Analyzing Discourse and Sentiment Surrounding Social Justice Movements: A Case Study of the Black Lives Matter Movement

A.C. Nanayakkara, B.T.G. Samantha Kumara

Abstract: The Black Lives Matter Movement ignited global discourse on systemic racism and police violence following George Floyd's death on May 25, 2020. This study used advanced Natural Language Processing to analyze BLM-related tweets from the public and fourteen major news channels. It identified key themes, sentiments, and emotional expressions to offer insight into the movement's complexity. Content analysis, sentiment comparison, and textual similarity revealed a mix of neutral and positive sentiments, with significant differences between public and news channel tweets. News channels often presented formal narratives, while public tweets displayed a broader range of emotions. Clustering techniques uncovered themes such as justice, solidarity, and activism, illustrating the movement's diverse nature. The study showed that while BLM's core message focused on social justice, the surrounding discourse varied, with different actors contributing varied narratives and emotions. This research enhanced understanding of social movements and provided empirical evidence on societal attitudes toward BLM. The findings highlighted how social justice movements were portrayed across media and suggested further exploration of these dynamics. Future research could include longitudinal studies and cross-platform comparisons to deepen insights into digital activism's evolving landscape.

***Index Terms*—Black Lives Matter Movement (BLMM); Clustering Techniques; Discourse Analysis; Natural Language Processing (NLP); News Channels; Social Media**

I. Introduction

This study advanced the analysis of BLM discourse during this critical period by employing a novel approach that integrated tweets from both individuals and news channels with advanced unsupervised NLP techniques [1], [2].

Using the GetOldTweets3 tool, BLM-related tweets were collected and categorized into those from individuals and news channels such as The New York Times, BBC, and CNN [3], [4]. The analysis involved identifying prominent words, clustering tweets to reveal patterns, and conducting sentiment analysis, including examining emoji usage to understand emotional expressions [5], [6]. A comparative analysis of sentiment between tweets from news channels and the general public was also conducted. This comprehensive approach provided insights into the diverse perspectives and sentiments within the BLM discourse, highlighted the evolving nature of social movements in the digital age, and contributed to a deeper understanding of contemporary social activism [7], [8].

A.C. Nanayakkara is a postgraduate candidate in Computer Science at the Sabaragamuwa University of Sri Lanka. Her research interests include social networking, data mining, big data, and artificial intelligence. E-mail: chethana@ccs.sab.ac.lk

B.T.G.S. Kumara is a professor at the Sabaragamuwa University of Sri Lanka. His areas of expertise include big data, data mining, and artificial intelligence. E-mail: kumara@appsc.sab.ac.lk

A. Problem Statement

The advent of social media revolutionized the dynamics of social movements, enabling individuals, communities, and media outlets to engage, mobilize, and shape public opinion on a global scale. Despite the critical influence of platforms like Twitter in movements such as BLM, a significant gap persisted in understanding the complexities of digital interactions within these movements. Specifically, challenges arose in effectively analyzing the massive volume of data generated, which included both textual content and non-verbal signals like emoji's, alongside discourse from online news channels. This research addressed the pressing need for a comprehensive analysis of sentiment and discourse surrounding the BLMM on social media.

It focused on uncovering how digital communication patterns, including the strategic use of emoji's and narratives from news channels, contributed to the movement's objectives and influenced public perception.

By applying data-driven methodologies, this study sought to reveal the underlying dynamics of social media engagement and media discourse in the context of BLM. The insights gained contributed to academic discourse and offered practical applications for advancing social justice advocacy.

B. Research Objectives

- **Main Objective**

RO1: To analyze the discourse surrounding the BLMM on social media and news channels.

ISSN: 2961-5992

- **Specific Objectives:**

RO2. To identify prevalent themes, emotions, and sentiments expressed within the BLM discourse.

RO3. To explore patterns of similarity and divergence among BLM-related tweets.

RO4. To compare and contrast sentiment between tweets from news channels and the general public.

RO5. To provide insights into the alignment or disparity in perspectives within the BLM discourse across different sources.

C. Research Significance

This research held significant value in advancing our understanding of how social media and online news channels shaped public discourse during pivotal social movements like BLM. By providing a comprehensive analysis of both textual and non-verbal communication patterns, such as the use of emoji's, this study offered new insights into the emotional and thematic dynamics that drove digital engagement. Additionally, by examining the perspectives shared by both the general public and news outlets, this research highlighted potential alignments or disparities in how the BLMM was portrayed and perceived across different platforms. These findings informed academic discussions on digital activism, contributed to the development of more effective social justice campaigns, and offered practical applications for media strategies in shaping public opinion. Ultimately, this research aimed to enrich the discourse on social media's role in activism and contributed to the broader field of social justice advocacy.

