International
JOURNAL OF SOCIAL,
HUMANITIES AND ADMINISTRATIVE SCIENCES
Open Access Refereed E-Journal & Refereed & Indexed
e-ISSN: 2630-6417

RESEARCH ARTICLE
Marketing

Political Polarization, Misinformation, and Sentiments: A Social Media Analysis About 'Capitol Hill 2021 Attack' by Tweets

ABSTRACT

January 6, 2021, was a noteworthy incident in the history of the United States of America. Dissidents attacked the Capitol building over the results of the 2020 presidential election. The purpose of this research is to understand how greater social media partisanship can create a greater division among many Americans leading to misinformation, political polarization, and mayhem in the society. The research analyzes queries generated through retrieved Twitter data on that specific data. In order to meet the study objective, a keyword analysis by using Python language programming and sentiment analysis was applied by using the Natural Language Processing method in the current study. According to the findings, the words such 'jobs', 'vaccine', 'president', 'spread' and 'pandemic' were the most dominant words in "retweets" and "likes" showing the concern of the society regarding the riot. Results also indicate words like 'love', 'stock', and 'market' were used in a positive manner to express the concerns while the words like 'president', 'unemployment', 'arrest', 'covid', and 'mask' were related to negative emotions. The research concludes that it is important to understand how social media could contribute to an attack on the Capitol building, the meeting place of the United States of America Congress and in order to better control this type of events, social media analysis outcomes can guide the society`s main areas of concerns.

Keywords: Twitter, sentiment, social media, word cloud, natural language toolkit (NLTK), Valence Aware Dictionary for Sentiment Reasoning (VADER)

Gulhan Bizel [1]
Amit Kumar Singh [2]

How to Cite This Article
Bizel, G. & Singh. A. K. (2023). "Political Polarization, Misinformation, and Sentiments: A Social Media Analysis About 'Capitol Hill 2021 Attack' by Tweets", Journal of Social, Humanities and Administrative Sciences, 9(61):2257-2266. DOI: http://dx.doi.org/10.29228/JOSHAS.64755

Arrival: 26 September 2022
Published: 28 February 2023

INTRODUCTION

Different examples of polarization appear to have been raised in both, Western and Eastern democracies. It is appropriate to hold social media platforms like Twitter and Facebook accountable for creating divisions within society. The primary reason is easy access to new information and content propounded by platforms such as Twitter or Facebook by manifesting users to like-minded people and by benefacting them with information customized to their interests, prejudices, and concerns. While social media dispense users with unlimited access to information from around the globe, the principles of diversity, discourse and debate are called into question by the rise of biased news and polarized discussion on digital platforms.

The specific objective of this study is understanding how social media can be used to understand the society`s perception for specific events like the Capitol Hill attack and how their justification on social media being led. The reason for the study is to find the answer to the question of whether social media communications could have an impact on other incidents such as local presidential elections. It demonstrates the significance of the study because events like presidential elections play a vital role in the future of a country, and their implications on societal issues may lead people into polarized directions. The theory point of view can be found with same concept in the past several research as well. In research from Guerra et al. (2013) polarization may be a key information for tasks such as opinion analysis. A biased opinion holder is likely to keep the same, extreme opinion over time, and the knowledge of the side in a discussion an opinion holder is (in favor or against an issue) can help predict the polarity of his/her opinions.

Literature Review

Social media has become a very important source for news consumption during the last decades as it becomes very easy to access due to fast development of internet. According to Matsa and Shearer (2018), as of August 2018, two-thirds of Americans report that they consume some news on social media – with two in ten doing so often. Among different social media, Facebook remains the most relevant: 45% of Americans consume news on Facebook (Abreu & Jeon, 2019). There is similar research of Campbell et al. (2019) mentions that social media (such as Twitter or Facebook) encompasses a much larger network than just direct friends, so that individuals are exposed to alternative political views through their weak ties and are thus less likely to hold extreme political opinions. Another example

[1] Ph.d., Saint Peter`s University, New Jersey, United States
[2] Ph.d., University of Essex, UK

from the recent past would be the 'Black Lives Matter' movement which continues to be especially prominent on Twitter and has intensified after the deaths of unarmed blacks resulting from police violence (Haffner, 2018).

2020 presidential elections outcome was on favor of Joe Biden against Donald Trump who was using actively Twitter and he tweeted a series of false allegations regarding the election and continues to do so even after the election and eventually communicating 'Save America Rally" to his supporters to march towards the Congress (Muhammad & Nirwandy, 2021) couple of hours before the attack. So, the attack to the Capitol was triggered by the continuous Twitter tweets with the history leading before the presidential elections.

