SOCIAL MEDIA'S TAKE ON THE CLIMATE CHANGE (1456 words)

Srichandana Chakilam, Spring 2023

INDIANA UNIVERSITY BLOOMINGTON

ILS-Z639 Social Media Mining

Instructor: Ali Ghazinejad

May 2023

1. INTRODUCTION

Climate change is a hot topic that has gradually become a growing concern today. It is a major global issue that needs a plan of action sooner than later to save the earth's life from the impending threat. Social media has played a substantial role since its inception in setting the public's perception of climate change by driving conversations around it. Today, many users on different social media platforms are discussing many aspects of climate change, its causes, impacts, and possible solutions. Many social media platforms have taken a step forward in guiding their users to support and promote sustainable products for a sustainable living. While some platforms allowed environmental activists to voice their notions on saving and healing mother nature, other platforms have incorporated features like stickers, stories, and filters that motivate the public and create awareness among them. In 2020 the tech-giant Meta announced that it had achieved net-zero emissions and had committed to being 100% supported by renewable energy by the end of that year. The ability to create, ideate, debate, and discuss climate-related conversations, posts, and tweets on social media has taken form recently after COVID attacked humans. Many users have initiated climate activism and advocacy, alongside sharing information about rallies, protests, and other actions to raise awareness and advocate for policy change. Infographics and data visualizations have also become a popular way to present complex information about climate change in a more accessible and easy-to-understand format. Sharing news articles and scientific reports about climate change to keep their followers informed about the latest research and developments has also become prominent among social media users.

Zander et al. (2023) explores the use of Twitter data as a tool for analyzing responses to heat waves[1]. The study analyzes tweets from different locations across the world to identify how individuals and communities respond to heat wave events. The results indicate that people tend to use Twitter to share information about heat waves and to express their emotional reactions to the events. It also identifies that individuals in areas with higher temperatures and higher heat indices were found to be more likely to use Twitter to share information and seek support during heat wave events. In their 2019 article, "The social media life of climate change: Platforms, publics, and future imaginaries," Pearce et al. explore the different ways that social media platforms are being used to disseminate information about climate change and to shape public attitudes towards it[2]. The authors examine the unique features of different social media platforms and the challenges and opportunities they present for climate change communication and discuss the potential for social media to promote more sustainable behaviors and to help build more positive future imaginaries around climate change.

2. RESEARCH QUESTION

As one of the most upcoming and evolving technologies of the modern era, social media has become an integral tool for transmitting information and shaping public opinion. The paper aims to answer how social media platforms influence public attitudes and behaviors towards climate change.

3. METHOD

3.1 DATA

The objective of this paper is to examine the conversations happening on social media platforms to gain insight into the most discussed topics related to climate change among the public. As Reddit is popular and easily accessible, many users discuss climate change and potential solutions on this platform. To continue the analysis, we have collected more than 3000 submissions using purposive sampling methodology from major subreddits like r/worldnews, r/climate, r/environment, r/ClimateActionPlan, r/climate_science, r/EcoNewsNetwork, etc. These submissions are collected using specific keywords related to climate change. Some of them are: 'heat', 'heat waves', 'torrid heat', 'climate change', 'greenhouse', 'flood', 'sea level rise', 'global warming' etc. A time range of 600 days is set to return the data from two years ago till now. To achieve this, we used Python Reddit API Wrapper (PRAW) and created a Reddit API instance.

3.2 ANALYSIS

The analysis of climate change discussions on social media platforms such as Reddit is becoming increasingly popular, as it provides insights into the public's attitude and perceptions of this critical issue. By using the unsupervised machine learning technique LDA, we have identified the key themes and topics that are being discussed by users on Reddit, which can help inform communication and outreach efforts related to climate change. The purpose is to identify the major topics or themes present in the Reddit titles. We begin with pre-processing the text data by removing URLs, emojis, stop words, punctuation, and numbers. Then we proceed to create a term-document matrix using a count vectorizer object to convert the text data into a numerical format that can be used for modeling.

