



Indexed in:



**Lembaga Masyarakat
Ahli, Pendidik, dan
Peneliti Indonesia**

Jalan Santun Gg Sumbar No 52,
Kelurahan Sudirejo I,
Kecamatan Medan Kota,
Sumatera Utara,
Indonesia

Simbisa Journal of Social Sciences and Humanities

A Case Study of Social Media Reach and Sentiment Analysis on *Karo* Ethnic Identity: *#karobukanbatak*

Loli Gunali Ginting*

Sekolah Tinggi Ilmu Manajemen Sukma, Medan, Indonesia

loli.gunali@gmail.com

ABSTRACT

This study investigates the sentiment expressed in the #karobukanbatak discussions on social media to understand public attitudes towards the cultural aspects of the Karo people in North Sumatra. By employing Deep Learning and Pretrained Language Models (PLM), the research analyzes how people perceive and respond to various forms of communication and cultural expressions by the Karo ethnicity. Data was collected from multiple social media platforms and analyzed through natural language processing and statistical analysis. The findings reveal a significant rise in the volume and reach of #karobukanbatak on social and non-social media platforms, with positive sentiments dominating up to 90.008% of all mentions. Platforms such as TikTok, Karogaul.com, YouTube, and sorasirulo.com were identified as key venues for these discussions. Additionally, certain public accounts and profiles were more active in discussing or sharing information related to the topic, highlighting the importance of the 'Karobukanbatak' concept in establishing the distinctive ethnic character of North Sumatra. The implications of these findings are significant, as they demonstrate how social media plays a crucial role in shaping and comprehending public perceptions and sentiments regarding cultural identity.

Keywords: #karobukanbatak, Sentiment Analysis, Social Media Influence, Cultural Identity, Pretrained Language Models (PLM)

1. INTRODUCTION

1.1 Background of the study

People now primarily communicate their thoughts, ideas, and expressions on social media. The *Karo* community's cultural identity and values are reflected in *#karobukanbatak* within this setting. Social media views and opinions can shape public perceptions of a particular subject or organization. A sentimental analysis of *#karobukanbatak* can show how society views and reacts to various aspects of *Karo* culture.

The knowledge and preservation of regional culture are becoming more and more crucial in the context of globalization. Examining the sentiment behind *#karobukanbatak* can reveal how the *Karo* community perceives and handles its cultural legacy in the digital age. The emergence of social media and technology has made cross-cultural communication easier. However, there are also issues with maintaining and interpreting cultural identity in the digital age. Sentiment analysis can help understand how social media affects people's experiences and views of *Karo* culture.

Social media usage has increased, which raises the possibility of interpersonal conflict or group cooperation. By using sentiment analysis to track *#karobukanbatak*, we can spot trends that could spark abuse or promote harmony amongst *Karo* communities.

Sentiment analysis of *#karobukanbatak* provides valuable insights into the social and cultural background, enabling a fuller understanding of social media's role in molding and reflecting cultural identity. Social media platforms like Twitter have

become crucial in expressing and shaping ethnic identities. They provide a space for individuals to share their cultural narratives and engage in discussions about their heritage (Estrada et al., 2022; Lachmar, 2017)

Hashtags such as *#karobukanbatak* assert and differentiate *Karo*'s identity from the broader Batak ethnic group. This digital activism helps mobilize community support and raise awareness about cultural distinctions (Estrada et al., 2022; Stella et al., 2020). Sentiment analysis is a powerful tool for understanding public opinion and emotional responses on social media. It involves categorizing text into positive, negative, or neutral sentiments, providing insights into users' collective moods and attitudes (Lachmar, 2017; Stella et al., 2020).

The study will use sentiment analysis to evaluate *#karobukanbatak*. This will involve using a customized dictionary and model tailored to the specific linguistic and cultural context of the *Karo* people (Lachmar, 2017). Hierarchical cluster analysis will identify conversation clusters and explore the semantic relationships between words. This method will help understand the main themes and topics discussed in *Karo* identity (Lachmar, 2017).

The study will also consider the impact of social media discussions on the mental well-being of the *Karo* community. Previous research indicates that social media can provide emotional support and a sense of connection, which is particularly important for minority groups (Estrada et al., 2022; Lachmar, 2017). Potential challenges include the brief nature of tweets, which may limit the depth of analysis, and the difficulty in capturing all

relevant tweets due to variations in hashtag usage. Additionally, the lack of user demographic information may limit the ability to draw comprehensive conclusions (Stella et al., 2020).

This study will contribute to understanding how social media influences ethnic identity and community dynamics. It will provide valuable insights into the role of digital platforms in cultural preservation and identity assertion (Estrada et al., 2022; Stella et al., 2020).

1.2 Formulation of the Problem

In this research, the author will identify and analyze the sentiments contained in *#karobukanbatak* on social media. Problem formulation will guide the research to answer the following questions:

1. What is the overall distribution of *#karobukanbatak* sentiments on social media, and how do these sentiments evolve?
2. How does social media influence public perception and communication about cultural aspects of *Karo* society, as reflected in *#karobukanbatak*?

1.3 Purpose of the Research

This research aims to gain an in-depth understanding of the sentiment associated with *#karobukanbatak* on social media. The objectives that will be achieved through this research are as follows:

1. This research aims to identify and analyze the sentiment of *#karobukanbatak* on social media. This includes determining whether the sentiment is positive, negative, or neutral and understanding how this sentiment is distributed and changes over time.

2. This research explores how social media platforms contribute to understanding and communicating cultural expressions relating to the *Karo* community in North Sumatra. The research highlights the role of social media in affirming ethnic identity and disseminating information on cultural topics.

2. LITERATURE REVIEW

2.1 Sentiment Analysis in Social Media

Subjective information such as opinions, attitudes, and emotions are expressed in the text through various natural language processing (NLP) techniques that classify text into positive, negative, or neutral sentiment categories. Sentiment analysis, one of the subfields of N.L.P., is crucial in extracting and assessing subjective information from unstructured data sources such as social media, reviews, and blogs. This process involves several methodologies and techniques, each contributing to accurate sentiment classification in text.

