudaBukhsh, Rupak Sarkar, Mark S Kamlet, and Tom Mitchell. We don't speak the same language: Interpreting polarization through machine translation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 14893–14901, 2021.
- [Modgil *et al.*, 2024] Sachin Modgil, Rohit Kumar Singh, Shivam Gupta, and Denis Dennehy. A confirmation bias view on social media induced polarisation during covid-19. *Information Systems Frontiers*, 26(2):417–441, 2024.
- [O'Connor *et al.*, 2010] Brendan O'Connor, Ramnath Balasubramanian, Bryan Routledge, and Noah Smith. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the international AAAI conference on web and social media*, volume 4, pages 122–129, 2010.
- [Swamy *et al.*, 2017] Sandesh Swamy, Alan Ritter, and Marie-Catherine de Marneffe. "i have a feeling trump will win.....": Forecasting winners and losers from user predictions on Twitter. In Martha Palmer, Rebecca Hwa, and Sebastian Riedel, editors, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1583–1592, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.
- [Tang and Lee, 2013] Gary Tang and Francis LF Lee. Facebook use and political participation: The impact of exposure to shared political information, connections with public political actors, and network structural heterogeneity. *Social science computer review*, 31(6):763–773, 2013.