II. LITERATURE REVIEW

The BLMM was a focal point for scholarly inquiry, particularly regarding its use of social media platforms like Twitter to amplify marginalized voices and foster collective action. This section explored key themes in the literature, including the movement's origins, the role of sentiment analysis in understanding public discourse, and the application of unsupervised NLP techniques to analyze BLM-related data. Furthermore, it highlighted the novelty of the research, which integrated advanced methodologies to bridge public sentiment with media narratives, offering a comprehensive view of the discourse surrounding the movement.

A. Black Lives Matter Movement (BLMM)

The BLMM attracted significant scholarly attention due to its impact on social discourse and activism.

Research focused on the movement's origins, objectives, and societal implications, with particular emphasis on social media platforms like Twitter. Studies highlighted Twitter's role in amplifying marginalized voices and facilitating collective action. Tufekci [9] discussed Twitter's influence in mobilizing protests and disseminating information during movements such as the Arab Spring and Occupy Wall Street, demonstrating its potential to galvanize social movements. Additionally, Twitter became a valuable data source for analyzing public discourse and sentiment surrounding movements like BLM. Researchers, including Nartey [2], and Edrington et al. [1], utilized NLP techniques to analyze tweets, uncovering patterns of sentiment, emotion, and discourse dynamics, and illustrating the evolving nature of BLM discourse and its societal impact.

B. Sentiment Analysis

Sentiment analysis was vital for understanding the emotional aspects of BLM discourse on Twitter. Researchers employed both supervised and unsupervised techniques to assess public sentiment. Goosen [10] used a supervised approach to categorize tweets about BLM into positive, negative, or neutral, highlighting sentiment shifts in response to major events. Field et al. [6], applied unsupervised methods to reveal the nuances of emotional expressions in BLM discussions. Klein et al. [5] examined how language and framing influenced public sentiment on Twitter. These studies underscored the importance of sentiment analysis in exploring public attitudes toward BLM on social media.

C. Unsupervised NLP techniques

Recent studies using unsupervised NLP techniques explored the dynamics of BLM discourse on Twitter. Zadeh and Cicekli [4] applied topic modeling to identify key themes such as police violence and systemic racism. Badaoui [7] used clustering to categorize tweets by content and sentiment, revealing distinct emotional perspectives. Jiang and Xu [3] employed graph-based methods to analyze the network of BLM discourse, highlighting key influencers.

Bolsover [8] examined temporal shifts in topics and sentiment. While these studies offered insights into sentiment and emotional expressions, they often missed broader linguistic patterns and thematic clusters.

D. Novelty of our approach

This research offered a novel perspective on the BLM discourse by integrating advanced unsupervised NLP techniques, setting it apart from previous studies. While prior research had largely focused on sentiment analysis and linguistic framing, our approach delved deeper by examining nuanced linguistic patterns, thematic clusters, and emotional expressions across a diverse range of sources. By comparing discourse on Twitter and 14 major news channels, including The New York Times and BBC, our study provided a comprehensive analysis that bridged public sentiment with media narratives. Additionally, our research introduced a comparative analysis of emoji usage, shedding

ISSN: 2961-5992

light on the emotional subtleties and attitudes that were often overlooked in traditional sentiment analysis. This multidimensional approach uncovered variations in discourse across different platforms, offering deeper insights into the dynamics of digital communication within the BLM movement.

III. METHODOLOGY

A. Data Collection

Data collection for this study was based on two primary sources: tweets from ordinary individuals and tweets from various news channels. The GetOldTweets3 library in Python facilitated the data collection process.

From Normal People

Predefined hashtags related to BLM and George Floyd's death, such as "blm" and "georgefloyd," were used as search queries. Tweets were collected from May 25, 2020, to August 1, 2020, with a maximum of 1500 tweets per hashtag. The tweets were stored in a data frame and exported to a CSV file.

From News Channels

Tweets from prominent news channels, including The New York Times, BBC Breaking News, The Economist, Reuters, The Wall Street Journal, Financial Times, The Guardian, Daily Wire, CNN, MSNBC, Blaze TV, Free Speech TV, and HGTV were collected using the same hashtags related to BLM and George Floyd's death. The collection period was from May 25, 2020, to August 1, 2020, with a limit of 1500 tweets per hashtag per news source. The tweets were compiled into a data frame and exported to a CSV file.

B. Data Pre-Processing

Before conducting sentiment analysis, several data pre-processing steps were undertaken to enhance the quality and reliability of the dataset:

Removal of URLs

A custom function was employed to identify and remove URLs from the dataset. This step was crucial as URLs could introduce noise and distract from the primary textual content, ensuring that the analysis focused solely on the relevant textual information.

Elimination of Duplicate Entries

To maintain the integrity of the analysis, duplicate entries in the news dataset were removed. This was achieved using a deduplication algorithm that compared entries based on text similarity and metadata, ensuring that each news item was unique and preventing redundancy that could skew results.