Techniques in Sentiment Analysis

Sentiment analysis often uses data-driven methods that do not rely on feature engineering. These methods include document-level sentiment analysis, aspect-based sentiment analysis, and word polarity recognition. These approaches help capture the complex semantic representation of text, which is crucial for understanding sentiment expressed in unstructured data (Ranjan et al., 2023). Sentiment analysis uses machine learning (ML) and deep learning (DL) models. ML models such as Logistic Regression, Random Forests, and DL based on Neural Networks are commonly used. These flexible models can achieve competitive levels of accuracy in classifying sentiment (Tsiligaridis, 2024).

Fast learning strategies have been introduced to overcome the challenge of data scarcity. These strategies utilize pre-trained models to inherit prior knowledge and learn high-level semantic features, thereby improving the recognition of subjective expressions in the text (Cheng et al., 2023).

Challenges and Applications

Sentiment analysis faces challenges such as dealing with informal language and detecting sarcasm. Advanced techniques address these issues, including rule-based and machine-learning approaches (Thakur et al., 2024). Sentiment analysis is widely used in business intelligence to understand customer sentiment, improve products and services, and develop effective marketing strategies. This analysis helps companies assess public attitudes toward products, services, and social issues (Chakraborty, 2023).

While sentiment analysis provides valuable insights into subjective information, it is important to consider this technique's limitations and ethical implications. The accuracy of sentiment classification can be affected by data quality and the complexity of human emotions. In addition, ethical considerations such as privacy and data security must be considered when analyzing personal information from digital text.

Sentiment analysis is a powerful tool for interpreting the emotional tone of large amounts of text data, especially on social media, where user-generated content abounds. This capability is critical for understanding public sentiment, monitoring brand reputation, and making informed decisions based on consumer feedback. Integrating advanced machine learning techniques and metadata increases the

precision and applicability of sentiment analysis, making it an indispensable tool in the digital age.

Neural Networking and Sentiment Analysis

Neural networks, including convolutional neural networks (CNN) and recurrent neural networks (RNN), are essential in sentiment analysis. They offer a powerful framework for detecting and categorizing emotional tendencies in text. These models address challenges such as data noise and model interpretability, improving the precision and comprehensiveness of sentiment analysis (Wang, 2024). Transfer learning and domain adaptation techniques further expand the applicability of sentiment analysis across multiple contexts, improving model performance and adaptability (Wang, 2024).

Text and Metadata Integration

<U.N.K.> Integrating metadata, such as reposts and followers, with textual data significantly enhances sentiment analysis results. A hybrid model combining B.E.R.T. output with metadata showed superior performance, achieving an F1 score of 69.4%, highlighting the effectiveness of this approach in a social media context (Guerrero-Contreras et al., 2024). This integration allows for a more nuanced understanding of sentiment, especially distinguishing between neutral and positive sentiment, which has increased in importance (Guerrero-Contreras et al., 2024).

Machine Learning Techniques

Machine learning models, such as gradient boosting classifiers, effectively automate sentiment analysis, achieving high levels of accuracy. These models are trained on labeled data to recognize patterns and derive meaningful insights from social

media content like Twitter posts (Agarwal et al., 2024). Ensemble methods such as Voting Classifiers and Decision Trees improve the efficiency of sentiment analysis, especially in dynamic social media environments such as Facebook (Srinidhi et al., 2024).

Meta-Learning and Emotion Detection

Meta-learning offers promising advances in sentiment analysis by improving algorithm generalization and data efficiency. This approach is particularly beneficial in scenarios where traditional sentiment analysis might fail, such as emotion detection (Kusal et al., 2024). The ability to detect emotions along with sentiment provides a deeper understanding of the content created by users, which is critical for applications in human psychology and consumer behavior analysis (Kusal et al., 2024).

While sentiment analysis provides valuable insights, it is crucial to consider its limitations, such as the need for large amounts of labeled data and the challenges of accurately interpreting nuanced emotions. Future advances in behavioral learning and deep learning techniques may overcome these challenges, increasing the effectiveness and scope of sentiment analysis in social media.

2.2 Cultural Communication and Social Media

Social media platforms have revolutionized how artistic expression and identity are exchanged and shared, offering dynamic communication channels that shape cultural narratives. These platforms facilitate the sharing of diverse content, enable user engagement, and support the creation of collective memory and identity. The following section explores the mechanisms by which social media platforms achieve these functions.

Content Diversity and User Engagement

Social media platforms such as Weibo have shown that content diversity and positive sentiment significantly increase the spread of cultural narratives. For example, the Palace Museum's Weibo account utilized content diversity and user interactions, such as reposts and comments, to effectively communicate cultural stories (Sun et al., 2024). The platform provides a graphical user interface (G.U.I.) that allows users to interact with digital content, encouraging meaningful social interactions. Relationship profiles and historical interaction data facilitate these interactions, improving the quality of online exchanges. (Lyons et al., 2019).

Narrative Sharing and Cultural Storytelling

Social media enables the sharing of narratives across platforms such as Wikipedia, Facebook, Twitter, and YouTube. These shared stories can unite or polarize communities, illustrating the complex dynamics of online communication. The 'Mediated Narrative Analysis' framework helps interpret these narratives, highlighting the micro- and macro-level politics involved in digital storytelling (Page, 2018). In the context of art and design, social media short videos utilize symbolism to convey cultural stories. For example, Chinese cultural symbols are used in videos to showcase traditional cuisine and handicrafts, effectively communicating China's diverse cultural disposition and image.

Collective Memory and Identity Formation

Social networking platforms are crucial in creating and disseminating collective memory and identity. Despite challenges, such as their dynamic and commercial nature, these platforms allow for the collection and semantic analysis of

discussions on heritage and identity, which contribute to shaping cultural narratives (Dallas et al., 2024). While social media platforms offer significant opportunities for the exchange and dissemination of artistic expression, they also present challenges, such as the potential for polarization and commercial influence on content. Researchers and cultural institutions must navigate these complexities to optimize the use of social media for cultural communication and identity formation.