Latent Dirichlet Allocation (LDA) model is applied to the term-document matrix to identify the major topics present in the data. LDA is a widely used topic modeling algorithm that seeks to identify underlying topics based on the frequency of words in the text data. In this analysis, the LDA algorithm is set to identify 10 different but closely related topics. Finally based on the submissions, we assign topic labels to each document based on the highest probability topic and create a function to label the topics. The labeled topics can be used to gain insights into the major themes present in the titles of the Reddit submissions related to climate change. One possible limitation of model is that there can be scenarios where different researchers may label the same topics differently based on different contexts or sets of documents. However, the goal of LDA is to provide an objective approach to topic modeling by identifying topics based on the patterns of word co-occurrences in the text corpus.

	title	clean_title	topic_label	topic
25	Suffering under the worst heat wave in more th...	suffer bad heat wave decade Argentina issue he...	5	Topic 5: Alerts and news
26	China braces for ‘big heat’ day with temperatu...	China brace big heat day temperature set soar	8	Topic 8: Human activities - climate impact
27	India suffered income loss of $159 billion in ...	India suffer income loss $ key sector extreme ...	1	Topic 1: Heatwaves and emergency
28	Historic heat wave, wildfires hit Europe as te...	historic heat wave wildfire hit Europe tempera...	1	Topic 1: Heatwaves and emergency
29	Extreme heat is slamming the world's three big...	extreme heat slam world big economy	9	Topic 9: Government laws, urban life
30	Global decarbonisation scenarios envisioned by...	global decarbonisation scenario envision oil g...	2	Topic 2: Extreme temperatures
31	UN chief, Gore, others give heated warnings in...	UN chief Gore heated warning climate talk	3	Topic 3: Globalwarming - relevant research
32	'Most serious' drought on record grips parts o...	drought record grip part Europe extreme heat hit	1	Topic 1: Heatwaves and emergency
34	‘Vulnerable’ South Asia least prepared to deal...	vulnerable South Asia prepare deal urban heat ...	8	Topic 8: Human activities - climate impact
35	Ottawa’s spring heat wave continues, with a th...	Ottawa spring heat wave continue straight day ...	1	Topic 1: Heatwaves and emergency
36	Extreme heat in oceans ‘passed point of no ret...	extreme heat ocean pass point return | ocean	1	Topic 1: Heatwaves and emergency
37	Canadians need to do more to prep for 'potenti...	canadian need prep potentially lethal extreme ...	5	Topic 5: Alerts and news
38	Ukraine warns of long outages after wave of Ru...	Ukraine warn long outage wave russian strike h...	1	Topic 1: Heatwaves and emergency
39	March 2022: Earth’s 5th-warmest March on recor...	March earth warm March record | month feature ...	1	Topic 1: Heatwaves and emergency
40	Heatwave: Ferocious European heat heads north	heatwave ferocious european heat head north	5	Topic 5: Alerts and news
41	An Extraordinary Heat Wave Exposes the Limits ...	Extraordinary Heat Wave expose Limits Protecti...	8	Topic 8: Human activities - climate impact
42	An Ominous Heating Event Is Unfolding in the O...	Ominous heating Event unfold Oceans	3	Topic 3: Globalwarming - relevant research
43	Extreme heat is the new normal for the world’s...	extreme heat new normal world ocean accord study	5	Topic 5: Alerts and news
44	Climate change, extreme heat and drought end w...	climate change extreme heat drought end way li...	4	Topic 4: Relief and help
45	‘I’m all for climate change’: Axel Springer CE...	climate change Axel Springer CEO face heat lea...	4	Topic 4: Relief and help
46	Greater glider now endangered as logging, bush...	greater glider endanger logging bushfire globa...	0	Topic 0: Ocean/Sea level
47	Stop talking, start acting, says Africa’s firs...	stop talk start act say Africa extreme heat of...	3	Topic 3: Globalwarming - relevant research
48	Blazes a ‘wake-up call’ on climate change, fir...	blaze wake climate change firefighter warn rec...	4	Topic 4: Relief and help
49	Exposure to deadly urban heat has tripled sinc...	exposure deadly urban heat triple 1980s affect...	8	Topic 8: Human activities - climate impact
50	Future generations face ‘climate carnage’ with...	future generation face climate carnage surge f...	3	Topic 3: Globalwarming - relevant research

Figure 1. A list showing the titles labeled based on the topics discussed.

