Tackling COVID-19 Misinformation: Sentiment Analysis and False Claims Detection on Twitter

Krishna Vamsi Dhulipallal[1] Siva Sagar Kolachina[2] Yaswanth Chakiri[3]

[1]Virginia Tech, Blacksburg kdhulipalla13@vt.edu

[2]Virginia Tech, Blacksburg sivasagar@vt.edu

[3]Virginia Tech, Blacksburg yaswanth22@vt.edu

May 5, 2024

Abstract

In our research, 'Tackling COVID-19 misinformation: Sentiment Analysis and False Claims Detection on Twitter,' we study the complex interplay between sentiment analysis, false claim detection, and misinformation transmitted on the Twitter network. Beginning with sentiment analysis, we identify patterns in sentiments expressed in COVID-19-related tweets, which provide insight into general opinions. We then use false claim detection model to identify instances of misinformation, which is critical for ensuring factual accuracy. Using community detection algorithms, we identify groups of people with similar sentiment patterns, providing insights into sentiment clusters. Moving on to user influence analysis, we investigate how influential users influence sentiment propagation within the network using centrality measures such as degree centrality. Our analysis expands to look at correlations between sentiment scores and network structure measurements, uncovering intriguing patterns. Furthermore, we investigate how misinformation influences sentiment dynamics, discovering relationships between sentiment scores and the spread of misinformation. This extensive study provides useful insights into understanding and countering COVID-19 misinformation on Twitter, ultimately helping to develop educated public conversation and ensure factual correctness.

1 Introductiom

Social media sites like Twitter have developed into effective instruments for information sharing in the digital age, especially in times of international emergency like the COVID-19 epidemic. On these sites, there is a lot of correct information shared, but there is also a lot of misinformation and misleading claims as well. In addition to misleading the public, this false information poses a risk to public health and may even fuel panic and terror. Our study, "Tackling COVID-19 Misinformation: Sentiment Analysis and False Claims Detection," attempts to address the critical problem of countering COVID-19 misinformation on Twitter.

It is critical to counter false information about COVID-19. False information not only erodes public confidence in reliable sources of information, but it also makes it more difficult to contain the pandemic. Furthermore, the quick transmission of untrue information on social networking sites can increase public unease and aid in the propagation of dangerous behaviors.

Many different stakeholders, such as public health officials, legislators, researchers, and the general public, can benefit from our study. Legislators and public health authorities depend on precise data to guide their choices and create winning public health plans. Scholars aim to comprehend the workings of disinformation and how it affects public opinion. For the general population to make educated decisions regarding their health and well-being, they must have access to trustworthy information.

We have a multifaceted approach as part of our broader plan to tackle the problem. To gauge public opinion on the epidemic, we will first perform sentiment analysis on tweets pertaining to COVID-19. In parallel, we will classify and identify instances of disinformation using false claim detection tools. We will use a BERT model trained on the LIAR dataset to compute factual scores, which will allow us to evaluate the accuracy of assertions made in tweets. In addition, we will study how sentiments spread throughout the network and examine community discovery methods to find user groups with comparable sentiment patterns. Lastly, we will investigate connections between sentiment scores and network structure measures as well as the impact of prominent users on sentiment dynamics. By using these strategies, our project hopes to support factual truth, educate the public, and aid in the battle against COVID-19 disinformation.

2 Background/Related Work

Sentiment analysis techniques have been thoroughly investigated by researchers in an effort to comprehend public perception and emotional reactions during the COVID-19 pandemic. In order to identify bogus news on social media, Shu et al. (2020) used data mining techniques, underscoring the significance of sentiment research in the fight against misinformation. Similar to this, Kouzy et al. (2020) underlined the importance of sentiment analysis in evaluating public perceptions toward the pandemic by conducting a study to quantify the COVID-19 misinformation epidemic on Twitter.

Understanding how information spreads on social media platforms and identifying influential users are the goals of information propagation analysis. Discourse-aware rumor stance categorization approaches were developed by Zubiaga et al. (2018) to assess the spread of information on Twitter during crises. Their research emphasizes how influential users shape public opinion and how crucial it is to identify them in order to develop successful crisis communication plans.