The purpose of this research is to provide an analysis of a notable attack on Capitol Hill on January 6, 2021, during the presidential election, in the United States of America, and the connection to social media platforms and disinformation. The Capitol attack was a signature event in the history of the United States and a testament to the power of misinformation (Walsh, 2021). Immediately after the start of the event, posts related to the riots started to trend on social media (Prabhu et al., 2021).

The research intends to analyze Twitter data to understand if increased Twitter following leads to greater spread of misinformation in society. Answering these questions is critical in understanding whether social media platforms like Facebook, Twitter, and Parler are contributing to increased political polarization of individuals. Social media platforms evolve and create a divided society with highly polarized individuals. Political polarization is spiraling both in the United States, and across the world. Aggregators and "content farms" have sprung up to produce low-quality, sensational, and often misleading news stories framed to maximize clicks (Hindman, et.al. October,2018). Polarization began growing in the U.S. decades before Facebook, Twitter, and YouTube appeared. Some polarization is inexorable and leads to some consequences such as conspiracy theories. Highly biased individuals use social media platforms as a stage to heave their hatred and falsify facts to portray an image of disgust to drive unrest in communal sensitive places. Social media microblogging site Twitter has become beneficial as a news resource for some and is changing the way news is obtained. The fact is that information fabrication is not new, misinformation has been featured in human communication throughout history, and with new technology it has never been easier to spread it (Walsh, 2021). Disinformation differs from misinformation because it is information that is false, and it is intentionally created to harm a person, social group, organization, or country (Journalism, "Fake News" and Disinformation: A Handbook for Journalism Education and Training, 2020)

Polarization takes place in every economy, political systems, and among communities. However, social media companies fuel it. With new technology, anyone can produce and broadcast content that can reach millions of people. The study is important because it will help to understand how Facebook and Twitter try to manipulate their algorithm to make certain news stories or sources of disinformation. Polarization caused by social media also can be seen as an important societal problem.

Social media platforms have faced scrutiny from various organizations. In defense, they have often implemented various ways to claim social media itself has a scrutiny system. Twitter has claimed repeatedly that it blocked accounts that spread fake news. The persistence of so many easily identified abusive accounts is difficult to square with any effective crackdown (Hindman, et.al. October, 2018). Tweets that received a soft intervention spread further, longer, and received more engagement than unflagged tweets (Sanderson et al., 2021).

With the guidance of literature review, this research will focus on analyzing the tweets data set of Twitter, extracted from Kaggle of January 6, 2021, during the presidential campaign. The study will analyze the queries and conduct a sentimental analysis using the Natural Language Toolkit (NLTK). The study will further create visualizations using Microsoft Excel and Google data studio. The research study also intends to analyze the data to form four-word clouds; positive, negative, neutral, and overall, through sentimental analysis.

METHODOLOGY

The micro-blogging social media platform Twitter began as an SMS-based platform on a mobile phone. The quality of information and debate is damaged as certain voices are impairing in nature, and the escalation of divisive "outrage" culture leads to angry and aggressive expressions (Gorrell, et.al.,10 June 2020) as also similarly observed in the Capitol Hill riots during the 2021 presidential election.

Twitter has been using a structured dataset of 82,036 tweets from its platform with respect to the Capitol Hill attack, extracted from Kaggle. The data was extracted in "comma separated values" or .csv format and converted into Microsoft Excel format. Keyword analysis and sentiment analyses have been applied for the dataset to find out the outcomes. Both keyword analysis and sentiment model have been applied to similar work in the literature. According to recent research done for understanding public health surveillance opinions keyword selection for Twitter data method has been used (Edo-Osagie et al., 2020). In research from Budiharto and Meiliana (2018) sentiment analysis

has been applied for analysis of Indonesia potential election from Twitter. To better visualize and understand the contents of the dataset, the tweets' text is preprocessed, from which word clouds are constructed according to the research of Zervopoulos et al. (2020) where similar method has noticed for Hong Kong protests analysis. That`s why it has been adapted the methodology in order to build results in a synthetic way for analyzing big data of Twitter.

Keyword Analysis

The methodology flow of keyword analysis is explained in detail below.

- ✓ The study conducted keyword analysis on data generated through Twitter by using Microsoft Excel, Google Data Studio, Lexicon, and Python Programming language.
- ✓ The study categorized all queries, calculated the overall query count, and tracked the percentage of tweets, likes and retweets of each query and created visualization.
- ✓ The study was categorized queries with respect to the follower count greater than or equal to 100,000 follower count (applying filter with a condition greater than or equal to) and track the percentage of user tweet count, retweets and likes of each query and created visualization.
- ✓ It was compared the effect of Twitter dataset queries with respect to follower count greater than or equal to 100,000 and less than 100,000 follower count.
- ✓ It was divided into two sections, unfiltered data for overall readings and filtered data with follower count >=100,000 to understand the effectiveness of queries which are present in the form of keywords.