2.3 Deep Learning and Trained Language Models

Integrating deep learning and trained language models (P.L.M.) in sentiment analysis has significantly improved the accuracy and efficiency of text classification tasks. These models, particularly those based on B.E.R.T. (Bidirectional Encoder Representations from Transformers), have revolutionized the field by automatically learning and predicting sentiment from text data. These advancements are based on theoretical studies exploring various machine-learning techniques for text classification, which offer substantial improvements over traditional methods.

Advances in Deep Learning for Sentiment Analysis

Deep learning has significantly improved sentiment analysis by allowing models to learn complex patterns in text data. Techniques such as CNN (Convolutional Neural Networks) and LSTM (Long Short-Term Memory Networks) have been effectively combined with B.E.R.T. to improve sentiment classification tasks. For example, the BERT-LSTM model has shown superior performance, achieving 95.91% accuracy in sentiment analysis tasks (V & Bargavi, 2024).

BERT and its variants, such as DistilBERT, have sophisticated sentiment analysis. These models utilize large-scale pre-training on a diverse text corpus to capture nuanced language features and improve sentiment prediction accuracy. Experiments with the SST2 dataset have demonstrated the effectiveness of DistilBERT in sentiment classification, underlining its transformative potential in NLP research. (Wu et al., 2024).

Applications and Challenges

Deep learning models are used in sentiment analysis in various domains, including social media, product reviews, and healthcare. In the medical field, BERT-based models have been used to analyze sentiment in clinical text, although there are still challenges, such as limited datasets and model design complexity. CNNs have been shown to perform well on smaller medical text datasets when combined with BERT, highlighting the importance of model selection in domain-specific applications. (Chen, 2024). Aspect-based sentiment analysis (ABSA) focuses on clustering sentiment based on certain document components. The use of BERT combined with CNN and LSTM has shown promise in handling domain-independent datasets, overcoming some of the challenges associated with traditional ABSA models (V & Bargavi, 2024)

Although deep learning and P.L.M. have significantly advanced sentiment analysis, many challenges remain, especially in domain-specific applications where data scarcity and model complexity can impact performance. Future research directions include developing more efficient model architectures and exploring new applications in various domains (Chen, 2024; Nikhil Sanjay Suryawanshi, 2024)

2.4 “The Context of *KarobukanBatak*”

The classification of ethnic groups in North Sumatra is unclear and confusing, particularly regarding the labels “Batak” and “Malay.” The origins of these ideas can be found in the writings of foreign ethnographers who traveled to northern Sumatra in the fifteenth century. First, the term 'Batak' was used to describe the communities residing in the interior of Aceh and North Sumatra in the writings of Pires (Winstedt, 1947) and Pinto (Damanik, 2018). In the meantime, Polo (Damanik, 2018) referred to the settlers living along the coast as “Malays.” Reid’s compilation of stories from foreign ethnographic trips to North Sumatra (1995) emphasizes this even more.

The homogeneity of the inside communities into a single category known as "Batak" and the designation of coastal communities as "Malay" are the causes of this ambiguity. It is critical to realize that both terms refer to basic traits of the communities in the hinterland and along the shore, not specific ethnic groupings. Important concerns come up about who exactly belongs to the "Batak and Malay" communities, how these categories are established, and who the basis for these categories' membership is. Are the terms “Batak” and “Malay” just ambiguous designations applied to communities that fall under both categories?

'*Karobukanbatak*' is a clear representation of the *Karo* community's wish to claim their own ethnic identity, different from the more general term 'Batak' that was previously applied by outsiders like ethnographers and colonial authorities (Loeb & Heine-Geldern, 1989). The name 'Batak' Needs to capture the significance of the *Karo* people's unique cultural and linguistic traits, which are emphasized in other contexts. (Kipp & Kipp, 1983; Loeb & Heine-Geldern, 1989; Putro, 1981)

Critics point out that the term 'Batak' has been applied to several different ethnic groups in North Sumatra, such as *Karo*, *Simalungun*, *Toba*, *Pakpak*, *Mandailing*, and *Angkola*, oversimplifying the region's complex ethnic landscape (Kipp & Kipp, 1983; Loeb & Heine-Geldern, 1989; Putro, 1981). Moreover, the name 'Batak' was initially generated by something other than the groups it was meant to characterize, suggesting that it is an external classification instead of a self-identified label (Loeb & Heine-Geldern, 1989). Social difficulties resulting from some groups—like the *Mandailing*—rejecting the title "Batak," while other groups—like the *Angkola*—accept it—illustrate further opposition to ethnic homogenization (Damanik, 2018).

The main points of *#karobukanbatak* are criticisms of the homogenization of ethnic groups in North Sumatra, the significance of appreciating the individuality of every group, and the fact that ethnic identity is built and dynamic (Damanik, 2018; Kipp & Kipp, 1983; Loeb & Heine-Geldern, 1989; Putro, 1981)

The term *#karobukanbatak* describes the social, cultural, and potentially political facets of the *Karo* community. The word refers to the *Karo* people's expression and cultural identity, seen in various web articles, social media posts, and promotional materials for their culture.

The following elements can be incorporated into *#karobukanbatak*:

- a. This refers to the traditional aspects of *Karo* culture, including music, dance, art, and customs.
- b. *#karobukanbatak* can include the *Karo* language and the unique communication methods the *Karo* community uses.

- c. The *Karo* community upholds certain values in its beliefs and customs.
- d. *Karo* culture has traditional food varieties and unique eating customs.

How *#karobukanbatak* may affect sentiments and opinions:

- a. The *Karo* community's pride and cultural identity can be strengthened if "*KarobukanBatak*" is positively portrayed.
- b. Some elements of *#karobukanbatak* may cause disagreements or stir up controversy, which could stoke polarized feelings and opinions within the community.
- c. Successful *#karobukanbatak* promotion can foster favorable perceptions and enhance Suku *Karo*'s reputation in the eyes of the general public.
- d. Current cultural events or trends receiving attention on social media or other platforms may impact public perceptions and sentiments regarding *#karobukanbatak*.
- e. Conversations and exchanges over *#karobukanbatak* on social media platforms can significantly influence perceptions and feelings. Criticism, as well as encouragement, may move fast and shape public opinion. Sentiment analysis can shed light on how societal reactions to these cultural characteristics in the digital age are shaped by an awareness of the background of *#karobukanbatak* and how these cultural factors can influence opinions and sentiments.