Understanding sentiment analysis among various user groups or communities on social media platforms is the main goal of community sentiment analysis. During the COVID-19 pandemic, Cinelli et al. (2020) carried out an infodemiology study to examine the main issues raised by Twitter users. Their results demonstrated the frequency of false information and the variety of sentiment profiles among various user groups, highlighting the significance of sentiment analysis at the community level.

More detailed research that is suited to particular situations is required, even though the literature that is now in the works offers insightful information about sentiment analysis, the spread of information, and prominent users on social media during the COVID-19 epidemic. By concentrating on sentiment analysis, information dissemination, and identifying prominent individuals on Twitter during the ongoing COVID-19 outbreak, our research seeks to close this gap. This research aims to give actionable insights for public health strategies, policy-making, and crisis communication efforts by identifying key influencers, understanding the dynamics of information transmission, and assessing sentiment patterns.

William Yang Wang is an assistant professor at the University of California, Santa Barbara. He has made significant contributions to the fields of machine learning and natural language processing, especially in the area of identifying fake news. His most well-known contribution is the LIAR dataset, which was first reported in the 2017 paper "Liar, liar pants on fire: A new benchmark dataset for fake news detection" that was presented at the 55th Annual Meeting of the Association for Computational Linguistics. This dataset, which was produced in partnership with Sen Palghat, fills a need in the market by offering a benchmark for assessing the effectiveness of fake news detection methods on a broad scale. Wang has made substantial contributions to the field of research by focusing on creating novel approaches and methodologies to counteract false information and deceit in textual content.

Wang has made contributions to numerous additional NLP and machine learning domains in addition to his work on the identification of fake news. Discourse analysis, propaganda identification, and fine-grained sentiment analysis are among his areas of interest in research. Wang hopes to overcome difficult issues in internet communication and information distribution with his multidisciplinary approach. His work continues to influence the advancement of NLP methods and machine learning algorithms, expanding the field's awareness of and ability to counteract disinformation in digital contexts.

3 Approach

Our approach includes a number of crucial analytics aimed at analysing sentiment dynamics and misinformation transmitted over the Twitter network during the COVID-19 pandemic. These analyses include network, influence, and misinformation classification. Each analysis provides useful insights into user behaviour and information distribution across the network.

3.1 Network Analysis

In our network analysis, we first collected a large dataset of tweets related to COVID-19 using the Twitter API. We then preprocessed the text data by removing URLs, mentions, and special characters, and converting text to lowercase. This preprocessed text data served as the basis for our sentiment analysis. Using a pretrained BERT model fine-tuned on the LIAR dataset, we classified each tweet into positive, negative, or neutral sentiment categories. Additionally, we constructed a network of users and their interactions, including mentions and retweets, to understand the relationships between users.

3.2 Influence Analysis

Our influence analysis focused on identifying influential users within the network and understanding how sentiment propagates through the network. We calculated centrality measures such as degree centrality to identify influential users. Degree centrality measures the number of connections a user has within the network, indicating their level of influence. Additionally, we analyzed how sentiment propagates within the network by tracing the spread of sentiment scores from user to user. By studying the impact of influential users on sentiment propagation, we gained insights into the mechanisms driving sentiment dynamics within the network.

3.3 Misinformation Analysis

In our misinformation analysis, we aimed to identify and classify tweets containing misinformation. We calculated fact scores for each tweet using the same pretrained BERT model used for sentiment analysis. Fact scores represent the likelihood of a tweet containing misinformation, with lower scores indicating a higher likelihood. By applying a threshold to the fact scores, we classified tweets as either misinformation or not. We then analyzed the spread of misinformation within the network by tracing the path of misinformation from user to user. This analysis provided insights into how misinformation affects sentiment dynamics within the network and its potential impact on public perception during the pandemic.

4 Experiment

4.1 Global Engagement During the Pandemic

The 2020 COVID-19 dataset's location distribution provides information about the debate around the pandemic around the world. Notably concentrated in areas such as the "USA," "India," "London," and "Nigeria," it suggests broad participation and concern over a variety of geographic locations. This distribution highlights the significance of international cooperation in the fight against the pandemic and probably reflects the differing degrees of COVID-19 impact worldwide. It also emphasizes how important it is to comprehend local viewpoints and difficulties in order to create mitigation and public health solutions that work. Overall, the geographic distribution of the information highlights the interconnectivity of global health issues and shows how the whole community is working together to address the COVID-19 situation.

