

Sentiment Analysis of Reddit Data and News Articles On Climate Change

Saketh Balaji Maddineni

*Department of Computer Science,
University of Exeter,
Exeter, UK
Sm1304@exeter.ac.uk*

Dr. Tristan Cann

*Department of Computer Science,
University of Exeter,
Exeter, UK
T.J.B.Cann@exeter.ac.uk*

Prof. Hywel Williams

*Department of Computer Science,
University of Exeter,
Exeter, UK
H.T.P.Williams@exeter.ac.uk*

Abstract—With heat waves and wild forest fires on one side and floods on the other, the world is experiencing an unparalleled climatic crisis. When formulating a strategy to lessen the effects of climate change, officials and researchers should always consider the general population's opinions. Social Media platforms like Twitter, Reddit, etc., provide a real-time snapshot of public opinion. This project seeks to explore how climate change-related events are discussed on Reddit, a community-driven platform, versus traditional news platforms, which often hold political and regional biases. Through sentiment analysis and comparative studies across four different periods (2016, 2020, 2021, and 2022) during which, a great debate regarding climate change happened in recent past. The analysis revealed a pattern of event-driven responses on Reddit, focusing on global incidents and often linking them to environmental impacts with a balanced sentiment distribution, reflecting a diverse and democratic portrayal. At the same time, traditional news articles were more influenced by political contexts and regional factors with a more optimistic perspective and emerging attention to negative aspects. Further analysis of news media by segregating news sources based on political alignment showed increased negative articles on both sides and right-leaning outlets publishing higher negative articles on climate change. Also observed intriguing human behavior on Reddit with similar reactions during similar events during different periods. The research emphasizes the need for an integrated analysis that considers multiple viewpoints by highlighting the complexities and biases within different media landscapes. The insights derived serve as valuable guidance for policymakers, academics, and practitioners aiming to foster more responsive climate strategies, thereby enriching our understanding of the complex narratives surrounding climate change.

I. INTRODUCTION

In an age characterized by swift technological advancements and global interconnectedness, climate change remains an unparalleled challenge, manifesting through extreme weather events and environmental degradation. As our world grapples with the harsh realities of climate change, with heatwaves, wildfires, floods, and a host of other environmental crises, the discourse surrounding it has never been more pertinent, which not only influences public policy and social behavior but also provide a glimpse into the shared values and concerns of communities worldwide. With its far-reaching impacts on ecosystems, economies, and societies, understanding and creating policies that will help mitigate climate change requires scientific inquiry and a robust understanding of public sentiment and opinion.

Climate policies often involve complex decisions and trade-offs that necessitate public support for effective implementation. Traditional methodologies for gathering public sentiment often fail to capture trends and the multidimensional nature of public attitudes, resulting in a gap in understanding that might impede policy design and implementation. This study undertakes a comprehensive sentiment analysis of climate change discourse on platforms like Reddit and news articles, serving as a mirror to the collective conscience. It aims to bridge the gap between public sentiments and policymaking by delving into conversations about climate change, environment, and sustainability issues across significant events and time frames in recent years. It also shows how climate change-related events are portrayed and perceived across different media platforms. The analytical approach provides an invaluable compass for policymakers and stakeholders, illuminating a path that resonates with the prevailing public opinion.

Social media platforms like Reddit serve as robust barometers of public sentiment Kirilenko et al. [1], often providing real-time insights into the perceptions, attitudes, and responses of diverse demographic groups. Conversely, news articles offer a more structured and professional perspective. Understanding these divergent portrayals can offer valuable insights into public sentiment and help inform more effective climate strategies. In essence, the sentiments reflected in these two sources offer a holistic view of the societal narrative around climate change.

In this study, Valence Aware Dictionary and Sentiment Reasoner (VADER) , an insightful lexicon-based sentiment analysis tool Hutto et al. [2], is used at both the sentence and article/post level, depending on the length of the article/post, to capture a more nuanced view of sentiments than a straightforward post-level analysis. This dual-level application represents one of the novel technical contributions of this study. This study performs a systematic sentiment analysis based on the discussions around climate change on Reddit and Traditional News platforms across four distinct periods: the 2016 and 2020 US Presidential Elections, COP26 in 2021, and COP27 in 2022. Also, this study mainly focuses on eight subreddits and news articles from various sources during those four critical periods. These forums and periods were chosen as they represent a broad spectrum of perspectives and events that ignited intense climate change debates in

the recent past. Through sentiment analysis and comparative studies, the project seeks to uncover the patterns and trends in how climate change is discussed across different media. Evaluate the sentiment distribution among Reddit and news platforms, offer a balanced and nuanced perspective, and examine any underlying pattern in both the Media sources. By integrating insights from sentiment analysis, social media analytics and traditional news analysis, the study contributes to a comprehensive and nuanced understanding of public sentiment toward climate change.

In essence, this endeavor goes beyond the confines of conventional sentiment analysis, offering a holistic and evolving perspective on public sentiment regarding climate change. Its insights are invaluable for researchers, decision-makers, and all participants in the climate conversation. This work not only aids in the overarching battle against climate change but also fosters a connection between public sentiment and policymaking. This paves the way for a more agile, inclusive, and effective strategy to address one of the paramount concerns in contemporary times.

II. BACKGROUND

Climate change stands as a paramount global concern in today's age. The public's views and feelings about this matter significantly influence policy directions and heighten awareness of the subject. Research by Leiserowitz et al. [3] analyzes the connection between public risk perception and policy preference regarding climate change. The research emphasizes the importance of understanding these perceptions in formulating effective and publicly resonant policies. In recent years, social media platforms like Reddit have become crucial hubs for climate change discourse, with an extensive user base discussing this pressing topic daily. Kirilenko et al. [1] employed text-mining to gauge climate sentiments on Twitter, shedding light on the multifaceted viewpoints surrounding the issue. This work offers a foundational method for sentiment discernment on digital platforms. Pearce et al. [4] explored the social media life of climate change, examining how platforms and publics contribute to shaping future imaginaries, providing a nuanced perspective on how social media contributes to the broader climate change narrative. Falkenberg et al. [5] discovered a growing polarization around climate change on social media platforms. This paper adds a new dimension to the discourse, emphasizing the role of social media in creating distinct and often conflicting public opinions on climate change. Newman et al. [6] examined how social media platforms frame climate change narratives, revealing that platforms like Facebook host more polarized discussions. They emphasized the need to consider platform-specific characteristics in climate change communication. When we consider platforms like Reddit in which the character limitations are non-existent when compared to that of Twitter, which allows for more in-depth discussions on Reddit. Treen et al. [7] delved into the nuances of climate change conversations on Reddit to understand their nature. Their investigation revealed a mix of polarized viewpoints and constructive dialogues.