Text Normalization

The text data was normalized to standardize the content. This included converting all text to lowercase to avoid case sensitivity issues, removing punctuation and special characters that do not contribute to the analysis, and correcting common misspellings to enhance the accuracy of sentiment analysis.

Tokenization and Stop Words Removal

The text was tokenized into individual words or phrases, and common stop words (e.g., "and," "the," "is") were removed. This process focused the analysis on meaningful terms and phrases, reducing the impact of non-informative words.

Stemming and Lemmatization

To standardize words to their root forms, stemming and lemmatization techniques were applied. This step reduced words to their base or dictionary form (e.g., "running" to "run"), which improved the coherence and accuracy of sentiment analysis.

C. Data Processing

The data processing phase of this study involved several key steps, as illustrated in Fig.1, aimed at extracting meaningful insights from the collected tweets.

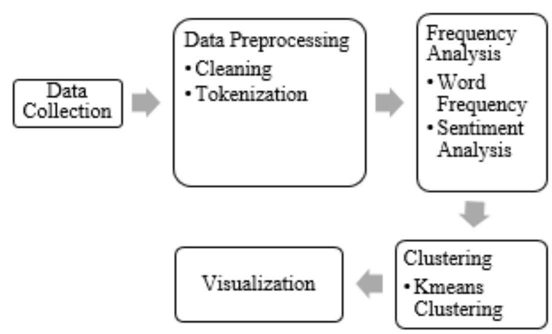

Fig. 1: Simplified Workflow of the research

Initially, the most prominent words were identified, followed by the removal of the least occurring nouns to refine the dataset. Frequency plots of these nouns were created to visualize their distribution within the discourse. To explore textual similarity, the top 25 tweets were identified based on similarity metrics, and a similarity heat map was generated to illustrate relationships between tweets. Elementary sentiment analysis was conducted, with particular focus on extracting emoji to gauge emotional expressions. Cluster identification was performed using K-Means clustering from SciKit-Learn, applied separately to tweets related to the BLMM and to news tweets. Further, cluster identification was carried out for individual news

ISSN: 2961-5992

networks as well as across all networks. The study also incorporated sentiment analysis based on emojis and a separate sentiment analysis specifically targeting tweets from news channels, offering a comprehensive view of the emotional tone and thematic structure present in the data. A more detailed explanation of each processing task is provided below.

D.K-means Clustering

The K-means clustering algorithm was applied to vectored tweet texts to partition them into cohesive clusters based on semantic similarities. By specifying the number of clusters, K-means iteratively assigned each tweet to the nearest cluster centroid, optimizing cluster cohesion and separation. The resulting clusters encapsulated tweets sharing common themes, sentiments, or linguistic patterns, facilitating a nuanced understanding of the discourse surrounding the BLMM.

Upon clustering, centroids of each cluster were examined to identify the most representative terms associated with each cluster. These representative terms offered insights into dominant themes or topics characterizing each cluster, providing researchers with a glimpse into the diverse range of discussions prevalent within the BLM discourse on social media platforms.

E. Sentiment Analysis

In this study, sentiment analysis was conducted on tweets related to the BLMM to discern prevailing attitudes and emotions within the discourse. Utilizing the Text Blob library, sentiment polarity of each tweet was quantified, assigning numerical values to indicate positivity or negativity. Descriptive statistics such as mean, standard deviation, minimum, maximum, and quartiles offered a detailed overview of sentiment distribution, revealing diverse attitudes and emotions within the tweet corpus. A histogram illustrating the frequency of tweets across different sentiment polarity ranges was generated to visually represent this distribution, facilitating a deeper understanding of emotional undercurrents within the BLM conversation on social media platforms.

F. Emoji Analysis

In analyzing the sentiment conveyed through emojis within BLM tweets, a specific subset of data containing emojis was isolated and refined to remove duplicates or anomalies. Emoji's, as non-verbal cues, play a significant role in conveying emotions and sentiments in digital communication. A systematic approach was implemented to extract emojis from tweets using a regular expression pattern, and the extracted emojis were integrated into the tweet dataset for analysis. Sentiment polarity scores, quantified on a scale from -1 (negative) to 1 (positive) using the Text Blob library, were calculated for the emoji's, and their distribution was visualized using a histogram. This visualization enabled the identification of clusters of positive, negative, or neutral sentiment expressions conveyed through emoji's, offering a unique perspective on the emotional dynamics and expressions within social media conversations surrounding the BLMM.

G.Data Validation

In this study, results were validated through a multi-step process that combined both quantitative and qualitative methods:

Data Quality Checks

Initially, the data was carefully preprocessed to remove noise, duplicates, and irrelevant content, ensuring a clean dataset for analysis. Validation started by ensuring that the data preprocessing steps (e.g., cleaning and filtering) were effective. This was done by manually inspecting samples of the cleaned data to verify that unwanted elements, like URLs and duplicates, were correctly removed.