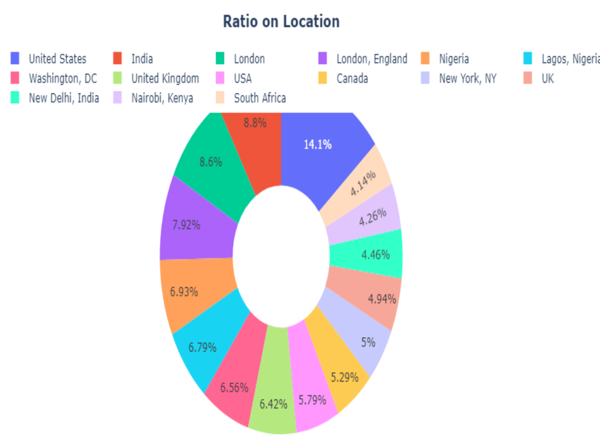

Figure 1: Ratio on Location

The sentiment analysis results demonstrate a wide range of reactions and provide insightful information about the dominant attitudes and emotions throughout the studied timeframe. Even though there is a minor bias toward positive thoughts, neutral statements are widely present, suggesting a complex emotional landscape. This intricacy is a reflection of the complex sociocultural context in which a variety of elements, including the state of the world at large, can influence public opinion. Comprehending these dynamics of mood is essential for deciphering social reactions, directing communication tactics, and assessing adaptability in the face of modern difficulties.

4.2 Word Count Analysis

The word cloud produced by our dataset—which consists of tweets related to the COVID-19 pandemic—provides a visual depiction of the dominant topics and debates in online discourse. Words like "pandemic" and "COVID-19" predominate the cloud, which is noteworthy since it reflects the general emphasis on the virus and its extensive effects on international communities. These well-known phrases highlight the seriousness of the situation and the pressing need for all-encompassing solutions to lessen its consequences.

Furthermore, the word cloud reveals a wide range of issues and subjects related to the pandemic, from talks about containment tactics like "lockdown" to analyses of the crisis's "impact" on society as a whole. Words like "support" and "relief" start to surface during these conversations, signifying a general understanding of the need for cooperation and teamwork in overcoming the pandemic's obstacles. Words like "patient" and "death" also serve as heartbreaking reminders of the human cost imposed by the virus, emphasizing how important it is to address both the immediate healthcare crisis and its long-term effects on people and communities.

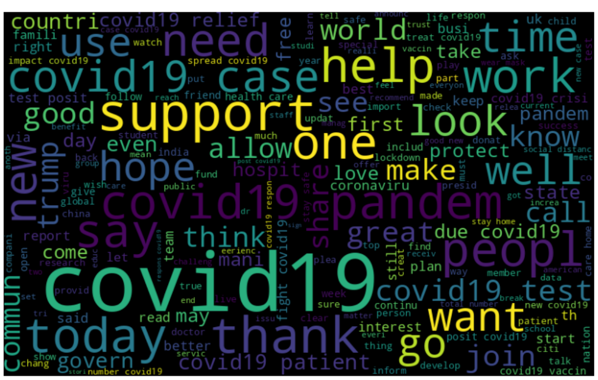

Figure 2: Wordcloud

4.3 Hastags analysis

The examination of hashtags present in the dataset provides valuable insights into the opinions shared on social media, especially with regard to the continuing COVID-19 pandemic. Interestingly, COVID-19-related hashtags like "coronavirus" and "covid19" predominate in discussions across all moods, highlighting the pandemic's widespread influence on online dialogues. This frequency points to a general obsession with COVID-19-related issues, demonstrating the disease's significant influence on public awareness and discourse regardless of personal opinion.

Analyzing the hashtags in the dataset offers insightful information on the views expressed on social media, particularly in relation to the ongoing COVID-19 pandemic. What's interesting is that COVID-19-related hashtags like "coronavirus" and "covid19" seem to be dominant in conversations of all kinds, indicating how much the pandemic has influenced conversations online. This frequency indicates a widespread preoccupation with COVID-19-related concerns, highlighting the disease's profound impact on public conversation and awareness irrespective of individual opinions.