METHOD

3.1: Sources of Data

The information needed for this study will come from various social media sites relevant to conversations about *#karobukanbatak*. It's possible to sample various platforms like Facebook, YouTube, TikTok, blogs, webpages, and culture-specific forums. These platforms were chosen based on how well-liked and varied the user interactions were about the cultural subject.

The selection of data will take place between 2018 and 2023. Additionally, an attempt will be made to get a representative sample from various sources, and data collected from platforms with particular biases will be avoided.

3.2 Methods of Analysis

Sentiment analysis based on Deep Learning and pre-trained Language Models (PLMs) is the analytical method to be applied. Using this technique, sentiment in texts can be automatically classified into three categories: positive, negative, and neutral.

3.3 The Methodology of Research

3.3.1 Data Collection

- a. Finding *#karobukanbatak*-related keywords to compile pertinent information.
- b. Accessing and retrieving data from social media sites using sentiment analysis platforms.

3.3.2 Sentiment Analysis

- a. Pre-trained language models (PLM) and deep learning models categorize sentiment for every token or sentence.

- b. Figuring out what proportion of opinions are neutral, negative, and positive.

3.3.3 Analysis of Statistical Data

Examining the distribution of sentiment generally and tracking variations over time.

3.3.4 Results Interpretation

Interpreting research in the *#karobukanbatak* cultural context and looking for patterns or trends to shed light on new information.

3.3.5 Final Thoughts

Presenting the research results in a paper format and summarizing the key findings.

ANALYSIS AND FINDINGS



Figure 1. Summary

The *#karobukanbatak* volume report shows a notable rise in volume over the given time frame. According to the data provided, the fact that there has been a 100% increase in volume indicates that this topic has grown in popularity and frequency of discussion across various venues. During the current period, compared to the prior one, there has been a noticeable rise in the volume mentioned in the graph. This suggests a rise in public

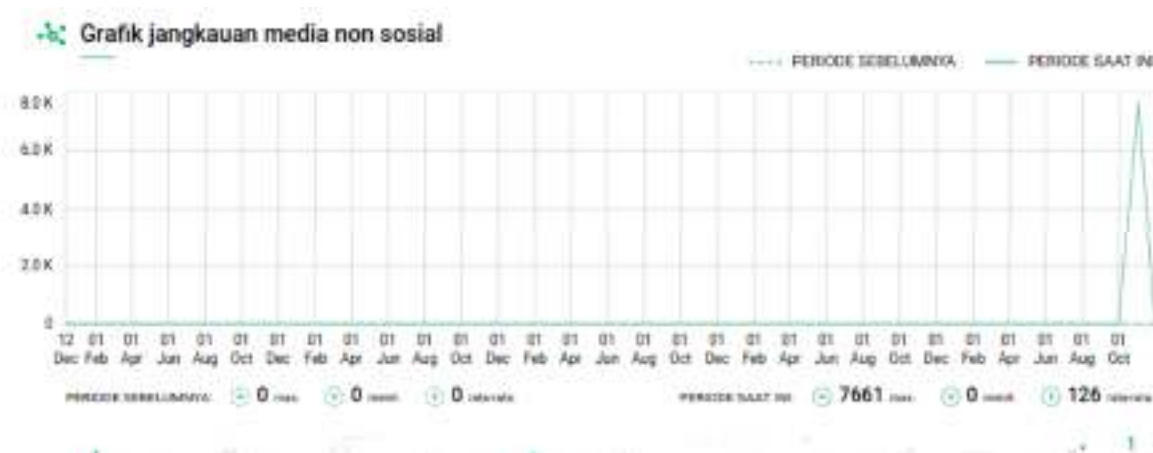
interest and conversation about *#karobukanbatak*. In addition, each category has had a notable surge, particularly on TikTok, with a 100% rise compared to the prior time. This implies that Twitter has emerged as one of the primary conversation forums for this subject. In conclusion, the summary of *#karobukanbatak* suggests that public interest in this topic has grown, especially on social media and other online platforms.



Graph 1. Social media reach graph

An examination of the social media reach graph reveals a notable rise in the topic *#karobukanbatak*'s reach throughout the given time frame. According to the available data, there has been a 100% rise in social media reach, with a recorded gain of 1.3 million over the preceding quarter.

This implies that during that time, there has been a greater focus on and debate on this topic on social media. With 1.3 million social media impressions, the reach increased significantly from the prior time frame.



Graph 2. Non-social media reach.

Analysis of non-social media reach charts indicates a significant increase in the reach of the topic *#karobukanbatak* on non-social media platforms. According to available data, there was an increase in non-social media reach of 7,661, marking a

100% increase compared to the previous period. This suggests that the topic is not only popular on social media but is also gaining wider attention on other platforms such as websites, forums, and possibly in print or broadcast media.



Figure 2. Hashtag by Category

Based on available data, analysis of mentions by category for *#karobukanbatak* discussions reveals significant changes in mention volume across several media categories.

1. TikTok experienced a 100% increase in mentions from the previous to the current period, indicating that this platform has become a primary hub for discussions related to this topic.
2. The podcast category also recorded a 100% increase, although the absolute number may be smaller than Twitter's. This suggests that the *#karobukanbatak* topic is gaining attention in various media types.
3. There was no change in mention volume in the TikTok and Instagram categories, which remained at 0% change. This could mean that discussions of *#karobukanbatak* are less popular or less frequently discussed on these platforms during the analyzed period.
4. Overall, the analysis of mentions by category indicates that *#karobukanbatak* discussions are most active on TikTok and emerging in podcasts. At the same time, other platforms like Twitter and Instagram have not shown a significant increase in activity related to this topic.



Figure 3. Numerical Summary

Based on the available numerical summary, sentiment analysis reveals that discussions related to *#karobukanbatak* have many positive mentions. With 42% of total mentions categorized as positive, the overall sentiment in these discussions tends to be supportive or has a favourable view of the topic. Moreover, there was a 100% increase in positive mentions from the previous to the current period, suggesting that positive sentiment towards this topic is growing. This could indicate that more people are engaging in discussions with

supportive views or that more content portrays the topic positively.