Auer et al. [7] stated the importance of utilizing analytical tools like sentiment analysis to social media content on climate change. Barachi et al. [8] proposed a novel sentiment analysis framework to monitor evolving public opinions in real time, focusing on climate change. This has profound implications for understanding immediate reactions to climate-related events. Rodríguez-Ibáñez et al. [9] provided a comprehensive review of sentiment analysis from social media platforms, including its application to climate change discourse, offering a foundational understanding of the methodologies employed in sentiment analysis. Dahal et al. [10] implemented topic modeling and sentiment analysis of global climate change tweets. Their work offers insights into how global perceptions are shaped and vary across regions. By analyzing tweets, they could identify specific trends and topics that dominate the climate change conversation on Twitter. Such an approach allows for more targeted analysis and can highlight specific areas of concern or interest in the global discourse. Zhan et al. [11] applied sentiment analysis to gauge perspectives on climate change through Twitter. Their study revealed that the emotions and sentiments conveyed in climate change-related tweets are intricately linked, offering a method to discern attitudes about the issue.. Ray et al. [12] implemented neural network-based sentiment analysis of Reddit data to analyze and predict the public attitude toward climate change during adverse climate events, which predicted the public attitude being negative during extreme weather events.

In contrast, Sham et al. [13] worked on "Climate Change Sentiment Analysis Using Lexicon, Machine Learning and Hybrid Approaches." They demonstrated how lexicon-based approaches could accurately classify sentiment by relying on a predefined set of words and expressions with associated sentiment values. Such lexicon-based methods effectively identify the general sentiment towards climate change but can be limited by language and context variations. Hamilton et al. [14] conducted an innovative approach to inducing domain-specific sentiment lexicons from unlabelled corpora. This method is particularly relevant for climate change, where specialized lexicons are essential for accurately capturing sentiment to understand attitudes towards climate change better.

Not only social media, traditional media sources also shape public perceptions. Boykoff et al. [15] revealed a false balance in climate change coverage that gave denialist voices excessive airtime after analyzing a decade's worth of news reporting, opening a debate on media responsibility in climate change discourse. Feldman et al. [16] analyzed how cable news frames climate change, discovering that different channels manifest varied biases and influences in public understanding of the issue. This serves as an example of the intricate interaction between the media and public perception. While comparing both, Schmidt et al. [17] identified substantial differences between traditional news media and social media in the focus and sentiment of climate change discussions. Their comparative study illustrates the increasingly divergent narratives between these platforms.

The methodology and approach of this project build upon

extensive research across various domains. By performing an integrated analysis that considers multiple viewpoints, highlighting the complexities and biases within different media landscapes, and comparing community-driven platforms like Reddit with traditional news outlets this study offers a fresh perspective on the democratic nature of public discourse around climate change. The project also stands out in its deep analysis across multiple time frames, reflecting the evolving nature of climate change discussion and providing a more comprehensive understanding of public sentiment toward climate change by understanding these different media portrayals and the underpinning factors that influence them. Such understanding is vital for fostering more informed, responsive, and effective climate strategies, filling an essential gap in the existing body of research.

III. AIMS & OBJECTIVES

This project aims to illuminate the intricate dynamics of public sentiment towards climate change and sustainability, leveraging both social media and traditional news sources. The project strives to provide actionable insights for policymakers, environmentalists, and other stakeholders by understanding these sentiments. To achieve this broad aim, the project is guided by the following specific objectives:

Objective 1 - Sentiment Analysis Across Platforms:

- Aim:** To conduct a comprehensive sentiment analysis to identify trends, patterns, and correlations between platforms and political alignments.
- Research Question:** How do sentiments differ between social media (Reddit) and traditional news media, and what are the underlying influences shaping these sentiments?

Objective 2 - Temporal Analysis:

- Aim:** To analyze sentiment over four distinct periods to understand how public opinion has evolved in relation to major political events and climate change discourse.
- Research Hypothesis:** There exists a consistent pattern in sentiment across different periods, reflecting the broader socio-political climate.

Objective 3 - Patterns in Sentiment Analysis of Reddit Across Time-Periods :

- Aim:** To explore any underlying similarities in Reddit sentiment across different time periods of the study.
- Research Question:** Are there any similarities in Reddit sentiment during similar events across periods?

Objective 4 - Correlation with Political Alignments:

- Aim:** To explore how left-leaning and right-leaning networks differ in their sentiment towards climate change and related policies.
- Research Question:** What trends can be observed in sentiment across political alignments, and how have they diverged or converged over time?

Objective 5 - Practical Implications:

- Aim:** To translate the insights from the analysis into practical strategies and interventions that resonate with public sentiment.
- Requirement:** The findings must be actionable, guiding the design of regulations, interventions, or campaigns that align with prevailing public opinion.

IV. EXPERIMENT DESIGN AND METHODS:

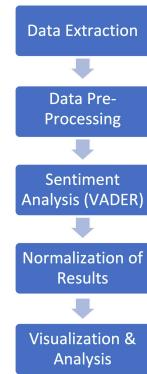


Fig. 1. Steps involved in the study

Methodology: Analyzing Sentiments of Climate Change Discussions on Reddit.

Data Extraction: To comprehensively analyze climate change discussions on Reddit, we utilize Pushshift API to extract the Reddit posts from 8 different subreddits (“Global warming”, “climate”, “climate change”, “climate policy”, “environment”, “politics”, “sustainability”, and “world news”) are selected. These are chosen as the majority of climate change discussions happen in these forums (Treen et al. [18]) on Reddit. Using Push shift API an open-source dataset that enables researchers to delve into Reddit’s rich user-generated content (Baumgartner et al. (2020) [19] historical Reddit posts’ key features are retrieved for four periods: the 2016 and 2020 US Presidential Elections, COP26 in 2021, and COP27 in 2022. Then using, PRAW “Python Reddit API Wrapper” (Boe et al. [20]), which allows for simple access to Reddit’s API to extract the key features in the comments based on the “Post_id” extracted from the Push shift API. Then both the comments and posts are combined.

Using News API [21] extracted the news from various sources for the Study Period using keyword filtering. Keywords used are Climate change, Global warming, Carbon emissions, Renewable energy, Sustainability, Climate Policy, Climate action, Paris Agreement, Carbon footprint, Greenhouse gas emissions, Climate adaptation, Renewable energy policies, Fossil fuels, Climate Justice, Environmental activism as these are the words that are frequently used in the discussions related to climate change.

Filtering and Extraction: At this stage, the extracted News and Reddit data was structured, and removed the data

that has no values in the “Title” and “Text” columns. Reddit data was again filtered to remove any unwanted data outside the study’s date range. News data is filtered to remove the articles that are not in English and any unwanted data that are outside the date range for the study. Finally, there are 80207, 73716, 103280, and 98161 Reddit posts and 128912, 254012, 485945, and 387225 from 2016, 2020, 2021, and 2022 respectively.