Sentiment Analysis

Sentiment analysis was performed by using the Natural Language Toolkit (NLTK) through Valence Aware Dictionary for Sentiment Reasoning (VADER) analyzer for a dataset of 82,036 unique tweet IDs comprising 3.6 billion user tweet count. Natural Language Processing can be run by several different stages for text and keyword analysis. Text is taken as an input for lexical analysis, and it generates tokens for the next phase **(**Bachate, R. P., & Sharma, A, 2020).

Figure 1 represents some of the NLP capabilities such as text clustering, information extraction, named entity resolution, text categorization, relationship extraction, topic modeling, and sentiment analysis. NLP is a branch of artificial intelligence within computer science aiding in understanding the way humans express themselves in the form of words and texts. In this study sentiment analysis was used for research methodology.

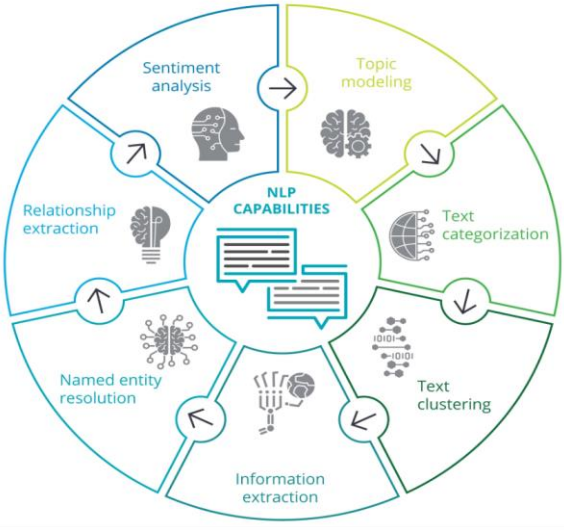

Figure 1: NLP applications (Key NLP Capabilities, 2019)

Sentiment analysis of Twitter data was done using Python Programming language to quantitatively analyze user tweets. The sentiment analysis was performed on Twitter dataset by using the NLTK 3.6.1 text classification process. NLTK's VADER analyzer categorizes text into three sentiments: positive, negative, or neutral where they are given a polarity score. The VADER lexicon correlates extraordinarily well in the social media domain. Sentiment analysis is a technique to measure users' emotion on social media. Figure 2 below shows the various steps of sentiment analysis using VADER.

Figure 2: Process of sentiment analysis using VADER

Figure 2 shows the process of sentiment analysis using VADER which is a simple rule-based model used in general sentiment analysis with seven steps. It determines the emotions of users who tweeted during the Capitol Hill attack. As the name suggests, sentiment analysis is used to identify people's sentiment. To better understand the mindsets of users on social media, sentiment analysis is a proven method. Natural language processing is used to understand the views of users and then sentiments of users are identified **(**Bachate, R. P., & Sharma, A, 2020)**.** If the sentimental analysis is applied to the text analytics, then transmuted properly into a specified structured text data format (Solanki, 2019). Lemmatization is established on its usage; and the machine looks for the relevant dictionary form of the word. In inflected languages search the problem of ambiguous words is common. The same word can have contrasting connotations depending on the context. In such a situation, lemmatization is crucial. To regulate the lemma of the word it considers its deliberate meaning. Depending on the context, the tool will axiomatically determine which is the right lemma for the word and therefore the results fetched for a query will be more accurate. There are numerous techniques to classify sentiments such as the machine learning approach and lexicon-based approach. ("Stemming and Lemmatization of Tweets for Sentiment Analysis Using R," 2019). In the lexicon approach, each of the words is grouped as positive, negative, or neutral and is given a polarity score. After obtaining the score of each letter in a sentence, a compounded score of the sentence is calculated. The lexicon-based approach involves calculating orientation for a document from the semantic orientation of words or phrases in the document (Turney 2002) (Taboada et al., 2011).

Word Cloud is an image composed of many words that implies the contents of the analyzed document. A word or tag cloud is a "method to visualize textual data, where the importance of each word in the text is highlighted by its font size, and/or color." (Hassler & Flinchbaugh, 2012).

RESULTS

The tweet data was analyzed in several ways to show sentiment on the Capitol Hill attack. Visualizations were created through keyword analysis of the queries in Google Data Studio, word cloud analysis, and sentimental analysis. The purpose of visualization is to show how these incidents affect thinking (in the form of queries, likes and retweets) and provide additional insights into the public view. The following observations have been made according to the analysis done.