4.4 Network Analysis Results

The network analysis offers fascinating new perspectives on the relationships and key players in the social media environment. Upon taking into account user mentions as well as original writers, well-known individuals such as @realDonaldTrump, @YouTube, and @JoeBiden become top-degree nodes, signifying their deep connections and interactions in the network. This implies that they have a major influence on social media conversations and are highly visible, which is indicative of their roles as influential people or focal points of online conversation.

On the other hand, when original authors are the only nodes considered, media organizations like @CNN, @SkyNews, and @WHO show up as top-degree tweeting nodes, demonstrating their importance in spreading news and sparking discussions on social media. This emphasizes how important media outlets are in influencing public opinion and disseminating news and updates, especially in emergency situations like the COVID-19 pandemic.

Additional information is provided by the average degree of nodes in the network, which shows the average number of connections or interactions per node. This metric provides insights into the general structure and dynamics of information distribution and social interactions within the observed dataset by measuring the network's overall density and connectivity.

This network is plotted using Gephi with the nodes

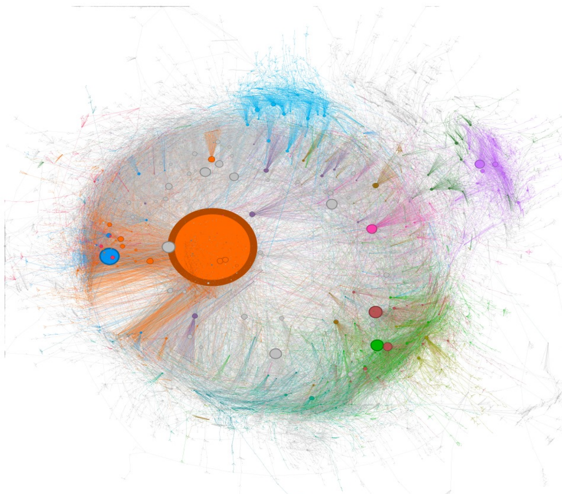

Figure 3: Gephi Network Plot

as the combination of the tweet authors and the author mentions with their interactions as edges. The colors to nodes are due to the modularity class to form communities and the size of the node are based on the degree of the node, so that the nodes with higher connections are significantly visible in the plot. In this case the orange node which is significantly large is realDonaldTrump who is the most influenced based on the covid-19 data we used.

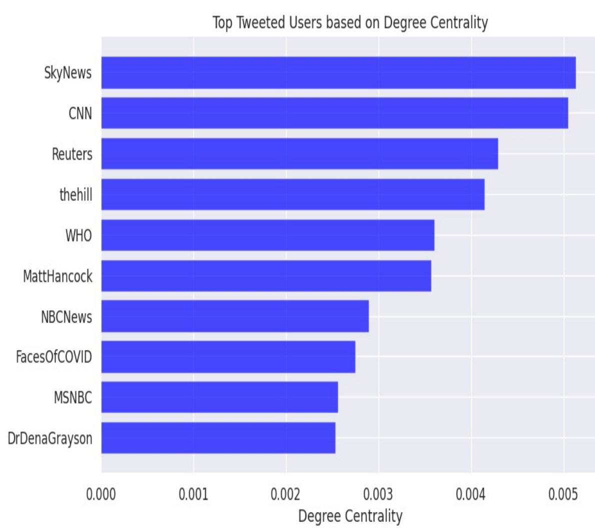

Figure 4: Tweeted Users Degree Centrality

The above plot includes prominent news organizations such as SkyNews, CNN, Reuters, and TheHill, indicating the significance of news outlets in driving discussions and disseminating information on Twitter. Additionally, the presence of health-related entities like WHO and public figures like Matt Hancock suggests a focus on health and political discourse within the network.

Similarly, for the most mentioned users feature notable political figures like Donald Trump, Joe Biden, and Boris Johnson, reflecting intense public interest and engagement with political discourse. YouTube's inclusion among the top mentioned users highlights the platform's role as a significant source of information and discussion on various topics.