During the analysis, News data was further divided into Left-leaning and Right-leaning. Which are categorized as below based on articles by Sanders [22] and Grieco [23] in “Today.yougov.com” and “Pew Research center” respectively.

Selected News Sources: The primary reason behind selecting below news sources in each category was to capture a broad and balanced view of the political spectrum in media coverage of climate change. The selected sources represent well-known media outlets widely recognized for their political alignments, per the articles cited above.

Left-Leaning Sources:

- **“The New York Times”:** Known for its comprehensive coverage and editorial stance, often perceived as liberal or progressive.
- **“Mother Jones”:** Recognized for its investigative journalism and often associated with progressive politics.
- **“CNN”:** Although aiming for neutrality, it has been critiqued for a liberal bias in its opinion pieces and commentaries.
- **“New York Post”:** Has a varied readership but includes left-leaning columns and perspectives.
- **“The Guardian”:** Internationally recognized for a progressive or liberal viewpoint, particularly on environmental issues.

Right-Leaning Sources:

- **“FoxNews”:** Widely recognized for conservative perspectives, especially in its opinion programming.
- **“NewsMax”:** Known for its conservative stance and support of right-wing political ideologies.
- **“National Review”:** A conservative publication focusing on political commentary and opinion.
- **“Breitbart”:** Often associated with right-wing populism and known for conservative viewpoints.
- **“Washington Examiner”:** A political journalism website and weekly magazine known for its conservative writing.

Sentiment Analysis: The core of our analysis involves determining the sentiments of the collected texts. We employ the VADER SentimentIntensityAnalyzer, a lexicon, and rule-based sentiment analysis tool developed for social media text. This choice is motivated by its ability to handle informal language, prevalent in Reddit discussions, and its effectiveness in identifying nuanced sentiments.

After analyzing the sentiment of a text, VADER produces the polarity scores as Positive, Negative, Neutral, and Compound, where positive, negative, and neutral scores indicate the ratio of the proportion of text that fall under each category.

The compound score was determined by summing up the value score of every word in the lexicon. This was then adjusted to fall within a range of -1 (highly negative) to +1 (highly positive). So, to classify our text, first, we need to decide on the thresholds; for this, multiple experiments are conducted by adjusting the neutral threshold. A few of the values used for the experiments are as follows, and these are the lower limit and upper limit of the neutral threshold [(-0.5 to 0.5), (-0.2 to 0.2), and (-0.05 to 0.05)]. Each threshold was tested on a sample data of 1000 news articles. Once sentiment is analyzed out of 1000, 100 articles are randomly selected. Without knowing how VADER classified the sampled 100 articles, they were manually classified to avoid bias and compared with the results from VADER. The first two sets produced high neutrals and accuracy rates of approximately 45% (-0.5 to 0.5) and 56%(-0.2 to 0.2) approximately on sample data. With the standard values for the threshold (-0.05 to 0.05), we got an accuracy of approximately 65% on the sample news data. So, in this study, the standardized thresholds to classify the text are used, which are as follows:

- Positive sentiment: compound score ≥ 0.05
- Neutral sentiment: (compound score ≥ -0.05) and (compound score ≤ 0.05)
- Negative sentiment: compound score ≤ -0.05

To increase the accuracy, further multiple experiments are conducted. The first was to form a baseline a direct sentiment of the entire news article/Reddit post was analyzed, and each post/article was classified as either Positive, Negative, or Neutral according to the compound score valuation as above, which yielded an accuracy score of approximately 65% as stated earlier.

In the second method, each article was tokenized into sentences, and then each sentence’s compound polarity value was calculated with VADER. It computes the average compound score by dividing the total compound score (sum of compound scores of all sentences) by the total number of sentences, and based on that compound score, sentiment is classified as per the threshold mentioned earlier, which increases the accuracy a little but not as much as hoped.

In the next method, similar to the above method each article was tokenized into sentences, and then each sentence’s compound polarity value was calculated with VADER. But instead of computing the average compound score in this method the sum of positive, negative, and neutral compound values (Intensity of each classification) was computed. If a compound value is greater than 0.05, it’s considered positive; if it’s less than -0.05, it’s negative; otherwise, it’s neutral. Then weighted average for each category is computed and based on the absolute value of the weighted average,i.e. whichever classification has the highest absolute weighted average value, the sentiment of the article is classified. This approach worsened the accuracy, so it was discarded.

In final method which was used in this study. First, the posts/articles are divided into two sets:

- For texts containing one or two sentences (Text with \leq

2 Sentences), we applied sentiment analysis directly to the entire text using VADER and classified the sentiment based on the threshold decided. This approach is suitable for shorter texts, where sentence-level sentiment analysis may not yield meaningful results due to limited context.

- For texts containing more than two sentences (Text with ≥ 2 Sentences), We adopted a sentence-level sentiment analysis approach. We first tokenized the text into individual sentences using the NLTK library's `sent_tokenize` function. Then applied, the VADER SentimentIntensityAnalyzer and classified each sentence separately. Then by aggregating the sentiment classification of individual sentences, i.e., counting the number of positive, negative, and neutral sentences and whichever was the highest, that was the sentiment of that post/article. For example, suppose an article has 10 sentences, out of which 5 are classified as positive. In that case, 3 are classified as neutral, and 2 are classified as negative, then the whole article is classified as Positive. In the case of ties, the article was classified as neutral.

This approach enables a more nuanced analysis of longer texts with varying sentiments across sentences and produced approximately 70% accuracy in the tests on both Reddit and News dataset samples. Based on this method, sentiment analysis on whole dataset was conducted.

Data Normalization and Aggregation: Once the data's sentiment was analyzed to gain insights, we normalize and aggregate the sentiment analysis results. We group the sentiment analysis results based on the '`Date_created`' and '`Sentiment`' columns. This step allows us to organize the data by date and sentiment type (positive, negative, or neutral), facilitating further analysis. Then we calculated the total number of posts (or comments) made on each date by summing up the counts of positive, negative, and neutral sentiments. This metric serves as the denominator in calculating the percentage values of each sentiment type. This normalization procedure ensures a fair comparison of sentiment proportions across different dates, despite potential variations in the total number of posts.

Data Visualization: Data visualization plays a crucial role in conveying the analysis results effectively. We generate a sentiment plot to visualize the percentage of positive, negative, and neutral sentiments over time for the Reddit posts and News articles in 2016, 2020, 2021, 2022. The horizontal axis displays the dates, whereas the vertical axis indicates the sentiment percentage values. The plot provides a clear and intuitive representation of the sentiment trends, enabling us to observe fluctuations and patterns in climate change discussions on Reddit and News articles during the specified period. For example: Consider figure 2, It depicts the sentiment changes for all 3 classifications overtime in 2020 for Reddit data, it clearly shows the fluctuation in reddit sentiment and also how classifications are distributed in 2020

for Reddit data. Similar graphs for news data are given can be seen in figures 13 to 16 in Appendix section.