Retweet Analysis

It has been formed retweet analysis for each query in total tweet count. Figure 3 represents the percentage of retweets in each query keyword for overall tweet data and the top five queries are: 'jobs', 'vaccine', 'president', 'spread', and 'pandemic' contributing to 69.9% of overall retweets.

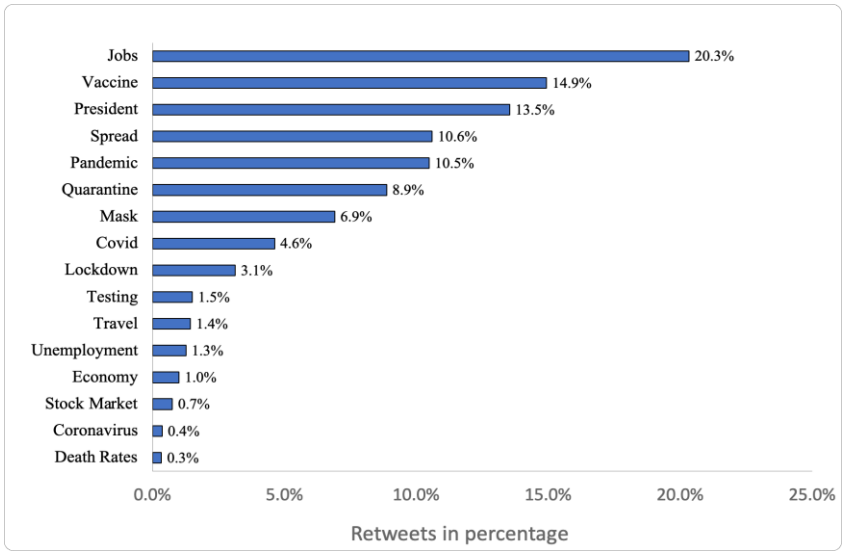

Figure 3: Keyword analysis in terms of percentage of retweets in total tweet count for each query for January 6th, 2021

Figure 4 represents the percentage of retweets on each query (on a scale of >=100,000 follower count) and the top three queries which are: "president", "jobs", "spread" contributing to 73.5% of total retweets.

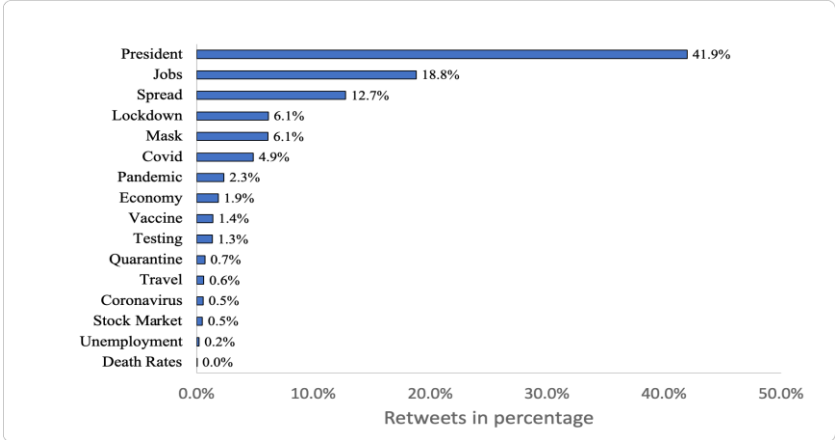

Figure 4: Retweets analysis on each query (on a scale of >=100,000 follower count)

'Likes' Analysis

In order to analyze 'likes' queries in >= 100,000 follower count have been analyzed. The Figure 5 pie chart indicates the majority like of the highest query to the other queries in the overall Tweet count. Likes on the 'president' query is 69.14% and the "others" is 30.86%. 'Others' query values with their percentage in descending order are: 'covid' represents 9.4 % followed by 'lockdown' 5.1%, 'mask' 4.4%, 'vaccine' 2.7%, 'pandemic' 2.0%, 'economy' 2.0%, 'spread' 1.20%, 'testing' 1.0%, 'stock market' 0.7%, 'jobs' 0.7%, 'coronavirus' 0.6% 'travel' 0.5%', quarantine' 0.4%, 'unemployment' 0.1% and 'death rates' is 0.04%.