Silhouette Score

The clustering results were quantitatively validated using the silhouette score, which measures the coherence of clusters. A high average silhouette score indicated well-defined and well-separated clusters, reinforcing the reliability of the clustering process.

Visual Validation

Data visualizations, such as similarity heat maps, frequency plots, and clustering diagrams, were used to visually inspect and validate the results. These visualizations helped to identify any anomalies and confirm that the clustering and sentiment trends aligned with expected patterns.

IV. RESULTS AND DISCUSSION

A. Most prominent words within tweets in the considered period

The analysis of prominent words in the tweet corpus, (Fig. 02, bar chart) exposed key themes and sentiments within the dataset. Initially the dataset was refined by filtering out infrequent nouns, focusing on significant linguistic elements. Nouns occurring fewer than ten times were removed, enhancing interpretability.

Key phrases and hashtags like "blacklivesmatter," "icantbreathe," "justiceforgeorgefloyd," and "justiceforbreonnataylor" highlighted ongoing dialogue and

ISSN: 2961-5992

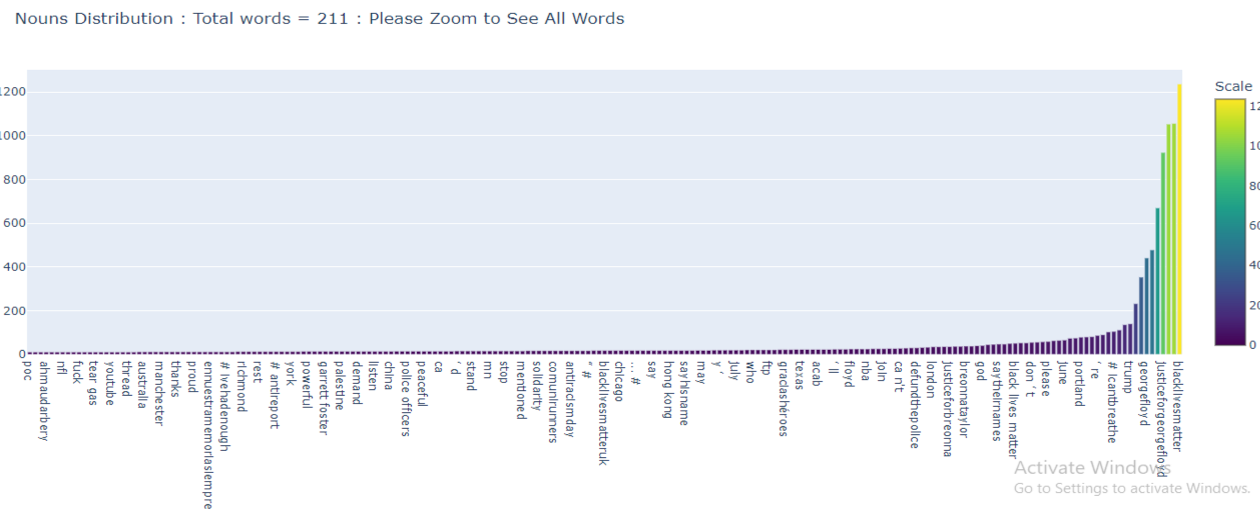

Fig. 2: Prominent Nouns distribution within blm tweets, Author created based on research data

advocacy for racial justice and against police brutality. Names like "George Floyd" and "Breonna Taylor" underscored the focus on individual victims, while references to political figures such as "Trump" and "Antifa" indicated the involvement of diverse ideologies. Themes like "police brutality" and "nojusticenopeace" emphasized societal accountability and demands for systemic change. Locations like "Portland" and "Minneapolis" pointed to specific geographical contexts of protests and significant events. Sentiments expressed through words like "love" and "please" provided insights into empathy, support, and calls for action. Overall, the analysis offered a comprehensive understanding of the themes, sentiments, and socio-political dynamics in contemporary social justice and activism discourse.

B. Cluster Identification

Cluster Identification on BLM Tweets

In examining a corpus of tweets related to the BLMM, the research employed cluster analysis to distill key themes and sentiments, identifying ten distinct clusters, each offering unique perspectives and focal points. The analysis of TABLE I revealed a diverse thematic composition within the BLM discourse on Twitter.

The predominant theme, captured in Cluster 2 (38.1%), focused on advocacy for justice and support, emphasizing justice for George Floyd and Breonna Taylor alongside expressions of love and solidarity for the movement.

