Opinion Polarization on Social Networks based on Political Discourse

Susmita Das
SCSET, Bennett University
Greater Noida, India
susmitad900@gmail.com

Sounak Sadhukhan
SCSET, Bennett University
Greater Noida, India
sounaks.cse@gmail.com

Arunabha Tarafdar
SCSET, Bennett University
Greater Noida, India
arunabhatarafdar@gmail.com

***Abstract*—Social media platforms have become the nerve center of campaign discussion in regards to global political landscape. Political candidates are wielding social media as an influential instrument to advance their own election campaigns. In this paper, we have perused the online exchanges and conversations among supporters of different political parties in the intense environment of US Presidential Election 2024. Sentiment scores, abusive speech and stance detection of the tweets have been considered for understanding the perspective of voters on social media platforms. Each post has been considered whether it is a hate speech which contributes in polarization in the various conversation obtained from mainstream social media platform X/Twitter. A novel model has been proposed that factors in sentiment, stance and and hate speech for opinion dynamics estimation. Our method has been evaluated on various publicly available datasets and satisfactory results have been obtained. It is observed that there is higher level of opinion polarization when hate speech is involved.**

***Index Terms*—Social Networks, stance Detection, Sentiment Analysis, Hate Speech**

I. Introduction

Engagement on conversations regarding the global political landscape has gained popularity on mainstream social media platforms. As a consequence, social media platforms are being used as a tool by multiple political parties across the globe for engagement. Users with different ideologies strive to establish the principles of their respective political parties. The increase in communication between users of like-mind and acquiring information adapting to their own ideologies gives social media rise to filter bubbles and echo chambers. There have been multiple areas of disagreement between researchers regarding the effect of online campaigns on the actual voting results, but the role of social networks cannot be refuted even by the political candidates. Sentiment analysis has been used extensively to gauge the opinions of voters [1] [2] on the political landscape. Topic modeling and opinion mining have been used for understanding the concerns of general citizens [3]. Multiple combinations of these techniques have been implemented for predicting the outcome of various elections [4] [5]. Sentiment and stance of a person can provide insights into the intent behind a certain political opinion. In respect to U.S. Presidential Election 2024, the social media posts majorly revolved around the rivalry between the top contenders Kamala Harris and Donald Trump. The social media posts from the point of stance can be majorly categorized as: a) *Support for The Democratic Party(USA) and Kamala Harris*, b) *Support for The Republican Party(USA) and Donald Trump*, c) *Support for third party* and d) *Neutral point-of-view*. In perspective of sentiment analysis, these posts can be organized as: a) *Positive sentiment towards The Democrats*, b) *Positive sentiment towards The Republicans*, c) *Negative sentiment towards Donald Trump and The Republican Party*, d) *Negative sentiment towards Kamala Harris and The Democratic Party*, e) *Negative sentiment towards both the Democrats and Republicans* and f) *Neutral standpoint*. Our principal contributions in this paper are: 1) Stance detection as well as sentiment analysis of the social media posts have been considered for analysing the perceptions of the online users. 2) Hate Speech has been considered for calculation of opinion dynamics and polarisation calculation. The research paper has been organised with literature review in *Section II*, proposed work discussed in *Section III* with dataset details, preparation and opinion dynamic scoring elaborated in subsections. Results have been encapsulated in *Section IV* and lastly conclusion is in *Section V*.

II. Literature Review

The online discourse regarding political elections has garnered interest in field of social network

TENCON 2025 - 2025 IEEE Region 10 Conference (TENCON) | 979-8-3315-3772-2/25/$31.00 ©2025 IEEE | DOI: 10.1109/TENCON66050.2025.11375442

analysis. Users with varying opinions have arguments and often engage in vilification of the election candidates, presenting an overview regarding the homophily in the social networks. Stance detection has been implemented using LSTM model [6]. Kuo et al. [7] studied integration of visual and textual features for stance classification and carried out on various posts on political fanpages regarding 2021 national referendums in Taiwan. Zhang et al. [8] proposed Bipartite Graph Neural Networks for stance detection based on 2020 US Presidential elections. Mets et al. [9] researches the efficiency of large language models(LLMs) for stance detection in complex policies and lower-resource language. Gul et al. [10] presents stance detection using advanced models LLaMa-2 and Mistral-7B. Chaudhry et al. [11] studied the shift of online sentiment before and after the 2020 US Presidential elections. Xia et al. [12] analyses the positive and negative sentiment ratioes related to Biden and Trump in US Presidential elections 2020. Rita et al. [13] researched the Twitter sentiment for election results in UK general elections 2019. Chauhan et al. [14] studied the collective sentiment and behaviour of voters in 2019 Indian general (Loksabha) elections. Ma et al. [15] explored the sentiment using multi-label classification of online review for Indonesian election 2024. Attai et al. [16] studied sentiment and public opinion on the candidates for Nigerian general elections 2023.