Upon visualizing the sentiment analysis results, we proceed with an in-depth interpretation and analysis of the findings. We aim to identify significant trends, sudden shifts in sentiment, and notable events that may have influenced public sentiments during the given time frame. By delving into the patterns, we can draw valuable insights regarding the attitudes and emotions related to climate change on Reddit and News.

For this, we have utilized word clouds to check on the discussions on both platforms during the days that show exciting patterns in the line graph. We have also utilized the panda's in-built visualizing capabilities to identify the overall trends in Reddit and News. Through these techniques, we have found some interesting insights that will help achieve the objectives of this study, which are explained in the following section.

V. RESULTS

Based on the above methodology, We have conducted an exhaustive sentiment analysis across two major platforms: News media and Reddit, focusing on the public and media responses to climate change. Below are the findings. The analysis revealed a distinct response pattern, with Reddit being event-centric and news articles more influenced by political and regional factors. This disparity can be observed across different periods of the study.

In October 2020, a clear divergence was noticed between Reddit and news articles' focus. From figure 2. A spike in negative sentiment on Reddit can be observed on 31st October 2020 , with posts primarily centered around the earthquake in Turkey and the war between Azerbaijan and Armenia. These posts often linked the events to environmental impacts or suggested climate change as a potential cause for the earthquake. This focus was visually apparent in the word cloud of negative Reddit posts from that day 3. Contrastingly, news articles from the same day were primarily concerned with forest fires' damage and political responses in the United States, with no significant mention of the global events that dominated Reddit. This can be seen in the word cloud of News articles figure 4 from that day.

A similar pattern emerged on 29th October 2016, where Reddit exhibited the highest percentage of negative posts (36.05%), can be observed from the figure 5, reacting to the North Dakota oil pipeline protests and the government's stern response. The word cloud figure 6 from that day highlights this focus. Conversely, news articles were more engrossed in stories surrounding the Presidential elections, mentioning the pipeline protests only tangentially, which is clear from the figure 7. This difference underscores news platforms' inclination towards the contemporary political climate rather than the event-driven response seen on Reddit.

This analysis demonstrates a unique dichotomy in the way climate change-related events are portrayed on Reddit versus

traditional news platforms. While Reddit's response is immediate and specific to global incidents, news articles appear more influenced by political contexts.

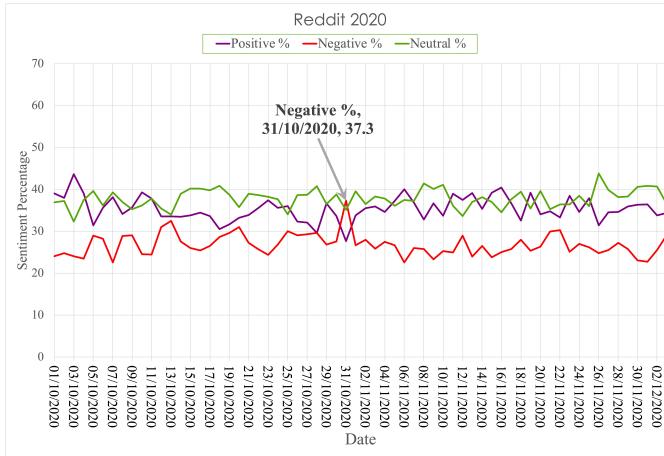


Fig. 2. Reddit 2020 Sentiment percentage overtime for each classification

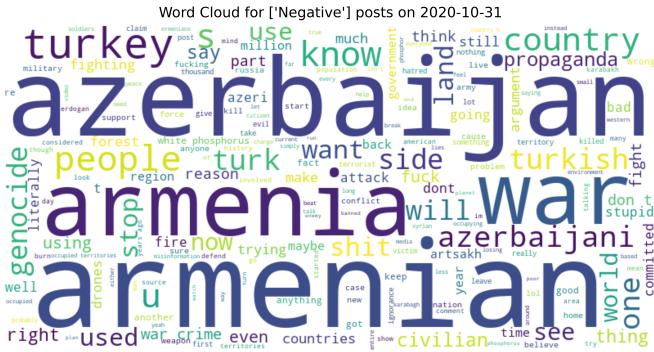


Fig. 3. Word Cloud for Negative Reddit Posts on 2020-10-31

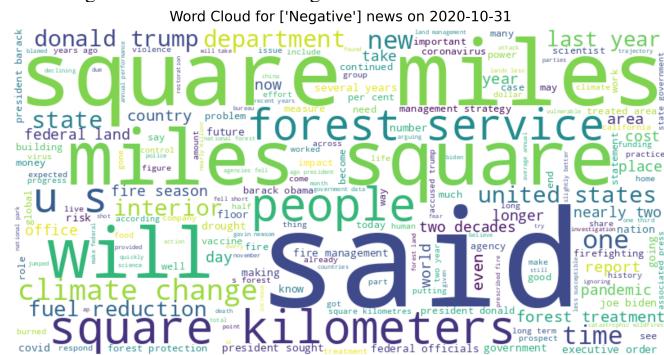


Fig. 4. Word Cloud for Negative News articles on 2020-10-31

When observing the figure 5 Reddit 2016 Sentiment percentage overtime and similar to one of News 2016, 13 in the Appendix, we can see the distribution of the sentiment percentages variations. To analyze this further, a comparative analysis across four distinct periods (2016, 2020, 2021, and 2022) was conducted, and that brings intriguing patterns in Reddit and News platforms, offering insights into different media landscapes surrounding climate change issues.