4. RESULTS

During our analysis, we attempted to visually represent the prevalence of different topics discussed on Reddit based on the titles of submissions. Our findings suggest that a significant number of users either seek help or provide assistance in response to various issues. Following that, the discussion revolves around extreme temperatures and heat waves. On the other hand, the least discussed topic was warning others about upcoming issues. It implies that people seem to focus more on debating and discussing solutions instead of alerting others, or perhaps there are gaps in the management of alerts when an impending issue arises.

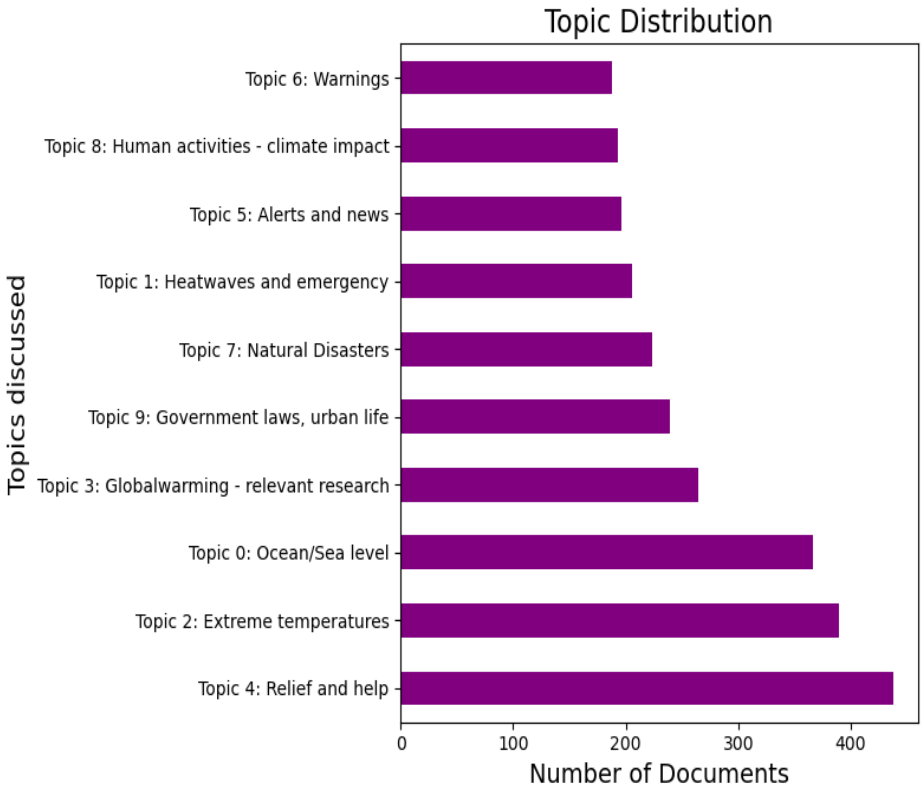

Figure 2. A bar chart showing the distribution of topics being discussed alongside the count.