III. PROPOSED WORK

Each user in social network is represented by a node and each node holds a particular stance and sentiment on a specific topic. When a new node integrates in the network, it is influenced by the opinions of its neighboring nodes. The content of the post, whether it is hate speech or not, further fuels the conversation and arguments online.

A. Stance Determination

Stance is defined as the political standpoint and support of the user writing the tweet. Stance differs from sentiment in that a person may express anger through their text but still be supporting a particular ideology. The stance of the tweets is categorized as 'FAVOR' for those supporting the ideology, political party or candidate, 'AGAINST' for those opposing it, and 'NONE' for tweets with a neutral perspective [6].

B. Sentiment Analysis

Political posts on social media disclose intense sentiments that assists in comprehending the perception of users regarding a political campaign. VADER sentiment analysis has been implemented for sentiment extraction from the posts on Twitter. Measures of sentiment scores of a few sample tweets is presented in Table I.

TABLE I: Sentiment Analysis of Political Tweets

It's time, for the toxic-Right to fuck us...	{'neg': 0.241, 'neu': 0.697, 'pos': 0.062, 'compound': -0.7685}
Let's bring back freedom of speech...	{'neg': 0.0, 'neu': 0.682, 'pos': 0.318, 'compound': 0.6369}
Kamala is the warmonger, Trump...	{'neg': 0.305, 'neu': 0.538, 'pos': 0.157, 'compound': -0.5106}
Her stupid MAGA ass was senten...	{'neg': 0.438, 'neu': 0.465, 'pos': 0.097, 'compound': -0.8316}
I've seem more Republicans on ...	{'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}

C. Dataset Details

People have different opinions on social media platforms which reflect their perspective on a certain issue and this information can help by researchers and strategists to understand the public opinions on the political landscape.

1) Dataset Description: We have used the extensive tweet dataset presented by Balasubramanian et al. [17]. This dataset consists of tweets related to the U.S. Presidential 2024 campaign between the time period May 2024 to November 2024. Significant keywords obtained from the crawled tweet dataset includes the presidential candidates for 2024 elections: *Trump*, *Kamala*, *Biden*, *president*, *MAGA (Make America Great Again)*, *GOP (Grand Old Party)* etc. The total dataset consists of 834 chunks of tweet data with $50k$ tweets for each chunk.

TABLE II: Sample Tweets with the Annotation

Tweet Text	Republican	Democrat	Sentiment	Hate
Dems uses military people for..	FAVOR	AGAINST	-0.78	-1
Kamala love your take on..	NONE	FAVOR	0.54	1
Okay white women voting sexist...	AGAINST	NONE	-0.65	-1

2) Preparation of Dataset: The extensive tweet dataset has been shortened for our paper. Selection of approximately $1k$ tweets randomly from each chunk has been conducted, accounting to a total of about $835k$ tweets aprroximately. The features of the dataset taken into account are *tweet id*, *tweet text* and *retweetCount*. We annotated every tweet according to their stance according to the political party, perform binary classification where each tweet has been annotated whether it is hate speech. When a tweet is hate speech, it is denoted by -1 and non-hateful speech is symbolized by $+1$. The sentiment of the tweet determined using VADER is included for each tweet. Sample of the annotation process has been displayed in Table II.

D. Opinion Dynamics

In social networks, users are affected by opinions of neighbouring nodes. The sentiment and stance of online posts depicts the perceptions of general public. In a charged political environment, hate speech also advances the cause for opinion dynamics. A certain post being retweeted by multiple online users alludes to approval by these users. For opinion change calculation, we have expanded on the Friedkin-Johnsen model [18] as following:

$$\Theta_i^t = \Theta_i^{t-1} + \frac{\varsigma_i^t * \mathcal{R}_i + H_p}{1 + (\mathcal{R}_i * |\varsigma_i^t|)} \quad (1)$$

where ς_i^t denotes the sentiment of i^{th} node at timestamp t, $\mathcal{R}_i$ represents the number of retweets of post by i, H_p denotes whether the post is considered a hate speech and Θ_i^t represents the opinion of i^{th} node at t timestamp. In case of social networks, opinion of the node neighbourhood is also significant. The above equation 1 can be modified accordingly as follows:

$$\Theta_i^t = \Theta_i^{t-1} + \frac{\sum_{\eta(i)} \varsigma_j^t * \mathcal{R}_i + H_p}{1 + (\mathcal{R}_i * \sum_{\eta(i)} |\varsigma_j^t|)} \quad (2)$$

where $\eta(i)$ represents the neighbourhood of a particular node and ς_j denotes the sentiment of the nodes present in the neighbourhood of node i.