Cluster 0 (12.4%) highlighted demands for systemic change with keywords such as "police," "justice," and "demand," while Cluster 3 (10.2%) underscored ongoing activism and resistance through terms like "protest" and "stand." Cluster 7 (9.8%) delved into addressing systemic racism, reflecting deeper analyses of societal structures. Clusters 1 (8.7%) and 4 (7.5%) dealt with expressions of solidarity and calls for action, respectively, highlighting community support and the urgency for change. Emotional responses and personal reflections were prominent in Cluster 5 (5.6%), whereas policy and reform discussions were captured in Cluster 6 (4.9%). Media coverage and public perception were the focus of Cluster 8 (2.3%), and historical context and legacy were explored in Cluster 9 (1.5%). Overall, the clusters provided a comprehensive understanding of the multifaceted nature of the BLM discourse on social media, emphasizing advocacy, systemic change, and the emotional impact on individuals and communities.

Cluster Identification for All news network

In analyzing tweets related to the BLMM extracted from 14 prominent news channels, we employed cluster analysis to uncover key themes and sentiments. This process resulted in the identification of ten distinct clusters, each representing unique perspectives and focal points within the discourse. The analysis of tweets from various news channels revealed diverse thematic threads within the BLM discourse. For example, as shown in TABLE II, Cluster 0 emphasized economic aspects, discussing financial implications and equity.

Cluster 2 focused on law enforcement and justice, reflecting conversations about police brutality and accountability. The distribution of tweets across these clusters offered insights into the prominence of certain themes. Cluster 2 emerged as the most dominant, representing 15.4% of the tweets, likely encompassing a broad range of discussions about police-related issues. In contrast, clusters like Cluster 9 had a smaller proportion, highlighting more specialized topics within the discourse.

ISSN: 2961-5992

TABLE I
CLUSTER IDENTIFICATION ON BLM TWEETS, AUTHOR CREATED BASED ON RESEARCH DATA

Cluster	Cluster Title	Description	Tweet %
2	Law Enforcement and Justice	Reflects conversations about police brutality and accountability with keywords such as "police," "killing," "officers."	15.40%
5	Emotional Responses	Captures emotional responses and personal reflections with terms like "feel," "heart," and "pain."	12.10%
8	Historical Context	Focuses on historical context and legacy with terms like "history," "legacy," and "past."	11.50%
6	Political Discussions	Discusses political dynamics and figures with terms like "Trump," "Biden," and "government."	11.00%
0	Economic Implications and Equity	Discusses financial aspects and equity with terms like "econ," "checks," and "balance."	10.30%
3	Advocacy and Activism	Highlights ongoing activism and resistance with terms like "protest," "stand," and "activism."	9.80%
1	Public Perception and Media Coverage	Focuses on media coverage and public perception with terms like "media," "coverage," and "public."	8.70%
4	Systemic Racism and Reform	Focuses on systemic racism and calls for policy reform with terms like "racism," "systemic," and "reform."	7.50%
7	Community Solidarity	Emphasizes solidarity and community support with terms like "solidarity," "community," and "support."	7.00%
9	Calls for Justice	Centres on demands for justice with terms like "justice," "demand," and "accountability."	6.70%

TABLE II
CLUSTERS IDENTIFICATION ON BLM NEWS TWEETS, AUTHOR CREATED BASED ON RESEARCH DATA

Cluster	Cluster Title	Description	Tweet %
2	Advocacy for Justice and Support	Focuses on justice for George Floyd and Breonna Taylor, along with overarching sentiments of love and support within the movement.	38.10%
0	Demands for Systemic Change	Emphasizes demands for justice and systemic change, with prominent terms like "police," "justice," and "demand."	12.40%
3	Ongoing Activism and Mobilization	Highlights ongoing activism and resistance, including mentions of "protest" and "stand," indicating continued mobilization.	10.20%
7	Addressing Systemic Racism	Focuses on the systemic roots of racial injustice, including terms like "racism," "systemic," and "brutality," reflecting deeper analyses of societal structures.	9.80%
1	Expressions of Solidarity and Community Support	Deals with expressions of solidarity and community support, featuring terms like "solidarity" and "community."	8.70%
4	Calls for Action and Change	Centres on calls for action and change, including terms like "action," "change," and "demand."	7.50%
5	Emotional Responses and Personal Reflections	Highlights emotional responses and personal reflections, including terms like "feel," "heart," and "pain."	5.60%
6	Discussions on Policy and Reform	Focuses on policy and reform discussions, including terms like "policy," "reform," and "change."	4.90%
8	Media Coverage and Public Perception	Features discussions on media coverage and public perception, including terms like "media," "coverage," and "public."	2.30%
9	Historical Context and Legacy	Emphasizes historical context and legacy, including terms like "history," "legacy," and "past."	1.50%

ISSN: 2961-5992

Cluster 2 focused on law enforcement and justice, reflecting conversations about police brutality and accountability. The distribution of tweets across these clusters offered insights into the prominence of certain themes. Cluster 2 emerged as the most dominant, representing 15.4% of the tweets, likely encompassing a broad range of discussions about police-related issues. In contrast, clusters like Cluster 9 had a smaller proportion, highlighting more specialized topics within the discourse.