4.5 Propagation Analysis Results

With a comparatively greater degree centrality and the highest propagated sentiment score, Narendramodi emerges as the most influential user. This implies that Narendramodi has considerable influence over opinions on the network and keeps close relationships with other individuals. His impact goes beyond only influencing sentiment; it has a broad reach and interaction with many network segments.

PMOIndia trails closely after, with a degree centrality that indicates a somewhat lower sentiment score but still significant influence. Because of their active participation and deep connections with other members of the network, this user continues to have a significant impact on mood and network dynamics.

In comparison to the top users, DrRPNishank, DG_NTA, and EduMinOfIndia have influential positions but have lower sentiment scores. Nonetheless, their noteworthy degree centrality values highlight their significance in promoting interconnectivity and easing the spread of information within the network. They may have less of an immediate effect on sentiment, but their crucial functions in preserving network cohesiveness and encouraging communication should not be undervalued.

Despite having lower degree centrality values and sentiment scores than the top users, transformIndia and PIBHindi nevertheless make significant contributions to the sentiment and connectivity of the network. Even if they have less of an impact, their existence in the network adds to the richness of viewpoints and encourages active idea sharing. In summary, the analysis highlights the multifaceted nature of influence within the network, with users such as Narendramodi and PMOIndia wielding sig-

	User	Propagated Sentiment Score	Degree Centrality
0	narendramodi	623.197838	0.001501
1	PMOIndia	530.465669	0.000874
2	DrRPNishank	448.343260	0.000535
3	DG_NTA	309.101022	0.000150
4	EduMinOfIndia	288.070956	0.000235
5	transformIndia	235.520718	0.000072
6	PIBHindi	219.964859	0.000046
7	Joee_Blackkk	207.078884	0.000124

Figure 5: Sentiment, Degree Centrality Scores

nificant sway over sentiment and network dynamics, while others such as DrRPNishank, DG_NTA, EduMinOfIndia, TransformIndia, and PIBHindi contribute to network cohesion and dialogue in varying capacities.

4.6 Community Analysis

4.6.1 Community Overview

Interesting insights into the sentiment dynamics among the top 10 Twitter communities were found by conducting a community analysis on our dataset. Every community displayed a unique sentiment profile that was indicative of the wide variety of viewpoints and beliefs that were present there. This investigation highlighted the complexity of sentiment generation across digital communities by illuminating the multifaceted character of online discourse concerning COVID-19 themes.

The top 10 communities found in the Twitter dataset have a varied landscape of sentiment trends, as shown by the community analysis. Every community displays distinct sentiment patterns that are indicative of the diverse viewpoints and attitudes that are common among its members. Interestingly, several communities exhibit a very even distribution of neutral, negative, and favorable opinions, suggesting a complex conversation around COVID-19 issues. Such balanced sentiment distributions, for example, are shown in Communities 2 and 7, indicating a lively flow of ideas and viewpoints within these bigger communities.

The community analysis reveals that the top 10 communities in the Twitter dataset exhibit a diverse range of sentiment trends. Each community exhibits unique sentiment patterns that reflect the range of opinions and attitudes that are shared by its constituents. It is noteworthy that a number of communities display a fairly balanced distribution of neutral, unfavorable, and positive viewpoints, indicating a nuanced discourse regarding COVID-19-related matters. For instance, Communities 2 and 7 display such balanced sentiment distributions, suggesting a vibrant exchange of ideas and points of view within these larger communities.