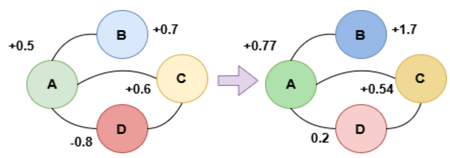

Fig. 1: Change of Opinions in a Social Network

As depicted in Fig. 1, the sample network shows the interaction between different nodes and their modification of opinion. The initial opinion of node A at timestamp t_0 is assumed to be $+0.5$ and after interaction with nodes B, C and D, A's opinion is updated to $+0.77$ in t_1. In same pattern, after interaction with nodes A and C, the opinion of D gets updated from -0.8 during t_0 to $+0.2$ during t_1.

Property 1: When the social network formed is from the interactions, the opinion of pendent vertex(if present) is completely influenced by the opinion of the node to which it is connected.

For a pendent vertex, the node neighbourhood $(\eta) = 1$, as only one node is present in its node neighbourhood. As a result, the opinion of pendent vertex is solely amplified by the connected node's opinion.

Property 2: In case of social network, the update of opinion dynamics Θ_i adheres to a linear equation f. Consider a function f that is continuous on [a,b] and $g(z)$ function is the convergence rate to a fixed point c. Then, f will converge linearly, in case of $g(c) \neq 0$.

E. Opinion Polarization

The variation in opinions is appraised to ascertain the comprehensive view of polarization for a particular set of users. The opinion polarization is calculated as follows:

$$\rho = \sum_{i=1}^{N} (\Theta_i - \Theta_{mean})^2 \quad (3)$$

where the total number of users considered for the study is denoted by N. Θ_{mean} is computed by the formula $\Theta_{mean} = \frac{1}{N}\sum_{j=1}^{N} \Theta_j$. The ρ value varies between $[0, 1]$, where values closer to 1 depicts high polarization and values adjacent to 0 denotes less polarization.

F. Illustration with Example

Consider a node A with an initial opinion at t_0, ie. $\Theta_A^{t0} = 0.7$. In timestamp $t = t_1$, the post by node A is a non-hate post and carries a negative sentiment -0.8, which is retweeted 10 times. So, for this timestamp t_1, $\varsigma_A^{t1} = -0.8$, $\mathcal{R} = 10$ and $H_p = 1$. Considering these values, we can calculate the opinion vector of A as follows:

$$\Theta_A^{t1} = 0.7 + \frac{-0.8 * 10 + 1}{1 + (10 * 0.8)} = -0.077$$

Suppose in the following timestamp $t = t_2$, user A uploads another non-hate post having positive sentiment 0.6 which is retweeted 11 times. In this case, $\varsigma_A^{t2} = 0.6$, $\mathcal{R} = 11$, $H_p = 1$ and $\varsigma_A^{t1} = -0.077$. The opinion vector for node A in timestamp t_2 is as follows:

$$\Theta_A^{t2} = -0.077 + \frac{0.6 * 11 + 1}{1 + (11 * 0.6)} = 0.923$$

Suppose at timestamp t_1, the opinions of certain users are obtained: $\Theta_A^{t1} = -0.077$, $\Theta_B^{t1} = 0.312$, $\Theta_C^{t1} = -0.18$ and $\Theta_D^{t1} = 0.47$. So, opinion mean at timestamp t_1, $\Theta_{mean}^{t1} = 0.13125$ and polarization $\rho^{t1} = (-0.077 - 0.13125)^2 + ... + (0.47 - 0.13125)^2 = 0.2$.

IV. Results and Analysis

We have employed social networks to study the communications and conversations among online users regarding U.S. Presidential Elections 2024.