C. Elementary Sentiment Analysis

This section explored elementary sentiment analysis applied to BLM tweets, revealing a significant proportion of neutral sentiment. The sentiment polarity varied widely, from strongly positive to strongly negative expressions. Fig. 3 demonstrated this distribution, with most tweets clustering around neutrality but significant outliers at the extremes. These insights offered a deeper understanding of the emotional landscape surrounding the BLMM on social media, facilitating a comprehensive analysis of societal attitudes and perceptions toward the issue.

Sentiment Analysis Based on Emojis

The research focused on emoji-based sentiment analysis, using the Text Blob library to quantify emotional tones in tweets (Fig. 4). Descriptive statistics revealed a generally neutral sentiment, with a mean polarity of approximately 0.015 and a wide standard deviation of 0.201, indicating substantial variability in emotional tones conveyed by emojis. Quartile analysis highlighted a significant portion of tweets with neutral emotional tones, but outliers demonstrated emojis' capacity to convey a broad spectrum of sentiments. Overall, this study pioneered understanding the complex interplay between linguistic expressions and emotional nuances in social media discourse, emphasizing the profound impact of emojis on shaping online interactions and perceptions.

Sentiment Analysis on News Channel Tweets

Analyzing sentiment in tweets from various news channels about the BLMM revealed diverse perspectives and stances within media discourse. The sentiment analysis, conducted using the Text Blob library, depictions nuanced portrayals of the BLMM by different news outlets. while The Economist maintained a marginally positive stance. For instance, according to the Figure 05, BBC News and The Guardian exhibited slightly negative sentiments, Reuters presented a neutral perspective, indicating balanced reporting. Contrasting sentiments were evident in sources like Real Daily Wire and Free Speech TV, with MSNBC displaying highly negative sentiment and CNN showcasing highly positive sentiment. These insights emphasized the media's influential role in shaping public perception of the BLMM, highlighting underlying biases and editorial perspectives.

D. Ensuring Data Integrity and Accuracy

The effectiveness of the data preprocessing steps and clustering process was thoroughly validated to ensure the integrity and accuracy of the results. The preprocessing steps, including the removal of noise, duplicates, and irrelevant content, were critical in creating a clean and reliable dataset for analysis. This foundation was further reinforced by the high silhouette score of 0.781 obtained during the clustering phase, indicating well-defined and well-separated clusters. Such a score underscores the effectiveness of the clustering model and suggests that the data points were appropriately grouped with minimal overlap between clusters.

Visual validation through similarity heat maps, frequency plots, and clustering diagrams further confirmed the robustness of the data processing and clustering techniques. These visualizations revealed clear patterns in sentiment trends and clusters, which aligned with expected outcomes and highlighted key insights into the digital dynamics of the studied social movements.

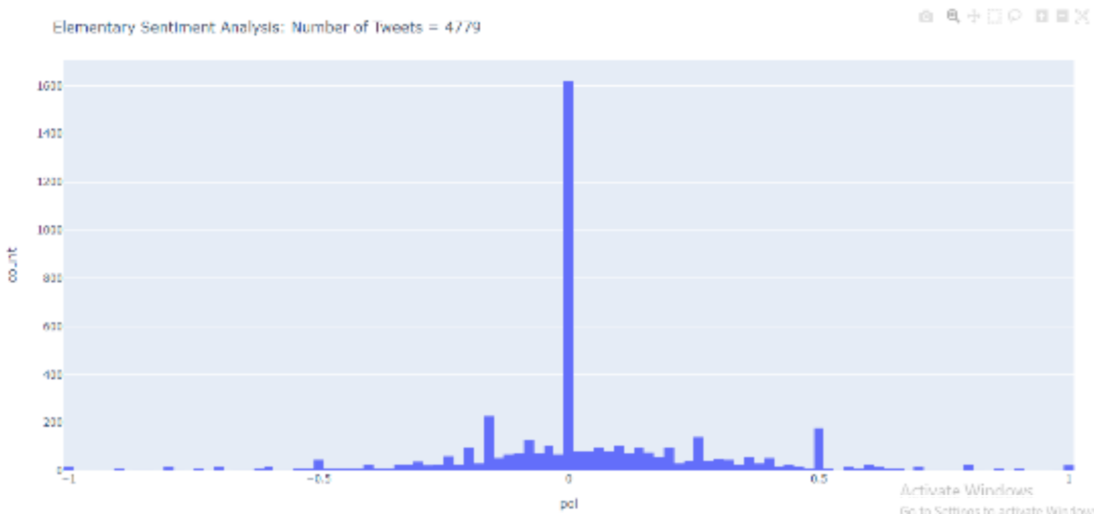

Fig. 3: Elementary sentiment analysis within BLM tweets, Author created based on research data