However, a 100% increase in negative mentions was also observed, indicating that while there is much positive sentiment, these discussions also attract attention from those who may disagree with or have a critical view of the *#karobukanbatak* topic. Overall, sentiment analysis on the numerical summary shows that *#karobukanbatak* discussions balance positive and negative views, with a stronger inclination towards positive sentiment.

Konteks diskusi

tdk toba suka lagu #SUKUKARO1838 dna #medan sejarah kalak uga #CapCut kaban #peranganingin balas beru org #pyf #mejuah #kalakkaro suku bahasa giting #KBBYES #sukukaro #ZONATIKTOKKALAKARO #LAGUKARO #medanviral charo #tkimalaysia #mandailingbukanbatak #medantiktok #adatkaru pembatakan nya batak sampainsekarang #karo #anakkangkak bkn hiburan adat #KALAKARO #mejuah_juah_kita_kerina_bujur

Figure 4. Discussion Context

Analysis of the discussion context summary related to *#karobukanbatak* reveals that these discussions have many mentions across various platforms. With 167 mentions, these discussions have garnered a relatively wide range of

attention. Positive mentions dominate with 42% of the total mentions, indicating that the general sentiment in these discussions tends to be supportive or has a favorable view of the topic.

The interactions recorded in this discussion context are also quite high, with 66,117 interactions indicating a high level of engagement from users in these discussions. This could include various forms of interaction, such as likes, comments, and shares. Additionally, trending hashtags like

#ustadztengkuzulkarnain, #kalakKaro, #karobukanbatak, and #mejuahjuah show the most talked-about topics and aspects within the #karobukanbatak context. The use of these hashtags helps disseminate information and facilitate discussions among users interested in the topic.

Table 1. Most Popular Hashtag

Date and Time	Account/Profile	Type of Social Media	Hashtag
2021-09-04 19:03	<i>jasantaperanginan</i>	tiktok.com	#ustadztengkuzulkarnain #kalakKaro #karobukanbatak #ginting #tarigan #sembiring #mejuahjuah #pera..
2023-09-01 05:55	<i>kudinmugusgus</i>	tiktok.com	Apa sih itu batak? part 1 #karobukanbatak #batak #toba #Karo #simalungun #mandailing #pakpak #an...
2022-06-29 21:58	<i>jasantaperanginan</i>	tiktok.com	#stitch dengan @arif.arval entahlah... 🤔🤔 #peranginangi n #kalakKaro #karobukanbatak #mejuahjuah
2023-02-18 06:38	<i>Karobukanbatak</i>	tiktok.com	#karobukanbatak
2023-09-12 04:27	<i>gintingbrebaktina</i>	tiktok.com	Bujur bg judika#SUKUKARO #karobukanbatak #KABUPATENKARO #KBBYES #ZONATIKTOKKALAKARO
2023-10-06 23:42	<i>amsal_sitepu</i>	tiktok.com	kemana arah diskusi KBB mu? #Karo #karobukanbatak #Karobagiandaribatak #batak #fyp
2023-06-07 06:56	<i>jasantaperanginan</i>	tiktok.com	#dosenganteng #dosenmarketing #dosenpariwisata #kalakKaro #karobukanbatak #peranginangin #mejuahjuah
2023-08-02 02:39	<i>anggibangunn</i>	tiktok.com	perpecahan. 🤔🤔🤔 #horas #mejuahjuah #Karo #batak #toba #nandenangin #kalakKaro #berusisada #Karobuka..
2022-08-18 06:34	<i>jasantaperanginan</i>	tiktok.com	#peranginangin #jasantaperanginangin #dosenganteng

2021-08-18 06:56	<i>jasantaperanginan</i>	tiktok.com	#dosenpariwisata	#kalakKaro
			#karobukanbatak	#m...
			#peranginangin	#kalakKaro
			#karobukanbatak	#mejuahjuah

Table 2. Most Active Profiles

No	Account/Profile	Data source	Follower	Number of Mentions
1	<i>gintingbrebaktina</i>	tiktok.com	3290	46
2	<i>bayolubisbukanbatak</i>	tiktok.com	817	17
3	<i>asantaperanginan</i>	tiktok.com	11100	14
4	<i>senoft26</i>	tiktok.com	1358	12
5	<i>rajanapal99</i>	tiktok.com	122	11
6	<i>akal_bulus</i>	tiktok.com	1071	8
7	<i>rimbadriver</i>	tiktok.com	76	5
8	<i>m3cryptogame</i>	tiktok.com	180	4
9	<i>karobukanbatak</i>	tiktok.com	6181	4
10	<i>suku.karo.bukan.batak</i>	tiktok.com	1235	3
11	<i>kudinmugusgus</i>	tiktok.com	3048	3
12	<i>eka_gintingp</i>	tiktok.com	2230	3
13	<i>kilap_sumagan</i>	tiktok.com	589	2
14	<i>khashita</i>	tiktok.com	123	2
15	<i>haganta.kaban</i>	tiktok.com	1118	2
16	<i>amsal_sitepu</i>	tiktok.com	4178	2
17	<i>zefritarigan581</i>	tiktok.com	2091	1
18	<i>vintage-hasasaon</i>	tiktok.com	5056	1
19	<i>t.s.biring.udha</i>	tiktok.com	6811	1
20	<i>sitastasnambur88</i>	tiktok.com	7525	1

Analysis of the most active profiles in the context of the *#karobukanbatak* discussion reveals that several accounts have played a significant role in disseminating information and sparking discussions. Based on the available data, public profiles such as *vintage_hasadaon*, *t.s.biring.udha*, *sitastasnambur88*, *jasantaperanginan*, *gintingbrebaktina*, *kudinmugusgus*, *Karobukanbatak*, *amsal_sitepu*, *senoft26*, *rajanapal99*, *anggibangun*, *eka_ginting_*, *mburakprojectofficial*, *sitastasnambur88*,

akal_bulus, and *bayolubisbukanbatak* have been recorded as the most active.