A. Polarization based on Tweet Timestamp

The election dataset [17] has been collected based on different timestamps throughout the duration of the election campaign starting from May 2024. We have determined the polarization for each month, that has been displayed in Table III. In

TABLE III: Polarization for each Month based on AGAINST Stance

Month/ Timestamp	Polarization	
	AGAINST Democrats	AGAINST Republicans
May	0.307	0.48
June	0.561	0.415
July	0.743	0.672
August	0.68	0.643
September	0.702	0.726
October	0.82	0.798
November	0.79	0.8

the timeline of May 2024 to November 2024, we observe that even though there has been certain variations in polarization, it has increased quite steadily as the election dates have approached. Certain events such as *cognitive decline of Joe Biden* during June-July 2024, *indictments against Donald Trump* during May-June 2024 etc. have influenced this polarization. As depicted in Table IV, the

TABLE IV: Polarization for each Month based on FAVOR Stance

Month/ Timestamp	Polarization	
	FAVOR Democrats	FAVOR Republicans
May	0.18	0.2
June	0.231	0.22
July	0.45	0.507
August	0.347	0.401
September	0.42	0.409
October	0.428	0.47
November	0.41	0.42

polarization is much less, as positive sentiment and positive stance often account for the lower opinion deviation dynamics. Polarization is also very high over the various months when we consider hate speech as shown in Fig. 2. The overall opinion

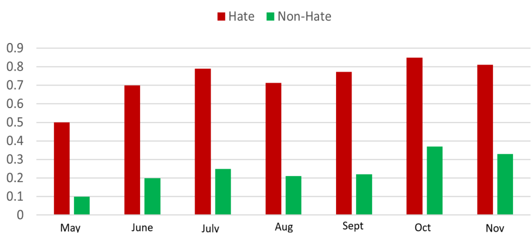

Fig. 2: Polarization based on Hate Speech

polarization over the whole timestamp of US Election campaign starting from May 2024 has been depicted in Fig. 3.

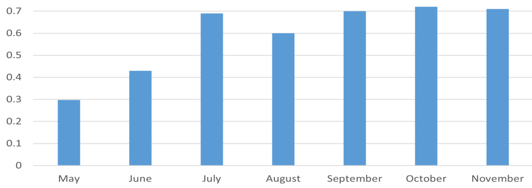

Fig. 3: Polarization in US Election 2024

B. Polarization based on Retweet Networks

The retweets for particular tweets have been obtained using Twitter/X API. As shown in Fig. 4, based on stance, interactions between users are quite clustered. In case of retweet networks of particular tweets, people with *AGAINST* stance, *NEUTRAL* stance and stance in *FAVOR* mostly exchange opinions among their groups. We formed

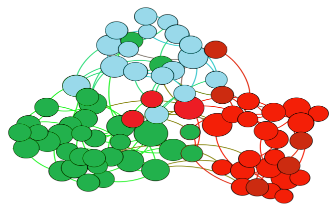

Fig. 4: Polarization in Sample Retweet Network

retweet networks for the tweets during *October* timestamp. Polarization for retweet networks in *FAVOR* of Democrats is 0.38, in *FAVOR* of Republicans is 0.42. Polarization *AGAINST* Democrats is observed to be around 0.786 and *AGAINST* Republicans to be around 0.8.

C. Analysis on Other Political Datasets

Various datasets based on different political events have been compiled and made publicly available. Chris Albon [19] presented Twitter dataset on the election day of US Presidential Election 2016 consisting of $6.5M$ tweets posted on that particular day. The results have been compiled in Table V. Grimminger et al. [20] have

TABLE V: Polarization on US Presidential Election Day 2016 Dataset

	AGAINST Trump	AGAINST Hillary	FAVOR Trump	FAVOR Hillary
Polarization	0.582	0.624	0.37	0.29

	HATE	NON-HATE
Polarization	0.65	0.19

compiled a dataset consisting of tweets based on 2020 US Presidential election combining both stance and hate speech. It contains the columns

TABLE VI: Polarization on US Presidential Election 2020 Dataset

	AGAINST Trump	AGAINST Biden	FAVOR Trump	FAVOR Biden
Polarization	0.69	0.653	0.45	0.32

	HATE	NON-HATE
Polarization	0.712	0.21

'Tweet Text', Stance towards candidates *'Donald'*, *'Biden'*, *'West'* in terms of *FAVOR*, *AGAINST* and *NEITHER*. As shown in Table VI, we can observe that political polarization have increased over time.

V. Conclusion

We have proposed our procedure for measuring the dynamics of opinion on social media. Stance, sentiment measurements, and hate speech have been factored in for formulation of the proposed approach. The prevailing opinion dynamics model of Friedkin-Johnsen has been extended and implemented to understand opinion polarization. Our proposed method has been applied to multiple political datasets and has yielded satisfactory results. We have studied the degree of polarization observed in an online political landscape. The analysis correlates to the observation that opposing stances and abusive language are significantly influential factors in creating opinion polarization. Future scopes include comprehension of opinion heterogeneity and echo chamber formation and their effect on opinion polarization over time.

References

[1] P. Chauhan, N. Sharma, and G. Sikka, "The emergence of social media data and sentiment analysis in election prediction," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 2601–2627, 2021.