ISSN: 2961-5992

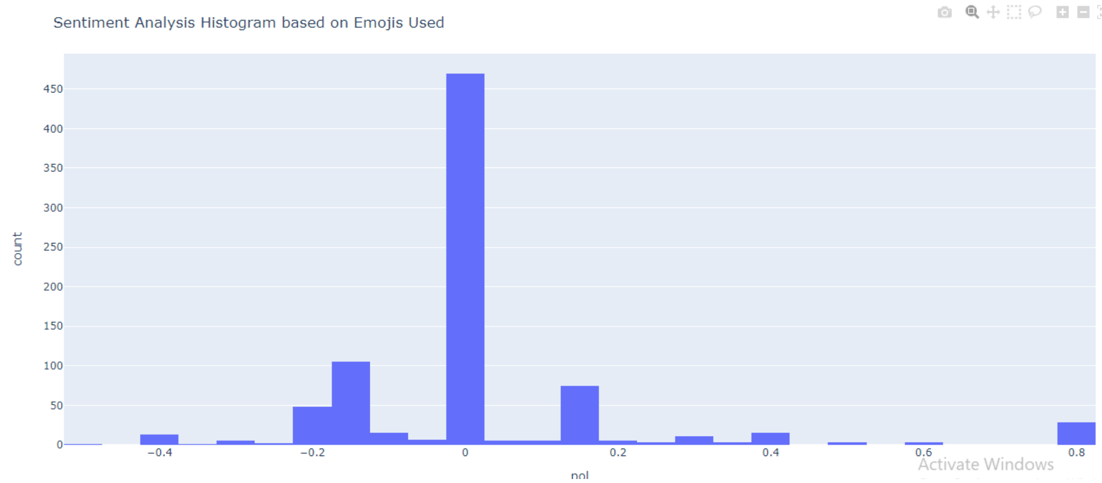

Fig. 4: Sentiment analysis Histogram based on Emoji's, Author created based on research data

E. Research Objectives Alignment

RO1 Alignment:

The discourse surrounding the BLMM on social media and news channels revealed a rich tapestry of linguistic elements that emphasized ongoing activism and demands for societal change. Hashtags such as "blacklivesmatter," "justiceforgeorgefloyd," and "justiceforbreonnataylor" emerged as prominent fixtures within the dataset, reflecting persistent dialogue and advocacy surrounding specific cases and broader movements addressing racial injustice and police brutality. References to political figures like "Trump" and "Antifa" further accentuated the intersectionality of socio-political dynamics within the conversation, indicating diverse ideologies and stakeholders involved.

RO2 Alignment:

Prevalent themes, emotions, and sentiments expressed within the BLM discourse encompassed solidarity, activism, frustration, and resilience. Tweets exemplified expressions of support for the movement ("Jews United for Justice affirmed unequivocally that Black lives matter") alongside frustration with misconceptions about its core message ("Anyone else tired of the 'tHE pROTeST iSN't aBOuT bLacK liVES aNYMorE!' takes?"). Moments of empathy and reflection, symbolizing collective resolve to confront systemic racism and police brutality, were also evident ("We knelt in silence for the 8 minutes that George Floyd was suffocated and held against his will").

RO3 Alignment:

Patterns of similarity and divergence among BLM-related tweets were discerned through robust methodologies like TF-IDF and cosine similarity.

Highly akin tweets indicated shared sentiments, narratives, or discussions within the BLM discourse, while a heat map analysis visually represented the intricate relationships within the discourse, highlighting clusters of cohesive discussions around central topics such as specific hashtags or keywords like "#justiceforgeorgefloyd" or "#icantbreathe."

RO4 Alignment

Sentiment comparison between tweets from news channels and the general public revealed nuanced portrayals of the BLMM. Tweets from various news outlets exhibited diverse perspectives and stances, with some sources displaying slightly negative sentiments (BBC News, The Guardian), while others maintained marginally positive stances (The Economist). Contrasting sentiments were evident in different news sources, underscoring the media's influential role in shaping public perception of the BLMM and highlighting underlying biases and editorial perspectives.

RO5 Alignment

Insights into the alignment or disparity in perspectives within the BLM discourse across different sources disclosed varying thematic threads and emphasis. While some clusters in news channel tweets focused on economic aspects or law enforcement and justice, others delved into ongoing activism, resistance, or the systemic roots of racial injustice. The relative emphasis placed on specific themes within the conversation varied among different news networks, shedding light on the diverse coverage of BLM-related topics and the broader socio-political implications of media representation.