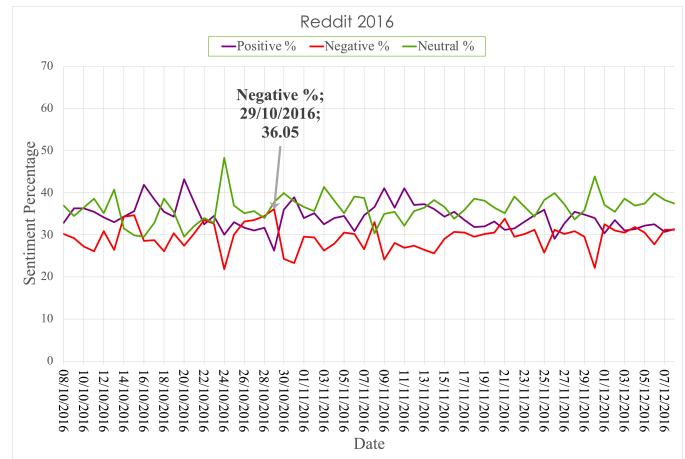


Fig. 5. Reddit 2016 Sentiment percentage overtime for each classification

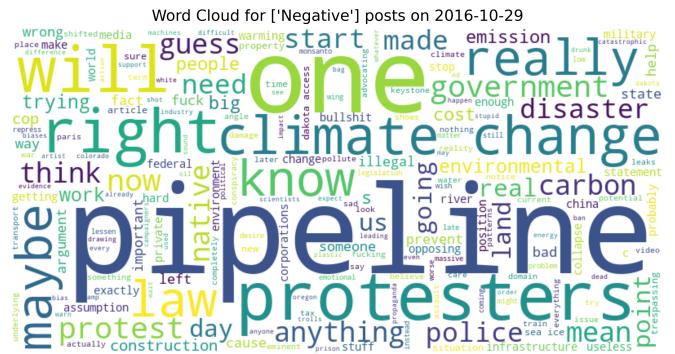


Fig. 6. Word Cloud for Negative Reddit posts on 2016-10-29



Fig. 7. Word Cloud for Negative News articles on 2016-10-29

Reddit's sentiment analysis exhibits a more balanced distribution among Positive, Neutral, and Negative sentiments. The average split across the four years is approximately 35% Positive, 35% Neutral, and 30% Negative, which can be observed from figure 8 and table I

This distribution showcases Reddit's diverse response, capturing a range of opinions and discussions. Reddit's slightly higher Neutral sentiment, reflecting more nuanced or informational posts, indicates a community engaged in substantive discussion and debate. The platform's democratic nature, allowing users from various backgrounds to voice their opinions,

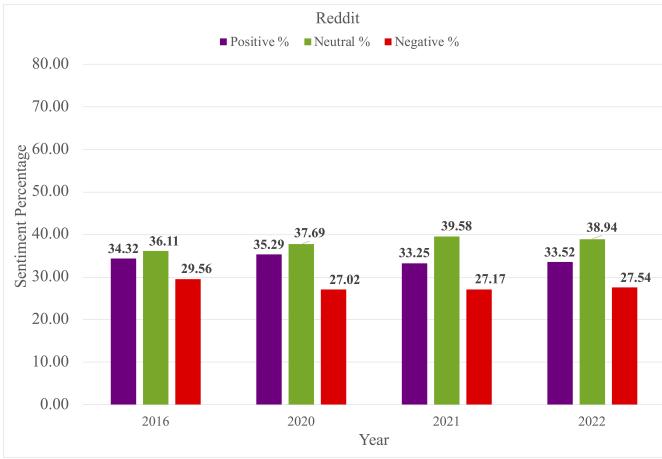


Fig. 8. Reddit overall sentiment distribution

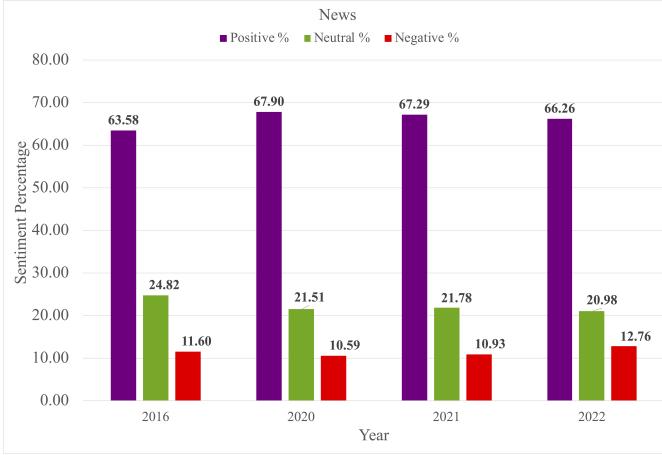


Fig. 9. News overall sentiment distribution

TABLE I
SENTIMENT DISTRIBUTION OVER YEARS IN REDDIT

Year	Positive (%)	Neutral (%)	Negative (%)
2016	34.32	36.11	29.56
2020	35.29	37.69	27.02
2021	33.25	39.58	27.17
2022	33.52	38.94	27.54

may lead to this more balanced distribution. Interestingly, the consistently high Neutral and Negative percentages reflect ongoing skepticism or concerns regarding climate change solutions, policies, or the very acknowledgment of the problem. The prevalence of Neutral sentiment suggests a broader array of topics, more factual or scientific.

Within the News category, the trend reveals a dominance of Positive sentiment, with a noticeable increase in Negative sentiment over time. The average split in News sentiment is approximately 65% Positive, 25% Neutral, and 10% Negative, which can be observed from table II and figure 9.

The detailed analysis underscores distinct approaches between Reddit's more democratic and multi-faceted portrayal and the news media's generally optimistic perspective on climate change, with emergent attention to negative aspects.

TABLE II
SENTIMENT DISTRIBUTION OVER YEARS IN NEWS

Year	Positive (%)	Neutral (%)	Negative (%)
2016	63.58	24.82	11.90
2020	67.90	21.51	10.59
2021	67.29	21.78	10.93
2022	66.26	20.98	12.76

The findings enrich our understanding of public sentiment, emphasizing the need for an integrated analysis that recognizes the complex narratives surrounding climate change. The contrast between the platforms serves as valuable insight for policymakers, academics, and practitioners aiming to foster more informed and responsive climate strategies.

To acquire a better understanding of the shift, an increase in negative articles in news sources, news data was sampled into left-leaning and right-leaning groups as mentioned earlier. These sources were chosen to provide a rich, diverse, and balanced representation of the media landscape. By including prominent media outlets known for their distinct political biases, the analysis aimed to uncover how political alignment might shape sentiment toward climate change, offering a more nuanced understanding of the multifaceted nature of climate discourse in the media. The in-depth analysis of sentiments within left and right-leaning sources depicted by the plot in Figure 10 and in Tables III and IV, unveils the complexity of climate change discourse in the media landscape. The nuanced changes in sentiment over time within each political alignment provide a rich understanding of how climate change is perceived, debated, and reported across different media spectrums. The emerging pattern of increased negativity across both political alignments reflects a shared concern or criticism but is likely driven by different underlying factors or ideologies. Right-leaning sources consistently showed higher negative percentages over the analyzed periods. This could reflect different stances on environmental policies, international agreements, or societal values that underpin the discourse on climate change.