These profiles may have significantly contributed to the number of mentions and interactions in the discussion, indicating that they strongly influence the community engaged with this topic. Their involvement may include various activities such as posting content, sharing information, and participating in conversations related to *#karobukanbatak*. In addition, profiles such as *khashita*, *gintingbrebaktina*, *bayolubisbukanbatak*,

jasantaperangan, haganta.kaban, amsal_sitepu, and zefritarigan581 are also active profiles in this context. Their involvement in this discussion suggests they may have an audience interested in this topic and contribute to disseminating information and opinions.

In conclusion, these most active profiles play a crucial role in the dynamics of the #karobukanbatak discussion and significantly impact the reach and engagement of this topic on social media.

Table 3. Top Public Profiles

No	Account/Profile	Data Source	Voice Distribution	Influence
1	<i>jasantaperangan</i>	tiktok.com	90.008%	626 656
2	<i>gintingbrebaktina</i>	tiktok.com	31.514%	219 410
3	<i>kudinmugusgus</i>	tiktok.com	14.832%	103 264
4	<i>karobukanbatak</i>	tiktok.com	8.693%	60 520
5	<i>amsal_sitepu</i>	tiktok.com	6.446%	44 880
6	<i>senoft26</i>	tiktok.com	5.034%	35 050
7	<i>rajapanal99</i>	tiktok.com	4.944%	34 420
8	<i>anggibangunn</i>	tiktok.com	3.66%	25 480
9	<i>eka_ginting_</i>	tiktok.com	2.58%	17 960
10	<i>mburakprojectofficial</i>	tiktok.com	2.565%	17 860
11	<i>sitastsnambur88</i>	tiktok.com	2.175%	15 140
12	<i>akal_bulus</i>	tiktok.com	2.111%	14 700
13	<i>bayolubisbukanbatak</i>	tiktok.com	1.903%	13 250
14	<i>kilap_sumagan</i>	tiktok.com	1.54%	10 720
15	<i>suku.Karo.bukan.batak</i>	tiktok.com	1.428%	9940
16	<i>m3cryptogame</i>	tiktok.com	1.2%	8352
17	<i>haganta.kaban</i>	tiktok.com	0.991%	6900
18	<i>adiginting397</i>	tiktok.com	0.618%	4300
19	<i>khashita</i>	tiktok.com	0.577%	4020
20	<i>ma.lingga2</i>	tiktok.com	0.539%	3750

Sentiment analysis related to #karobukanbatak reveals a strong dominance of positive mentions, with 90.008% of total mentions having positive

connotations. Negative mentions only account for 31.514%, indicating that the discussion generally tends to support or have a favourable view of the topic.

Table 4. Most Active Sites

No	Site	Number of Mention
1	tiktok.com	157

2	<i>karogaul.com</i>	6
3	<i>youtube.com</i>	2
4	<i>sorasirulo.com</i>	2

Websites like *tiktok.com*, *karogaul.com*, *youtube.com*, and *sorasirulo.com* have been recorded as the most active sites in the context of 'KaroBukanBatak'

Table 5. Most Influential Sites

No	Site	Number of Mention	Influence Score
1	<i>youtube.com</i>	34 B	10/10
2	<i>tiktok.com</i>	2.4 B	10/10
3	<i>sorasirulo.com</i>	34 553	3/10
4	<i>karogaul.com</i>	13 970	3/10

Several public profiles and social media accounts have been identified as the most active in discussing or disseminating information about '*karobukanbatak*'. The

influence score for some of these accounts reaches 10/10, indicating a high level of influence.

Table 6. Trending Hashtags

No	Hashtag	Number of Mention
1	<i>#karobukanbatak</i>	153
2	<i>#kbbyes</i>	47
3	<i>#zonatiktokkalakaro</i>	46
4	<i>#kalakkaro</i>	44
5	<i>#kalakaro</i>	34
6	<i>#sukukaro</i>	32
7	<i>#sukukaro1838</i>	27
8	<i>#mejuahjuah</i>	23
9	<i>#duet</i>	17
10	<i>#pyf</i>	16
11	<i>#mandailingbukanbatak</i>	16

12	#lagukaro	16
13	#karo	15
14	#peranginangin	15
15	#kalakkarotiktok	10
16	#batak	10
17	#kbb	9
18	#capcut	7
19	#mejuah_juah_kita_kerina_bujur	7
20	#dosenganteng	7

These hashtags reflect the most talked-about topics and aspects on social media in the context of *#karobukanbatak*. Using these hashtags aids in disseminating information and facilitating discussions among users interested in the topic.

DISCUSSION

The research uses sentiment analysis to evaluate tweets with the hashtag *#karobukanbatak*, crucial for understanding public opinion and emotional responses on social media. This analysis helps identify whether the sentiments are positive, negative, or neutral, providing insights into users' collective mood and attitudes toward *Karo* ethnic identity (Estrada et al., 2022; Lachmar, 2017; Stella et al., 2020). Social media platforms like Twitter are instrumental in expressing and shaping ethnic identities. They offer a space for individuals to share their cultural narratives and discuss their heritage. The hashtag *#karobukanbatak* is used to assert and differentiate *Karo*'s identity from the broader Batak ethnic group, highlighting the role of digital activism in mobilizing community support and raising awareness about cultural distinctions (Estrada et al., 2022; Lachmar, 2017; Stella et al., 2020).

The study highlights the importance of preserving regional culture in the context of globalization. Social media and technology have facilitated cross-cultural communication and pose challenges in maintaining and interpreting cultural identity. A sentimental analysis of *#karobukanbatak* can reveal how the *Karo* community perceives and handles its cultural legacy in the digital age. The research explores how social media influences public perception and communication about cultural aspects of *Karo* society. By analyzing the sentiment of *#karobukanbatak*, the study sheds light on how society views and reacts to various aspects of *Karo* culture, reflecting the community's cultural identity and values.