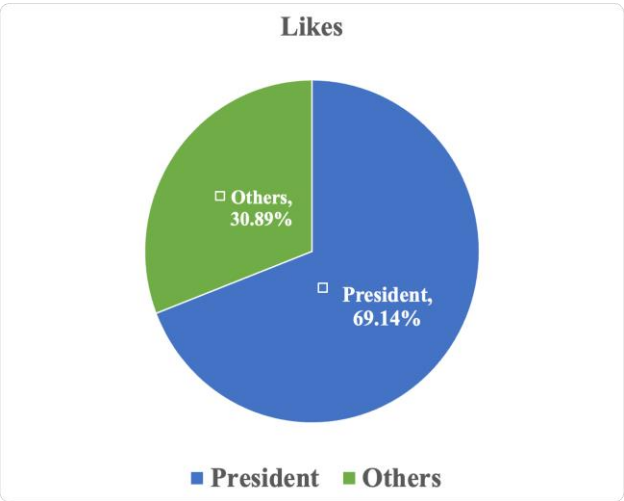

Figure 5: "Likes" analysis for queries in >= 100,000 follower count

Comparison Analysis of User Tweet Count Likes and Retweets

The likes and tweets analysis has brought the research to the point where it may be thought interesting to look for comparisons. Figure 6 represents the comparison between greater than or equal to 100,000 follower count and less than 100,000 follower count, taking in consideration parameters like user tweet count, like, and retweets. Figure 8 shows greater than or equal to 100,000 follower count has 5.4% user tweet count, 92.2% likes and 0.4% retweets. On the other hand, less than 100,000 follower count has 94.6% user tweet count, 7.8% likes, and 99.6% retweets.

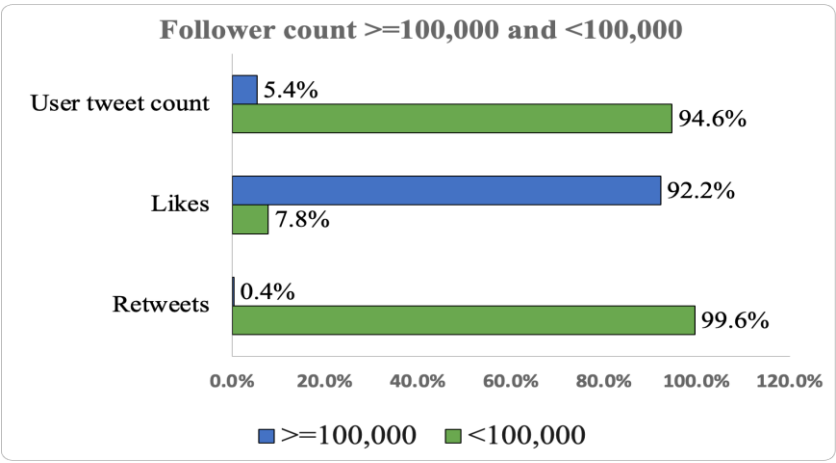

Figure 6: Comparison of user engagement of tweet count, likes, and retweets in >=100,000 follower count and in <100,000 follower count

Sentiment Analysis

Sentiment analysis was run through the dataset and the keyword dictionary created for three different clusters accordingly. Below Figure 7 illustrates the various categories of sentiment analysis with example keywords.

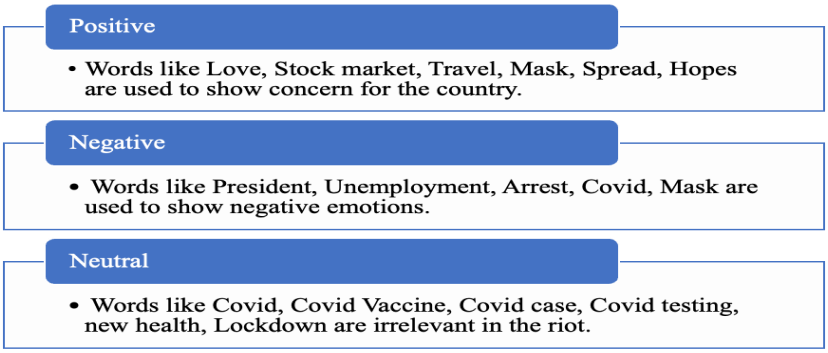

Figure 7: Three categories of sentiment analysis with example keywords

Figure 7 shows the overall sentiments for positive, negative, and neutral categories. Words such as 'love', 'stock market', 'travel', 'spread', and 'hope' are used in a positive manner to express concern for the country. On the contrary, words like 'president', 'unemployment', 'arrest', 'covid', and 'mask' are related to negative emotions. Words like 'covid', 'covid vaccine', 'covid case', 'covid testing', 'new health', and 'lockdown' have no connection with the January 6, 2021, incident. They are irrelevant and considered neutral words in the overall sentiment analysis.

Figure 8 represents the overall sentiment analysis and according to the analysis, the positive word count is 38,407, negative word count is 28,175, and neutral word count is 15,727. It is concluded that the nature of the queries has maximum positive word count, representing users' concern for the country during the Capitol Hill attack.