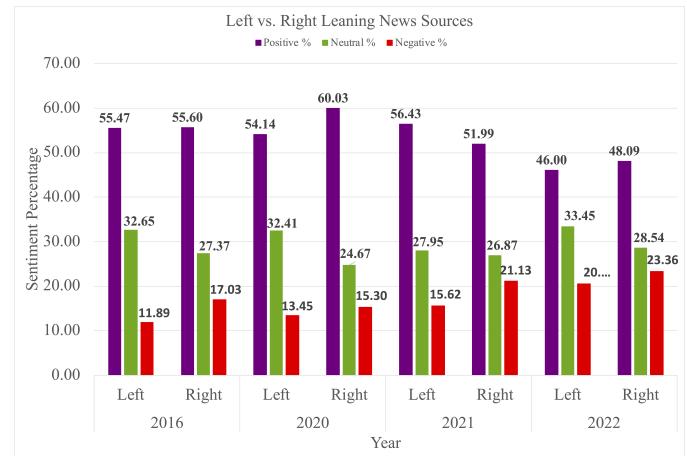


Fig. 10. Left vs. Right Leaning News Sources Sentiment Distribution

TABLE III
LEFT-LEANING NEWS SOURCES OVERALL SENTIMENT DISTRIBUTION
OVERTIME

Year	Positive (%)	Neutral (%)	Negative (%)
2016	55.47	32.65	11.89
2020	54.14	32.41	13.45
2021	56.43	27.95	15.62
2022	46.00	33.45	20.55

TABLE IV
RIGHT-LEANING NEWS SOURCES OVERALL SENTIMENT SENTIMENT
DISTRIBUTION OVERTIME

Year	Positive (%) Positive %	Neutral (%) Neutral %	Negative (%) Negative %
2016	55.60	27.37	17.03
2020	60.03	24.67	15.30
2021	51.99	26.87	21.13
2022	48.09	28.54	23.36

An intriguing consistency emerges across different periods when similar events occur on Reddit. For instance, as the COP26 summit commenced on 31st October 2021, there was an evident spike in negative posts on Reddit. Interestingly, this wasn't an isolated instance. When the following year's COP27 began on 6th November 2022, a similar upswing in negative sentiments was detected, as visualized in figures 11 and 12. This consistency across years points to the possibility that global events like the COP summits stir particular emotions or concerns in the Reddit community. A parallel consistency was observed during the US Presidential elections of 2016 and 2020, as showcased in figures 5 and 2. Around the election days, which fell on 8th November 2016 and 4th November 2020 respectively, there was a discernible decrease in negative posts and a concurrent increase in positive sentiments. This shift might suggest an underlying hope or optimism associated with the potential for political change or a new direction in climate policy. Such consistent sentiment responses to globally significant climate events underscore Reddit's capacity as a live barometer of public sentiment. Its community-driven nature allows for the instantaneous capturing of public emotions, reflecting the platform's role as a social media tool and a significant data source for understanding public sentiment during global climate events. The predictability of these sentiment shifts offers valuable insight for stakeholders, from policymakers to researchers, into the collective consciousness of the global online community during pivotal moments in climate discourse.

VI. DISCUSSION

The central theme of our investigation lies in deciphering the sentiment variations concerning climate change discussions across diverse mediums, specifically Reddit and traditional news platforms. Our findings vividly highlight a distinguishing pattern: While sentiments on Reddit predominantly resonate with specific global incidents irrespective of the ongoing news narrative, mainstream news narratives, Conversely traditional

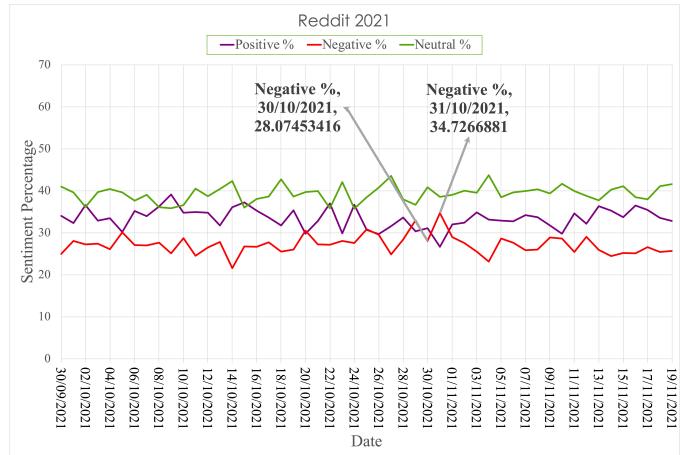


Fig. 11. Reddit 2021 Sentiment percentage overtime for each classification

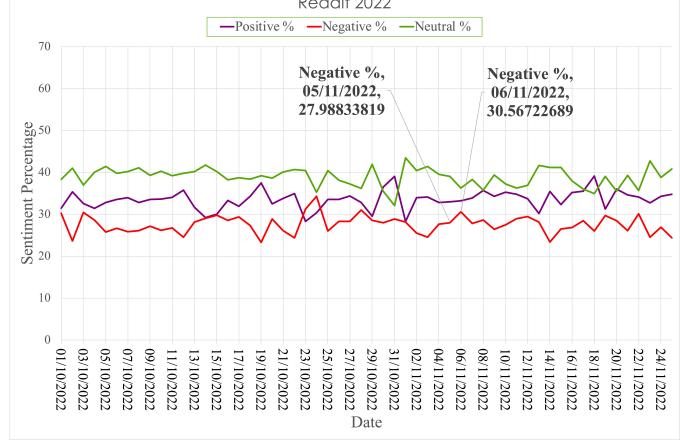


Fig. 12. Reddit 2022 Sentiment percentage overtime for each classification

news platforms often appear influenced by the extant political environment.

Study's observations echo findings from Kirilenko et al. [1], who analyzed climate change discourse on Twitter and found that such discussions are mainly reactive to current events. This reinforces the idea that real-time platforms, like Reddit and Twitter, serve as immediate barometers for public sentiment in response to global incidents. On the other hand, our insight about the politicization of climate discourse in mainstream media mirrors the results from Boykoff et al. [15], who found that U.S. prestige-press coverage of global warming often frames it in a political context. Our findings are congruent with this observation, reinforcing that such biases often eclipse significant global occurrences pivotal to climate discussions.

The broader implication of our research is its potential to shape climate policy. Understanding sentiment dynamics on social media platforms like Reddit can offer insights into public perception, informing policymakers, and recognizing political underpinnings in news narratives can help craft more informed and resonant public communications. The major challenge for the researchers going forward would be data

availability, as the landscape of data extraction from social media platforms has been transforming. Recent restrictions and policy shifts by platforms such as Twitter and Reddit pose a challenge to the academic and research community. However, researchers can overcome this challenge in a few ways, like utilizing new emerging media platforms or working on an extension of previous studies, like utilizing one of the findings from this study that similar patterns can be observed across different periods when similar events occur—granted these methods have limitations, like not as many discussions happening on the new platforms, etc.