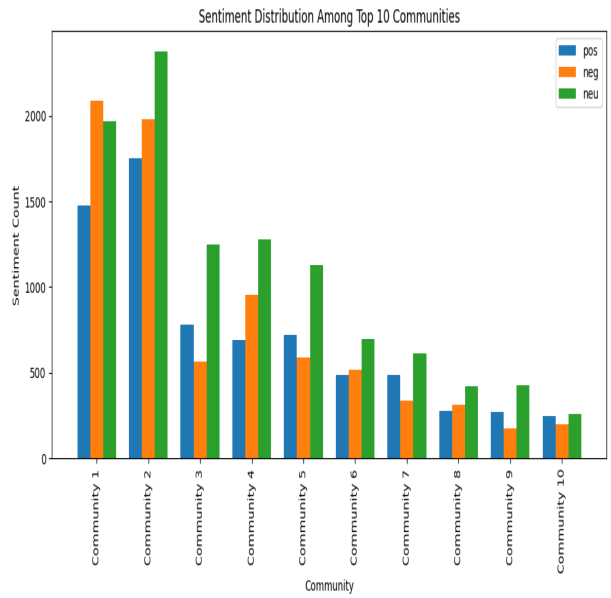

Figure 6: Sentiments Over Communities

4.6.2 Influencer Analysis in the Communities:

Our examination of the most influential members of the top 10 COVID-19 Twitter communities has shed light on the opinions influencing conversations around the epidemic. Through the assessment of influence scores in these groups, we have acquired practical knowledge about the people influencing stories and promoting interaction. This analysis contributes to our understanding of social dynamics and provides stakeholders with useful advice on how to interact with these communities in a productive way.

The wide range of influences present in the top 10 communities emphasizes how complex the conversation about COVID-19 is. Influential voices ranging from politicians like Donald Trump to media outlets like CNN and reputable organizations like the World Health Organization (WHO) demonstrate the diversity of sources influencing public attitudes and reactions to the pandemic. This variety highlights the intricate interactions between institutions and players that have shaped the online conversation around COVID-19.

The diverse array of impacts seen in the top 10 communities highlights the complexity of the COVID-19 discussion. Reputable institutions like the World Health Organization (WHO), politicians like Donald Trump, and media outlets like CNN are just a few examples of the influential voices that shape public perceptions and responses to the pandemic. This range demonstrates the complex relationships between organizations and participants that have influenced the online discourse surrounding COVID-19.

Finding the most influential people offers chances for deliberate interaction, boosting messages, and promoting conversation. Comprehending the dynamic character of influence enables interested parties to modify tactics, leverage patterns, and skillfully navigate conversations. Finally, examining influencers provides information on COVID-19 conversations, allowing interested parties to participate actively and tackle epidemic issues.

4.7 Misinformation Analysis Results

We thoroughly trained a BERT model, a cutting-edge transformer-based architecture, to categorise tweets based on factuality, which was measured using fact scores from the LIAR dataset. Following extensive fine-tuning, our model achieved a commendable average Mean Squared Error (MSE) score, indicating its ability to distinguish falsehood from accurate content. Subsequently, applying this trained model to our dataset revealed a worrying reality: nearly 20% of the tweets contained misinformation, a troubling proportion suggestive of the widespread propagation of misinformation throughout the pandemic.

Our analysis further revealed insights into the interplay between sentiment and misinformation within the Twitterverse. Despite the prevalence of misinformation, our findings unveiled a nuanced landscape, with misinformation distributed across various sentiment categories without a discernible bias towards any specific sentiment. Remarkably, the correlation between sentiment and misinformation was found to be negligibly low, hovering around -8.91e-05, thereby underscoring the limited impact of misinformation on shaping the overall sentiment within the network.

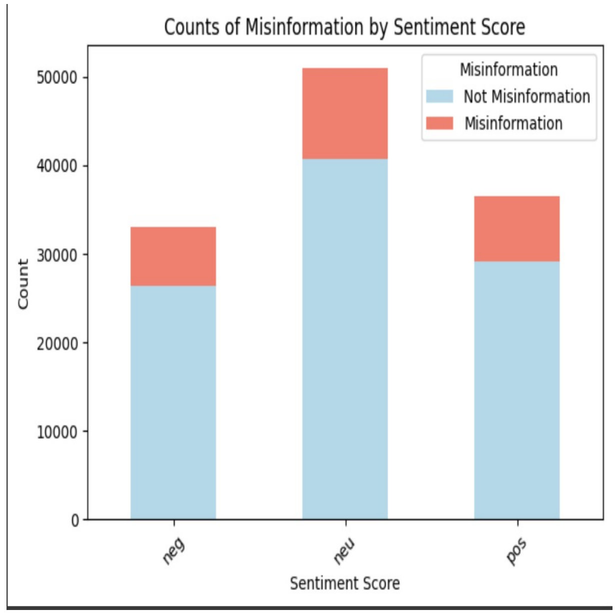

Figure 7: Count of Misinformation by Sentiment Score

4.7.1 Misinformation Conclusion

The findings of our study give an overview of the complex processes that regulate misinformation and sentiment diffusion inside the Twitter ecosystem during key global events.The BERT training yielded a robust model capable of discerning factual content from misinformation, a critical asset in combating the spread of falsehoods online. Despite the prevalence of misinformation, users appear to exercise discernment and skepticism, as evidenced by the lack of a strong correlation between sentiment and misinformation. Moving forward, these findings underscore the imperative of fostering digital literacy and promoting the dissemination of factual information to empower users in navigating the vast landscape of online information.