The study acknowledges potential challenges, such as the brief nature of tweets limiting the depth of analysis and the difficulty in capturing all relevant tweets due to variations in hashtag usage. Despite these challenges, the research provides valuable insights into the role of digital platforms in cultural preservation and identity assertion, contributing to understanding how social media influences ethnic identity and community dynamics (Estrada et al., 2022; Lachmar, 2017; Stella

et al., 2020). The impact of social media discussions on the mental well-being of the *Karo* community is also considered. Social media can provide emotional support and a sense of connection, which is particularly important for minority groups. This aspect underscores the dual role of social media in both challenging and supporting cultural identity (Estrada et al., 2022; Lachmar, 2017; Stella et al., 2020)

The study found that positive sentiments dominated up to 90.008% of all *#karobukanbatak* on social media platforms. This indicates a strong positive perception and support for the *Karo* ethnic identity among users, highlighting the effectiveness of social media in fostering cultural pride and community solidarity. The literature review emphasizes the role of sentiment analysis in extracting subjective information from social media, which is crucial for understanding public opinion and emotional responses (Estrada et al., 2022; Lachmar, 2017; Stella et al., 2020). The findings align with this by demonstrating how sentiment analysis can reveal positive attitudes toward *Karo*'s identity, thus validating the use of NLP techniques in cultural studies. The study employed machine learning (ML) and deep learning (DL) models to analyze the sentiment of *#karobukanbatak*. These models include Logistic Regression, Random Forests, and neural networks, commonly used for sentiment classification due to their flexibility and competitive accuracy.

The study employed Aspect-Based Sentiment Analysis, which focuses on identifying and analyzing sentiments related to specific aspects or features within the text. It allows for a more nuanced understanding of how different elements of the *Karo* culture are perceived. The aspect-based sentiment analysis highlighted

traditional elements of *Karo* culture, such as music, dance, art, and customs. These elements are integral to the cultural identity of the *Karo* people and are often discussed in social media posts under the hashtag *#karobukanbatak*. The unique language and communication methods of the *Karo* community were also emphasized. The *Karo* language is a significant part of their cultural identity, and its use in social media discussions reflects its importance to the community. The analysis revealed that the *Karo* community upholds certain values and beliefs, which are frequently mentioned in discussions. These values are crucial for maintaining the cultural identity and cohesion of the *Karo* people (Damanik, 2018; Kipp & Kipp, 1983; Loeb & Heine-Geldern, 1989; Putro, 1981).

Sentiment Distribution and Evolution of *#karobukanbatak* on Social Media

- a. The sentiment analysis of *#karobukanbatak* on social media reveals a strong dominance of positive mentions. Specifically, 90.008% of the total mentions have positive connotations, indicating a generally favourable view of the topic within the online community
- b. In contrast, negative mentions account for only 31.514% of the total, suggesting that while there are some negative perceptions, they are significantly outnumbered by positive sentiments
- c. The research aims to understand how these sentiments evolve. This involves analyzing the distribution of sentiments and observing any changes or trends in public perception as expressed through social media discussions
- d. Social media platforms play a crucial role in shaping and reflecting public

perception and communication about the cultural aspects of *Karo* society. The discussions under *#karobukanbatak* help in affirming ethnic identity and disseminating information on cultural topics, which can influence the evolution of sentiments over time

- e. Public and social media accounts have been identified as the most active in discussing or disseminating information about *#karobukanbatak*. These profiles significantly impact the reach and engagement of the topic, potentially influencing the sentiment distribution and its evolution

Influence of Social Media on Public Perception and Communication of *Karo* Culture

- a. Social media significantly influences how the public perceives and communicates about the cultural aspects of *Karo* society. The hashtag *#karobukanbatak* serves as a focal point for discussions, allowing users to express their thoughts and opinions, which collectively shape public perception
- b. The discussions around *#karobukanbatak* reflect the *Karo* community's cultural identity and values. Social media provides a platform for the *Karo* people to express pride in their heritage and engage in conversations that reinforce their cultural legacy in the digital age
- c. By analyzing sentiments associated with *#karobukanbatak*, researchers can understand societal reactions to *Karo* culture. This analysis helps reveal how social media discussions influence public opinion and the perception of cultural characteristics
- d. Social media facilitates cross-cultural communication, making sharing and interpreting cultural identities easier.

However, it also presents challenges in maintaining cultural identity, as diverse opinions and sentiments can lead to polarized views within the community

- e. Certain public accounts and profiles actively discuss and share information related to *#karobukanbatak*, amplifying its reach and influence. These active participants play a crucial role in shaping the narrative and public perception of *Karo* culture on social media

CONCLUSION

1. The sentiment analysis of the hashtag *#karobukanbatak* on social media reveals a strong dominance of positive mentions, with 90.008% of the total mentions having positive connotations. This indicates a generally favourable view of the topic within the online community. In contrast, negative mentions account for only 31.514% of the total, suggesting that while there are some negative perceptions, they are significantly outnumbered by positive sentiments. The research also aims to understand how these sentiments evolve, observing any changes or trends in public perception expressed through social media discussions. The discussions under *#karobukanbatak* help affirm ethnic identity and disseminate information on cultural topics, which can influence the evolution of sentiments over time. Social media platforms play a crucial role in shaping and reflecting public perception and communication about the cultural aspects of *Karo* society.
2. Social media significantly influences how the public perceives and communicates about the cultural aspects of *Karo* society. It provides a platform for the *Karo* people to express pride in their heritage and engage in

conversations reinforcing their cultural legacy in the digital age. By analyzing sentiments associated with #karobukanbatak, researchers can understand societal reactions to Karo culture. This analysis helps reveal how social media discussions influence public opinion and the perception of cultural characteristics. The hashtag #karobukanbatak is a focal point for discussions, allowing users to express their thoughts and opinions, collectively shaping public perception. The discussions around #karobukanbatak reflect the Karo community's cultural identity and values. Social media facilitates cross-cultural communication, making sharing and interpreting cultural identities easier, although it also presents challenges in maintaining cultural identity due to diverse opinions and sentiments.