Our study as well has limitations. The predominant focus on English-based sources might introduce a linguistic bias, neglecting non-English sentiments and perspectives as VADER is an English-based lexicon. Additionally, by concentrating mainly on a few subreddits and eliminating the posts that have only images, which have become the primary form of showing emotions in social media in recent days, we may have overlooked influential narratives. Current sentiment analysis tools including VADER used in this study doesn't take context in the text into consideration to improve on this future studies can use the Context sensitive sentiment analysis for short-form text (CIDER) depicted in Young et.al [24].

Even with these limitations, the insights of this study have substantial implications not only for researchers and policy-makers, who can develop this even further by converting the scope of the study from a global scale to a specific region and building policies and making decisions, but also to industries pivoting towards sustainable initiatives, understanding these sentiments can guide public engagement strategies. Furthermore, tech enterprises could glean the value of real-time sentiment analysis, broadening its applications from market analytics to predictive modeling.

VII. CONCLUSION

In an era where public discourse on climate change is increasingly pivotal, this study aimed to comprehend the interplay of sentiments across diverse platforms – specifically Reddit and traditional news articles. Through rigorous sentiment analysis, our project unveiled a unique dichotomy: while Reddit sentiment often responded dynamically to global events, traditional news articles were influenced more by regional political climates.

Our multi-dimensional approach, which seamlessly integrated real-time social media data with conventional news analysis, highlighted this disparity, drawing attention to the broader intricacies in portraying climate change narratives. By employing the Valence Aware Dictionary and sentiment Reasoner (VADER) at both the sentence and post levels, our study captured a more layered and nuanced perception of sentiments than conventional methodologies might offer.

The temporal depth of our analysis, concentrating on significant events and periods like the US Presidential Elections of 2016 and 2020, COP26 in 2021, and COP27 in 2022, enabled us to chart the evolving nature of public sentiment

over time. These insights can significantly influence policy-makers, scholars, and the general public, offering a clearer perspective on how climate discourse shapes and is shaped by varying narratives. For example, By performing these kinds of analyses after major policy implementations policymakers can find out about public reaction, and if there are any concerns pre-emptive steps can be taken to address them. Such as creating informative posts or conducting interactive discussions between the public and experts, or just drafting a better policy that is in accordance with public opinion.

Looking ahead, there lies immense potential in expanding this analytical framework that was used in this study. Future endeavors could dive deeper into other social media platforms, investigate the role of influencers in shaping sentiments, or even explore the potential of real-time sentiment analysis in influencing policy decisions. This research can be extended further by focusing on particular geographical locations or particular languages like French, Spanish, Hindi, etc., utilizing the suitable lexicons for that language and media outlets to extract the public discourse specific to that location and extend the accuracy using a context-based lexicon. How these all pan out in a time where access to significant data sources for social discussions is being limited by the profit-driven mentality of corporate heads of these sources is yet to be seen.

The research presented here contributes significantly to the realms of sentiment analysis, climate change discourse, and policy formulation by providing a detailed, multi-dimensional view of public opinion, urging stakeholders to consider public sentiment as an invaluable compass in navigating the complexities of climate change and other global discourses. Finally, as the digital realm continues to shape public opinion, our methodology offers a foundation upon which future research can build, evolve, and offer more insights into the intricate dance of public sentiment in our rapidly changing world.

VIII. DECLARATIONS

Originality Declaration: I hereby declare that the work presented in this report is my own and has not been submitted elsewhere for assessment, publication, or any other purpose. Any material used from other sources, including books, journals, online sources, or any other kind of publication, has been duly cited and referenced in accordance with standard practices. **Ethical Declaration:** This research abides by all ethical standards and guidelines required for research in this field. No personal data was accessed or used without consent, and all the data gathered was treated with confidentiality and respect. I confirm that there were no conflicts of interest, financial or otherwise, that could have influenced the outcomes of this study.

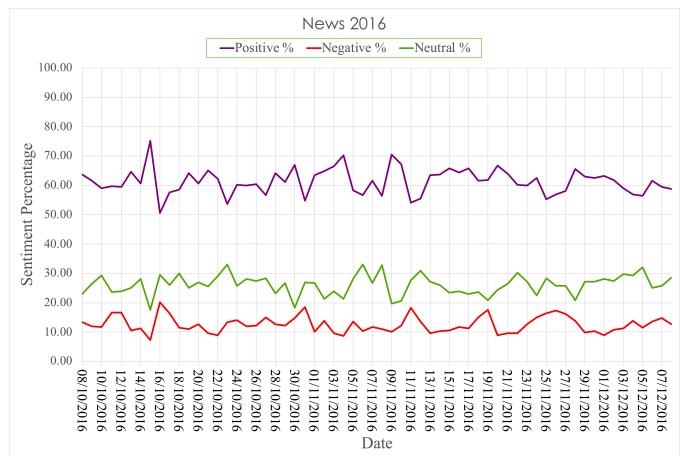
REFERENCES

- [1] A. P. Kirilenko and S. O. Stepchenkova, "Public microblogging on climate change: One year of twitter worldwide," *Global Environmental Change*, vol. 26, pp. 171–182, 2014. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0959378014000375>