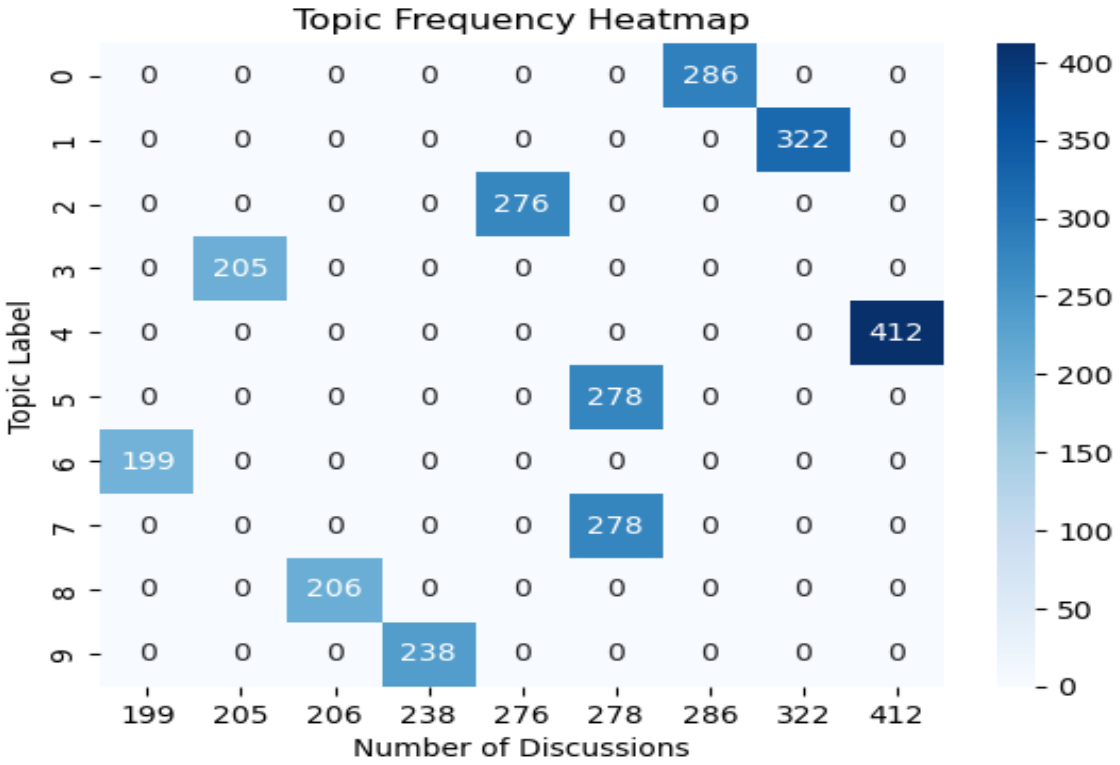

Figure 3. A heatmap showing the distribution of topics being discussed with the count.

When trying to visualize the most frequently discussed topics using word cloud, we see the words ‘climate, ‘change’, ‘sea level’, ‘extreme heat’, and ‘wildfire’ with the highest probability.