ISSN: 2961-5992

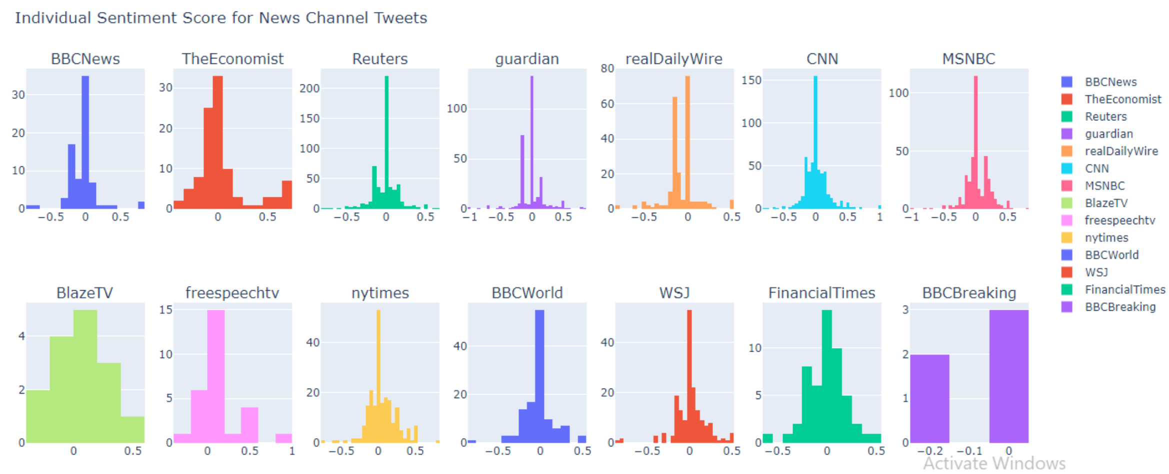

Fig. 5: Individual sentiment analysis for each and every news channel based on tweets

V. Conclusion

This research provided a comprehensive and multifaceted analysis of the BLM discourse following George Floyd's death, leveraging a range of NLP techniques to uncover diverse perspectives, thematic threads, and sentiments within tweets from both individuals and news channels.

By employing content analysis, critical themes such as social justice, police brutality, and systemic racism were identified, highlighting the movement's complex and varied discourse. The study's textual similarity analysis revealed shared linguistic patterns, while sentiment analysis, including emoji-based assessments, offered nuanced insights into the emotional landscape surrounding the BLM movement.

The findings underscored the significant role that social media played in shaping public discourse and the diverse range of emotions and opinions that characterized the BLM conversation. The sentiment analysis of news channel tweets further illustrated the media's influential role in portraying the movement, with varying degrees of positivity and negativity across different outlets. This comprehensive approach provided a deeper understanding of how the BLMM was perceived and discussed across different platforms, highlighting both the alignment and disparity in perspectives.

Future research should extend this analysis longitudinally to track the evolution of BLM discourse over time and across different social media platforms for a broader understanding. Additionally, examining the intersections of race, gender, and class within the discourse, as well as conducting cross-cultural analyses, will provide further insights into the global impact of the BLMM. Network analysis and media framing studies will also be valuable in exploring the interconnectedness of users, topics, and sentiments, and in assessing the broader socio-political implications of media coverage on the movement.

VI. References

[1] C. L. Edrington and N. Lee, "Tweeting a social movement: Black Lives Matter and its use of Twitter to share information, build community, and promote action," The Journal of Public Interest Communications, vol. 2, no. 2, pp. 289-289, 2018.

[2] M. Nartey, "Centering marginalized voices: A discourse analytic study of the Black Lives Matter movement on Twitter," in Voice, Agency and Resistance, Routledge, 2023, pp. 65-80.

[3] K. Jiang and Q. Xu, "Analyzing the dynamics of social media texts using coherency network analysis: a case study of the tweets with the co-hashtags of #BlackLivesMatter and #StopAsianHate," Frontiers in Research Metrics and Analytics, vol. 8, p. 1239726, 2023.

[4] M. Zadeh and I. Cicekli, "Protest Event Analysis: A New Method Based on Twitter's User Behaviors," Information Technology and Control, vol. 52, no. 2, 2023.

[5] C. Klein, R. Reimann, I. O. Quintana, M. Cheong, M. Ferreira, and M. Alfano, "Attention and counter-framing in the Black Lives Matter movement on Twitter," Humanities and Social Sciences Communications, vol. 9, no. 1, pp. 1-12, 2022.

[6] A. Field, C. Y. Park, A. Theophilo, J. Watson-Daniels, and Y. Tsvetkov, "An analysis of emotions and the prominence of positivity in #BlackLivesMatter tweets," Proceedings of the National Academy of Sciences, vol. 119, no. 35, p. e2205767119, 2022.

[7] S. Badaoui, "Black Lives Matter: A new perspective from Twitter data mining," Policy Center for the New South, pp. 2021-01, 2020.

[8] G. Bolsover, "Black Lives Matter discourse on US social media during COVID: polarised positions enacted in a new event," arXiv preprint arXiv:2009.03619, 2020.

[9] Z. Tufekci, Twitter and Tear Gas: The Power and Fragility of Networked Protest, Yale University Press, 2017.

[10] D. Goosen, "An Analysis of Black Lives Matter Social Media Activism from 2014-2020," unpublished, 2021.

ISSN: 2961-5992