[2] M. Rodríguez-Ibáñez, F.-J. Gimeno-Blanes, P. M. Cuenca-Jiménez, C. Soguero-Ruiz, and J. L. Rojo-Álvarez, "Sentiment analysis of political tweets from the 2019 spanish elections," *IEEE Access*, vol. 9, pp. 101 847–101 862, 2021.

[3] Y. Matalon, O. Magdaci, A. Almozlino, and D. Yamin, "Using sentiment analysis to predict opinion inversion in tweets of political communication," *Scientific reports*, vol. 11, no. 1, p. 7250, 2021.

[4] K. Kawintiranon and L. Singh, "Knowledge enhanced masked language model for stance detection," in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Online: Association for Computational Linguistics, Jun. 2021, pp. 4725–4735. [Online]. Available: https://aclanthology.org/2021.naacl-main.376

[5] P. Gunhal, A. Bashyam, K. Zhang, A. Koster, J. Huang, N. Haresh, R. Singh, and M. Lutz, "Stance detection of political tweets with transformer architectures," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2022, pp. 658–663.

[6] K. Dey, R. Shrivastava, and S. Kaushik, "Topical stance detection for twitter: A two-phase lstm model using attention," in *European Conference on Information Retrieval*. Springer, 2018, pp. 529–536.

[7] K.-H. Kuo, M.-H. Wang, H.-Y. Kao, and Y.-C. Dai, "Advancing stance detection of political fan pages: A multimodal approach," in *Companion Proceedings of the ACM Web Conference 2024*, 2024, pp. 702–705.

[8] C. Zhang, Z. Zhou, X. Peng, and K. Xu, "Doubleh: Twitter user stance detection via bipartite graph neural networks," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 18, 2024, pp. 1766–1778.

[9] M. Mets, A. Karjus, I. Ibrus, and M. Schich, "Automated stance detection in complex topics and small languages: the challenging case of immigration in polarizing news media," *Plos one*, vol. 19, no. 4, p. e0302380, 2024.

[10] I. Gül, R. Lebret, and K. Aberer, "Stance detection on social media with fine-tuned large language models," *arXiv preprint arXiv:2404.12171*, 2024.

[11] H. N. Chaudhry, Y. Javed, F. Kulsoom, Z. Mehmood, Z. I. Khan, U. Shoaib, and S. H. Janjua, "Sentiment analysis of before and after elections: Twitter data of us election 2020," *Electronics*, vol. 10, no. 17, p. 2082, 2021.

[12] E. Xia, H. Yue, and H. Liu, "Tweet sentiment analysis of the 2020 us presidential election," in *Companion proceedings of the web conference 2021*, 2021, pp. 367–371.

[13] P. Rita, N. António, and A. P. Afonso, "Social media discourse and voting decisions influence: sentiment analysis in tweets during an electoral period," *Social Network Analysis and Mining*, vol. 13, no. 1, p. 46, 2023.

[14] P. Chauhan, N. Sharma, and G. Sikka, "Application of twitter sentiment analysis in election prediction: a case study of 2019 indian general election," *Social Network Analysis and Mining*, vol. 13, no. 1, p. 88, 2023.

[15] A. N. Ma'aly, D. Pramesti, A. D. Fathurahman, and H. Fakhrurroja, "Exploring sentiment analysis for the indonesian presidential election through online reviews using multi-label classification with a deep learning algorithm," *Information*, vol. 15, no. 11, p. 705, 2024.

[16] K. Attai, D. Asuquo, K. Okonny, A. Johnson, A. John, I. Bardi, C. Iroanwusi, and O. Michael, "Sentiment analysis of twitter discourse on the 2023 nigerian general elections," *European Journal of Computer Science and Information Technology*, vol. 12, no. 4, pp. 18–35, 2024.

[17] A. Balasubramanian, V. Zou, H. Narayana, C. You, L. Luceri, and E. Ferrara, "A public dataset tracking social media discourse about the 2024 us presidential election on twitter/x," *arXiv preprint arXiv:2411.00376*, 2024.

[18] N. E. Friedkin and E. C. Johnsen, "Social positions in influence networks," *Social networks*, vol. 19, no. 3, pp. 209–222, 1997.

[19] C. Albon, "Election day 2016 twitter," https://github.com/chrisalbon/election_day_2016_twitter, 2016, accessed: 2025-04-25.

[20] L. Grimminger and R. Klinger, "Hate towards the political opponent: A twitter corpus study of the 2020 us elections on the basis of offensive speech and stance detection," *arXiv preprint arXiv:2103.01664*, 2021.