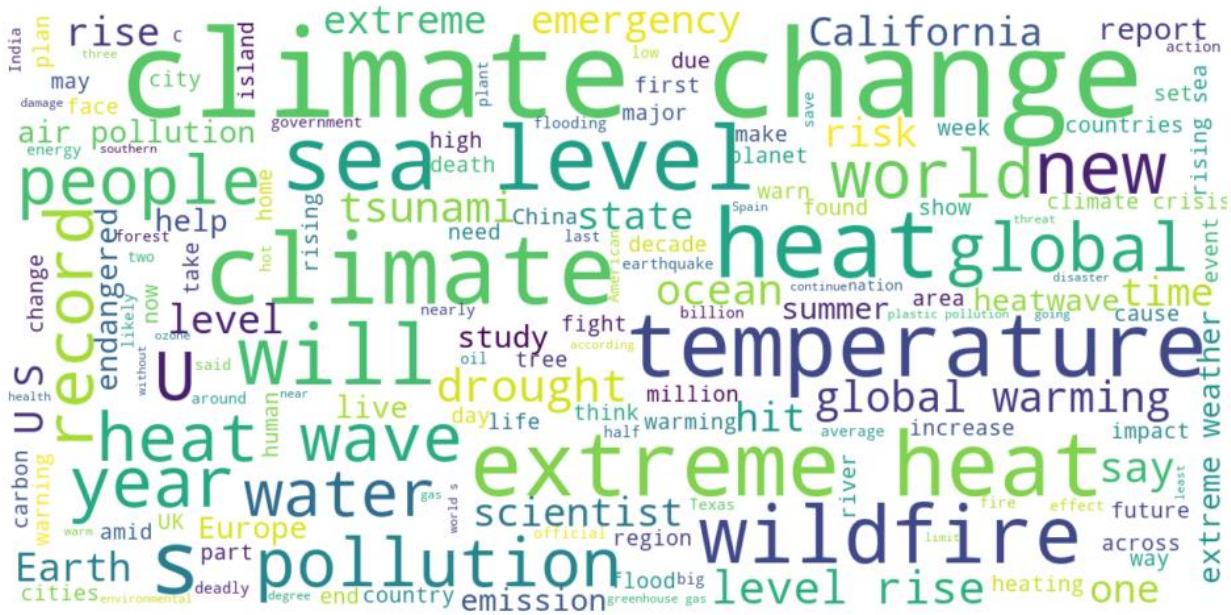

Figure 3. A wordcloud showing the most frequent topics being discussed.

5. CONCLUSION AND LIMITATIONS

Based on the word cloud and topic distribution analysis, we can conclude that the discussions on Reddit mainly revolve around seeking help or offering relief solutions, with a secondary focus on extreme temperatures and heat waves. These findings show that people are likely to find solutions to problems and help others rather than warn them about nearing issues. However, it is worth noting that the discussion may not be balanced or representative of the entire population, and there may be other issues or topics not being discussed as frequently. To improve the analysis, we could consider conducting sentiment analysis to determine the overall sentiment of the discussions related to different topics. Additionally, we could also include the comments section and analyze them to get a more comprehensive view of the topics being discussed. That said, our current approach has limitations like its inability to capture the nuances in the language used in the submissions and the risk of missing out on important keywords that may be indicative of a certain topic.

As a future implementation, the GSDMM algorithm could be applied to topic modeling on Reddit titles. The GSDMM algorithm is a probabilistic clustering algorithm that can be used to discover latent topics in text data. The Reddit platform has a vast amount of text data that could benefit from topic modeling techniques. R. Sharma and V. Srivastava in their paper titled "Discovering Latent Topics in Social Media using Gaussian LDA and the GSDMM Algorithm", proposed a

hybrid approach that combines Gaussian LDA with the GSDMM algorithm for identifying latent topics in Reddit posts. This approach outperforms other topic modeling techniques in terms of topic coherence and interpretability and provides a more fine-grained analysis of the topics discussed in social media[3]. By applying GSDMM to Reddit titles, it is possible to identify hidden topics and clusters within the data that are not immediately apparent. One potential challenge in applying GSDMM to Reddit titles is the need to preprocess the text data properly to ensure the algorithm can effectively identify topics. Additionally, the hyperparameters of the algorithm, such as the number of topics and the concentration parameters, would need to be tuned to achieve the best results. If implemented successfully, the use of GSDMM for topic modeling on Reddit titles could lead to a better understanding of the topics and discussions taking place on the platform. The potential practical applications of this approach are numerous, including but not limited to targeted advertising and climate apocalypse alerts.

REFERENCES

[1] Zander, K. K., Rieskamp, J., Mirbabaie, M., Alazab, M., & Nguyen, D. (2023). Responses to heat waves: What can Twitter data tell us? Natural Hazards, 1-21. https://doi.org/10.1007/s11069-023-05826-0

[2] Pearce, W., Niederer, S., Özkula, S. M., & Lüders, M. (2019). The social media life of climate change: Platforms, publics, and future imaginaries. Wiley Interdisciplinary Reviews: Climate Change, 10(2), e569. https://doi.org/10.1002/wcc.569

[3] Sharma, R., & Srivastava, V. (2019). Discovering Latent Topics in Social Media using Gaussian LDA and the GSDMM Algorithm. In Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 640-643). IEEE.