- [2] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, May 2014. [Online]. Available: <https://ojs.aaai.org/index.php/ICWSM/article/view/14550>
- [3] A. Leiserowitz, "Climate change risk perception and policy preferences: The role of affect, imagery, and values," *Climatic Change*, vol. 77, no. 1, pp. 45–72, 2006. [Online]. Available: <https://doi.org/10.1007/s10584-006-9059-9>
- [4] W. Pearce, S. Niederer, S. M. Özkula, and N. Sánchez Querubín, "The social media life of climate change: Platforms, publics, and future imaginaries," *WIREs Climate Change*, vol. 10, no. 2, p. e569, 2019. [Online]. Available: <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcc.569>
- [5] M. Falkenberg, A. Galeazzi, M. Torricelli, N. Di Marco, F. Larosa, M. Sas, A. Mekacher, W. Pearce, F. Zollo, W. Quattrociocchi, and A. Baronchelli, "Growing polarization around climate change on social media," *Nature Climate Change*, vol. 12, no. 12, pp. 1114–1121, 2022. [Online]. Available: <https://doi.org/10.1038/s41558-022-01527-x>
- [6] T. P. Newman, E. C. Nisbet, and M. C. Nisbet, "Climate change, cultural cognition, and media effects: Worldviews drive news selectivity, biased processing, and polarized attitudes," *Public Understanding of Science*, vol. 27, no. 8, pp. 985–1002, 2018, pMID: 30253695. [Online]. Available: <https://doi.org/10.1177/0963662518801170>
- [7] M. R. Auer, Y. Zhang, and P. Lee, "The potential of microblogs for the study of public perceptions of climate change," *WIREs Climate Change*, vol. 5, no. 3, pp. 291–296, 2014. [Online]. Available: <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcc.273>
- [8] M. El Barachi, M. AlKhatib, S. Mathew, and F. Oroumchian, "A novel sentiment analysis framework for monitoring the evolving public opinion in real-time: Case study on climate change," *Journal of Cleaner Production*, vol. 312, p. 127820, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0959652621020382>
- [9] M. Rodríguez-Ibáñez, A. Casámez-Ventura, F. Castejón-Mateos, and P.-M. Cuénca-Jiménez, "A review on sentiment analysis from social media platforms," *Expert Systems with Applications*, vol. 223, p. 119862, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417423003639>
- [10] B. Dahal, S. A. P. Kumar, and Z. Li, "Topic modeling and sentiment analysis of global climate change tweets," *Social Network Analysis and Mining*, vol. 9, no. 1, p. 24, 2019. [Online]. Available: <https://doi.org/10.1007/s13278-019-0568-8>
- [11] Z. Mi and H. Zhan, "Text mining attitudes toward climate change: Emotion and sentiment analysis of the twitter corpus," *Weather, Climate, and Society*, vol. 15, no. 2, pp. 277 – 287, 2023. [Online]. Available: <https://journals.ametsoc.org/view/journals/wcas/15/2/WCAS-D-22-01231.xml>
- [12] S. Ray and A. M. Senthil Kumar, "Prediction and analysis of sentiments of reddit users towards the climate change crisis," in *2023 International Conference on Networking and Communications (ICNWC)*, April 2023, pp. 1–17.
- [13] N. Mohamad Sham and A. Mohamed, "Climate change sentiment analysis using lexicon, machine learning and hybrid approaches," *Sustainability*, vol. 14, no. 8, p. 4723, Apr 2022. [Online]. Available: <http://dx.doi.org/10.3390/su14084723>
- [14] W. L. Hamilton, K. Clark, J. Leskovec, and D. Jurafsky, "Inducing domain-specific sentiment lexicons from unlabeled corpora," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Austin, Texas: Association for Computational Linguistics, Nov. 2016, pp. 595–605. [Online]. Available: <https://aclanthology.org/D16-1057>
- [15] M. T. Boykoff and J. M. Boykoff, "Climate change and journalistic norms: A case-study of us mass-media coverage," *Geoforum*, vol. 38, no. 6, pp. 1190–1204, 2007, theme Issue: Geographies of Generosity. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0016718507000188>
- [16] L. Feldman, E. W. Maibach, C. Roser-Renouf, and A. Leiserowitz, "Climate on cable: The nature and impact of global warming coverage on fox news, cnn, and msnbc," *The International Journal of Press/Politics*, vol. 17, no. 1, pp. 3–31, 2012. [Online]. Available: <https://doi.org/10.1177/1940161211425410>
- [17] A. Schmidt, A. Ivanova, and M. S. Schäfer, "Media attention for climate change around the world: A comparative analysis of newspaper coverage in 27 countries," *Global Environmental Change*, vol. 23, no. 5, pp. 1233–1248, 2013. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S095937801300126X>
- [18] K. Treen, H. Williams, S. O'Neill, and T. G. Coan, "Discussion of climate change on reddit: Polarized discourse or deliberative debate?" *Environmental Communication*, vol. 16, no. 5, pp. 680–698, 2022. [Online]. Available: <https://doi.org/10.1080/17524032.2022.2050776>
- [19] *The Pushshift Reddit Dataset*. Zenodo, Jan. 2020. [Online]. Available: <https://doi.org/10.5281/zenodo.3608135>
- [20] B. Boe, *PRAW: The Python Reddit API Wrapper Documentation*, 2012, last modified: 2023, Accessed online: 25-05-2023. [Online]. Available: <https://praw.readthedocs.io/>
- [21] (2023) Newsapi.ai. Accessed: 25-05-2023. [Online]. Available: <https://www.newsapi.ai/>
- [22] L. Sanders. (2022) Trust in media 2022: Where americans get their news poll. Accessed: 25-05-2023. [Online]. Available: <https://today.yougov.com/topics/politics/articles-reports/2022/04/05/trust-media-2022-where-americans-get-news-poll>
- [23] E. Grieco. (2020) Americans' main sources for political news vary by party and age. Accessed: 25-05-2023. [Online]. Available: <https://www.pewresearch.org/short-reads/2020/04/01/americans-main-sources-for-political-news-vary-by-party-and-age/>
- [24] J. C. Young, R. Arthur, and H. T. P. Williams, "Cider: Context sensitive sentiment analysis for short-form text," 2023.

APPENDIX

Graphs of News Sentiment percentage overtime for each classification.



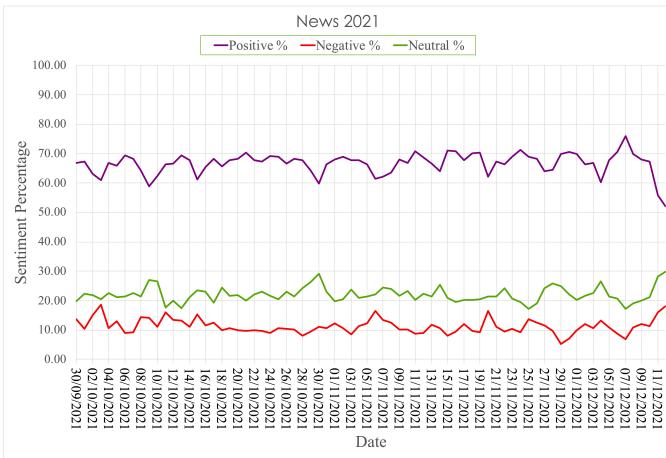


Fig. 15. News 2021 Sentiment percentage overtime for each classification

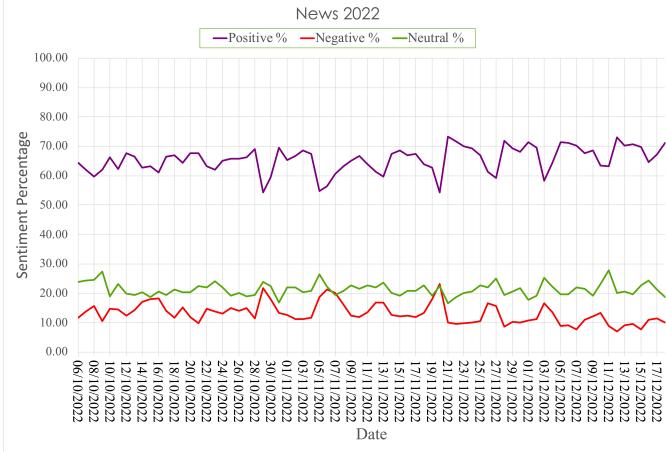


Fig. 16. News 2022 Sentiment percentage overtime for each classification