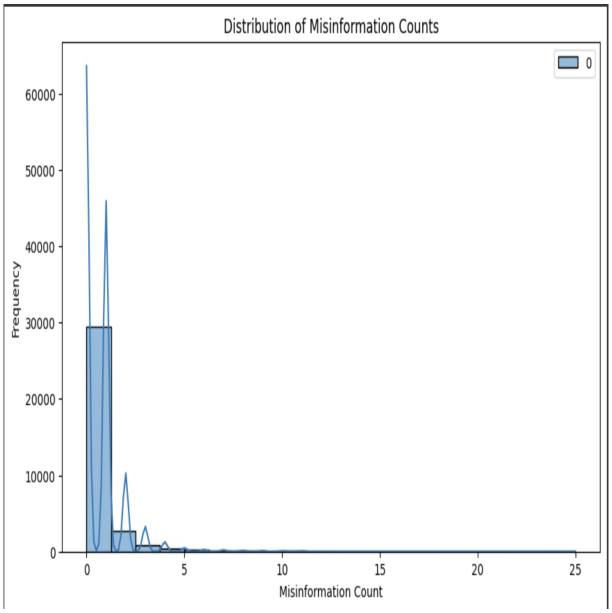

Figure 8: Distribution of Misinformation Counts

5 Conclusion

Our analysis into COVID-19 discussions on Twitter reveals the complicated interplay of mood, influence, and disinformation. By studying sentiment patterns, we can reveal a range of popular beliefs. Identifying significant personalities such as Narendramodi highlights their critical role in shaping conversations. Further, our study demonstrates instances of disinformation and the channels via which it spreads, underscoring the crucial need of accurate information. The community-level research reveals varied sentiment dynamics, reflecting the multidimensional character of online talks. Finally, our examination of powerful voices across groups highlights the diverse sources that shape public opinions. In summary, our findings provide important insights for countering disinformation, developing informed debate, and increasing public health literacy in digital environments.

References

[1] Wang, W.Y (2018)."Liar: A New Benchmark Dataset for Fake News Detection." IEEE Intelligent Systems, 33(4), 83-87

[2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova (2018). "Bert:Pre-training of Deep Bidirectional Transformers for Language Understanding" arxiv.org/abs/1810.04805

[3] Xie, J., Sreenivasan, S., Korniss, G., Zhang, W., Lim, C. (2021). "Information cascades in complex networks." Physics Reports, 900, 1-78.

[4] Zhang, X., Zeng, A., Shang, M. (2020). "Information cascade prediction in microblogging platforms: A dynamic Bayesian approach." Information Sciences, 511, 277-289.

[5] Bovet, A., Makse, H. A. (2020). "Influence of fake news in Twitter during the 2016 US presidential election." Nature communications, 11(1), 1-11.

[6] Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R. (2020). "Automatic Detection of Fake News about COVID-19." arXiv preprint arXiv:2006.09916.

[7] Bovet, A., Makse, H. A. (2020). "Influence of fake news in Twitter during the 2016 US presidential election." Nature communications, 11(1), 1-11.

[8] Gallotti, R., Valle, F., Castaldo, N., Sacco, P., De Domenico, M., Caldarelli, G. (2020). "Assessing the risks of 'infodemics' in response to COVID-19 epidemics." Nature Human Behaviour, 4(12), 1285-1293.

[9] Albahli, S., Abdulhay, E. (2021). "Fighting COVID-19 Misinformation on Twitter: Detection and Classification of Micro-Influencers and Fact-Checkers Using Graph Embeddings." In Proceedings of the 2021 4th International Conference on Information Science and Systems (pp. 1-6).

[10] Chen, E., Lerman, K. (2020). "Covid-19: The First Public Coronavirus Twitter Dataset." arXiv preprint arXiv:2003.07372.