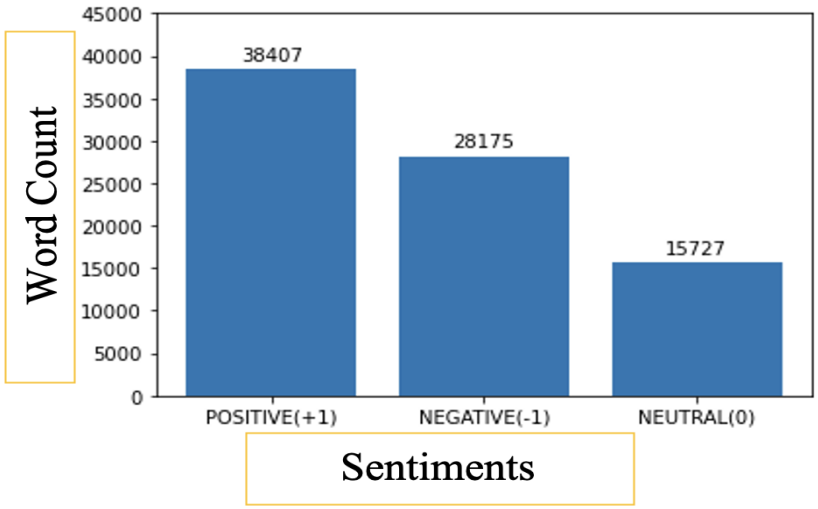

Figure 8: Overall sentiment analysis

Word Clouds

After completing the sentiment analysis, it was decided that showing the keywords by using word clouds was necessary. According to Figure 9, the overall word cloud analyzed by sentiment analysis using the NLTK important words was visualized. The most prominent words were ‘people’, ‘stock market’, ‘pandemic’, ‘economy’, ‘lockdown’ ‘covid’, ‘quarantine’, ‘need’, ‘better’, ‘spread’, ‘country’ and 'mask’. It is observed that ‘people’ has the maximum significance in the Capitol Hill attack.

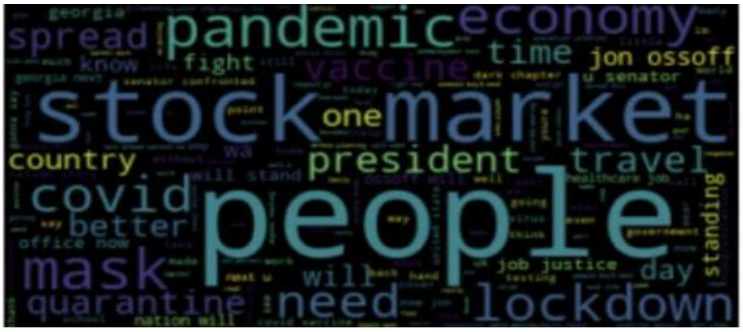

Figure 9: Overall word cloud

Figure 9a represents the results of the word cloud for best sentiment words. Words like ‘love’, ‘travel’, and ‘stock market’ are used to show concern for the country.

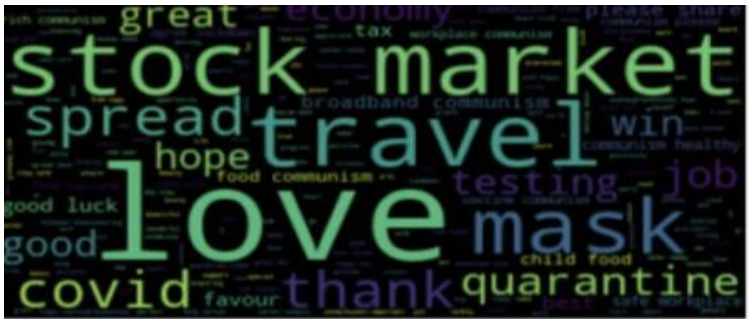

Figure 9a: Positive Sentiment word cloud

When the world cloud has been analyzed for negative words, the outcomes are different. Figure 9b represents the worst sentiment words. Words like ‘president’, ‘unemployment’ and ‘arrest’ show the negative emotions of the users.

Figure 9b: Negative Sentiment word cloud

Figure 9c represents the neutral sentiment words like ‘covid’, ‘covid vaccine’, and ‘testing’. These words are irrelevant with respect to the incident of the Capitol Hill attack.

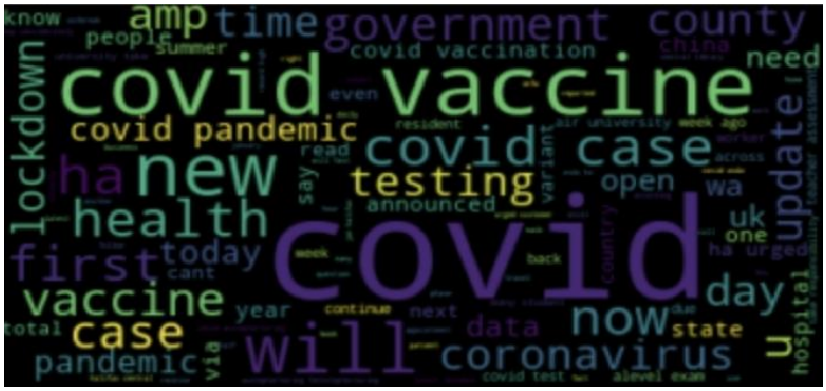

Figure 9c: Neutral Sentiment word cloud

DISCUSSION OF FINDINGS

Research findings reveal that the words like ‘jobs’, ‘vaccine’, ‘president’, ‘spread’, and ‘pandemic’ represent 69.9% of the total retweets posted during the day of the attack. When retweets analysis was run for >=100,000 follower count, the ratio became even higher and is 73.5% indicating the words where the society focused on that day. Likes analysis shows that ‘president’ query portion is 69.14% for >=100,000 in the follower count. The common word for both analysis; ‘president’ is evidence that the incident has been attached to the presidential elections.

Overall, it is important to note a second conclusion that likes are greater with 92.2% for Twitter accounts that has >=100,000 follower count, while retweets are 96.6% for Twitter accounts <100,000 follower count. That`s why it is vital to check both retweets and likes while analyzing the Twitter dataset.

Society`s emotions that has been revealed on social media are a mix of positive and negative sentiments: 'love’, ‘stock market’, ‘travel’, ‘spread’, and ‘hope’ are used in a positive manner to express concern for the country while ‘president’, ‘unemployment’, ‘arrest’, ‘covid’, and ‘mask’ are related to negative emotions. Overall sentiments have been expressed positively as the positive word count is 38,407, negative word count is 28,175, and neutral word count is 15,727. Words like ‘covid’, ‘covid vaccine’, ‘covid case’, ‘covid testing’, ‘new health’, and ‘lockdown’ have no connection with the attack. They are irrelevant and considered as neutral words in the overall sentiment analysis.

CONCLUSION

This research has studied user engagement and sentiments of users in relation to the Capitol Hill attack of January 6, 2021, during the presidential election with respect to the tweets by users on Twitter in the United States. Obtaining news through social media is increasing, but the effect of social media on public judgment remains contended. The study shows that social media is very powerful because consumers are exposed to different news on social media, and they fortuitously consume news when it becomes accessible in their feed, and so exposure to news on social media affects attitudes. The results show how higher follower count leads to a higher engagement rate among users. It has been also conducted sentiment analysis which is a computational study that analyzes mankind feelings, reaction, response, or emotions for specific concepts. In order to increase the potency of the analysis, the classification of the text based on natural language processing was deployed. Lastly, even though social media platforms like Twitter are associated with pro-attitudinal content, they can expose individuals to varied perspectives which often leads to political polarization as observed in the Capitol Hill attack.

Literature of polarization on different schools of thoughts in social networks has devoted attention to domains previously known to influence polarization. There is a growing body of work demonstrating how people's lives are getting affected by social media. Similarly, this research also exhibits the effect of social media users with a particular incident, the Capitol Hill attack. Therefore, the study will be significantly useful to various stakeholders such as policymakers, government communication teams and governmental organizations for understanding the sentiments of the society during a specific incident such as an attack. Theoretically, this research will be adding an extension to scholarship for how to create outcomes from social media channels by using methods like keyword analysis, word cloud analysis, and sentiment analysis especially for communication organizations, marketing agencies and such.

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

The study can procure data from only one day (i.e., January 6, 2021) which is the day the Capitol Hill attack took place. If more data points were available in the data set, of the entire week or month from the day the Capitol Hill attacks, the study could have affirmed a more detailed result on people's opinion.

The research study has a limitless scope in the form of lasting effects of the incident on other matters concerning politics and economy. Future research could study the effects of political unrest created through social media and its effects on various other economic factors like wages, governmental activities, laws, policies, tax rates, interest rates and unemployment. This research could be a potential future research approach when possible similar incidents has happened. This research could also guide to how it can be adapted to other social media channels by using the same approach.

CONFLICT OF INTEREST

The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

FUNDING

This research did not receive any outside funding or support. The authors report no involvement in the research by the sponsor that could have influenced the outcome of this work.

AUTHORS` CONTRIBUTIONS

All authors have participated in drafting the manuscript. All authors read and approved the final version of the manuscript. All authors contributed equally to the manuscript and read and approved the final version of the manuscript.

ACKNOWLEDGMENT

We would like to thank Dr. Joseph W. Gilkey Jr. from Frank J. Guarini School of Business, Data Science Institute at Saint Peter's University for his contribution to this research paper.

REFERENCES

Abreu, L., & Jeon, D. S. (2019). Homophily in Social Media and News Polarization. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3468416

Bachate, R. P., & Sharma, A. (2020). Acquaintance with Natural Language Processing for Building Smart Society. E3S Web of Conferences, 170, 02006. https://doi.org/10.1051/e3sconf/202017002006

Budiharto, W., & Meiliana, M. (2018). Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. Journal of Big Data, 5(1). https://doi.org/10.1186/s40537-018-0164-1

Campbell, A., Leister, C. M., & Zenou, Y. (2019). Social Media and Polarization. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3419073

Edo-Osagie, O., Iglesia, B. D. L., Lake, I., & Edeghere, O. (2020, November). An Evolutionary Approach to Automatic Keyword Selection for Twitter Data Analysis. In International Conference on Hybrid Artificial Intelligence Systems (pp. 160-171). Springer, Cham.

Gorrell, G., Farrell, T., & Bontcheva, K. (2020). Mp twitter abuse in the age of covid-19: White paper. arXiv preprint arXiv:2006.08363.

Guerra, P., Meira Jr, W., Cardie, C., & Kleinberg, R. (2013). A measure of polarization on social media networks based on community boundaries. In Proceedings of the international AAAI conference on web and social media (Vol. 7, No. 1, pp. 215-224).

Haffner, M. (2019). A place-based analysis of# BlackLivesMatter and counter-protest content on Twitter. GeoJournal, 84(5), 1257-1280.

Hassler, H. C., & Flinchbaugh, M. (2012). Biz of Acq-Using Tag Clouds that Visualize Circulation Patterns and Inform Acquisitions. Against the Grain, 24(4). https://doi.org/10.7771/2380-176x.6204

Hindman, M., & Barash, V. (2018). Disinformation, and influence campaigns on twitter. Knight Foundation: George Washington University.

Journalism, "Fake News" and Disinformation: A Handbook for Journalism Education and Training. (2020). Unesco. https://en.unesco.org/fightfakenews/modules

Key NLP capabilities. (2019, January 16). https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/natural-language-processing-examples-in-government-data.html

Matsa, K. E., & Shearer, E. (2018). News use across social media platforms 2018. Pew Research Center, 10.

Muhammad, M. R. A., & Nirwandy, N. (2021). A Study on Donald Trump Twitter Remark: A Case Study on the Attack of Capitol Hill. Journal of Media and Information Warfare Vol, 14(2), 75-104.

Prabhu, A., Guhathakurta, D., Subramanian, M., Reddy, M., Sehgal, S., Karandikar, T., ... & Kumaraguru, P. (2021). Capitol (Pat) riots: A comparative study of Twitter and Parler. arXiv preprint arXiv:2101.06914.

Sanderson, Z., Brown, M. A., Bonneau, R., Nagler, J., & Tucker, J. A. (2021). Twitter flagged Donald Trump's tweets with election misinformation: They continued to spread both on and off the platform. Harvard Kennedy School Misinformation Review. https://doi.org/10.37016/mr-2020-77

Solanki, M. (2019). Sentiment Analysis of Text using Rule Based and Natural language Toolkit. International Journal of Innovative Technology and Exploring Engineering, 8(12S), 164–168. https://doi.org/10.35940/ijitee.l1049.10812s19

Stemming and Lemmatization of Tweets for Sentiment Analysis using R. (2019). International Journal of Recent Technology and Engineering, 8(2), 2038–2040. https://doi.org/10.35940/ijrte.b2157.078219

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. Computational Linguistics, 37(2), 267–307. https://doi.org/10.1162/coli_a_00049

Turney, Peter. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings of 40th Meeting of the Association for Computational Linguistics, pages 417–424, Philadelphia, PA.

Walsh, D. R. (2021). Neutral Isn't Neutral: An Analysis of Misinformation and Sentiment in the Wake of the Capitol Riots. The Research Repository @ WVU. https://researchrepository.wvu.edu/etd/8055/

Zervopoulos, A., Alvanou, A. G., Bezas, K., Papamichail, A., Maragoudakis, M., & Kermanidis, K. (2020). Hong Kong Protests: Using Natural Language Processing for Fake News Detection on Twitter. IFIP Advances in Information and Communication Technology, 408–419. https://doi.org/10.1007/978-3-030-49186-4_34