Bibliography

- Agarwal, V., Tayal, A., Dixit, R., Chugh, B., & Jain, S. (2024). Exploring ML Methods for Sentiment Analysis in Text Data. *2024 International Conference on Intelligent Systems for Cybersecurity (ISCS)*, 1–6. <https://doi.org/10.1109/ISCS61804.2024.10581249>
- Chakraborty, S. (2023). Sentiment Analysis in the Perspective of Natural Language Processing. *International Journal for Research in Applied Science and Engineering Technology*, 11, 2235–2241. <https://doi.org/10.22214/ijraset.2023.56925>
- Chen, Y. (2024). A Sentiment Analysis of Medical Text Based on Deep Learning. *ArXiv*, abs/2404.1. <https://api.semanticscholar.org/CorpusID:269157234>
- Cheng, B., Lu, S., Duan, Y., Wang, S., & Zhang, Q. (2023). Prompt Learning for Subjective and Objective Recognition of Text. *2023 IEEE 6th International Conference on Computer and Communication Engineering Technology (CCET)*, 1–5. <https://doi.org/10.1109/CCET59170.2023.10335159>
- Dallas, C., Kelpšienė, I., Laužikas, R., & Gribovskis, J. (2024). Socialinių tinklų tyrimų duomenų archyvo funkciniai reikalavimai, projektavimas ir įgyvendinimas. *Information & Media (Vilnius University)*, 100, 22–37. <https://api.semanticscholar.org/CorpusID:270550285>
- Damanik, E. L. (2018). Menolak Evasive Identity: Memahami Dinamika Kelompok Etnik di Sumatera Utara. *Anthropos: Jurnal Antropologi Sosial Dan Budaya (Journal of Social and Cultural Anthropology)*, 4(1). <https://doi.org/10.24114/antro.v4i1.9970>
- Estrada, M. S., Juarez, Y., & Piña-García, C. A. (2022). Toxic Social Media: Affective Polarization After Feminist Protests. *Social Media and Society*, 8(2). <https://doi.org/10.1177/20563051221098343>
- Guerrero-Contreras, G., Balderas-Díaz, S., Serrano-Fernández, A., & Muñoz, A. (2024). Enhancing Sentiment Analysis on Social Media: Integrating Text and Metadata for Refined Insights. *2024 International Conference on Intelligent*

- Environments (IE)*, 62–69.
<https://doi.org/10.1109/IE61493.2024.10599899>
- Kipp, R. S., & Kipp, R. D. (1983). *Beyond Samosir : recent studies of the Batak peoples of Sumatra*.
<https://api.semanticscholar.org/CorpusID:152558912>
- Kusal, S., Pirbhai, A., & Patil, S. (2024). Meta-Learning in Textual Sentiment and Emotion Analysis: A Comprehensive Review. *2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon)*, 1–8.
<https://doi.org/10.1109/MITADTSoCiCon60330.2024.10574942>
- Lachmar, E. M. (2017). #MyDepressionLooksLike: Examining public discourse about depression on Twitter. *JMIR Mental Health*, 4(4).
<https://doi.org/10.2196/mental.8141>
- Loeb, E. M., & Heine-Geldern, R. (1989). *Sumatra: Its History and People*. Oxford University Press.
<https://books.google.co.id/books?id=jqFGAAAAMAAJ>
- Lyons, J. B., Beck, J. C., Dumas, R. E., Rice, C. L., & Moe, P. J. (2019). *Social media platform* (Patent US 20190037264A1).
<https://patents.google.com/patent/US20190037264A1/en>
- Nikhil Sanjay Suryawanshi. (2024). Sentiment analysis with machine learning and deep learning: A survey of techniques and applications. *International Journal of Science and Research Archive*, 12(2), 005–015.
<https://doi.org/10.30574/ijrsra.2024.12.2.1205>
- Page, R. (2018). *Narratives online: shared stories in social media*. Cambridge University Press.
<https://research.birmingham.ac.uk/en/publications/narratives-online-shared-stories-in-social-media>
- Putro, B. (1981). *Karo, dari jaman ke jaman*.
<https://api.semanticscholar.org/CorpusID:160812806>
- Ranjan, M., Tiwari, S., Sattar, A. M., & Tatkar, N. S. (2023). A New Approach for Carrying Out Sentiment Analysis of Social Media Comments Using Natural Language Processing †. *Engineering Proceedings*, 59(1), 1–9.
<https://doi.org/10.3390/engproc2023059181>
- Srinidhi, K., Harika, A., Voodarla, S. S., & Abhiram, J. (2024). Sentiment Analysis of Social Media Comments using Natural Language Procesing. *2024 International Conference on Inventive Computation Technologies (ICICT)*, 762–768.
<https://doi.org/10.1109/ICICT60155.2024.10544849>
- Stella, M., Restocchi, V., & De Deyne, S. (2020). #Lockdown: Network-enhanced emotional profiling in the time of covid-19. *Big Data and Cognitive Computing*, 4(2), 1–23.
<https://doi.org/10.3390/bdcc4020014>
- Sun, W., Ye, B., & Li, Y. (2024). Analysis of factors influencing cultural communication on social media and prediction of dissemination effectiveness. *MIPPR 2023: Remote Sensing Image Processing, Geographic Information Systems, and Other Applications*, 13088, 130–144.
<https://journals.spiedigitallibrary.org/>

- conference-proceedings-of-spie/13088/130880K/Analysis-of-factors-influencing-cultural-communication-on-social-media-and/10.1117/12.3000618.short
- Thakur, M., Arora, A. M., Sandhu, Y. S., & Singh, G. (2024). Sentimental Journeys: Novel Insight into Textual Information. *Advancements in Communication and Systems*, 305–313. <https://doi.org/10.56155/978-81-955020-7-3-26>
- Tsiligaridis, J. (2024). *Approaches of Classification Models for Sentiment Analysis*. 81–86. <https://doi.org/10.5121/csit.2024.141007>
- V, V., & Bargavi, S. K. M. (2024). A Intensified Approach on Aspect Based Sentiment Analysis Using Pre-Trained Deep Learning Model. *2024 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, 1–5. <https://doi.org/10.1109/RAEEUCCI61380.2024.10547866>
- Wang, X. (2024). Analysis of neural networks in text sentiment. *Applied and Computational Engineering*, 76(1), 240–244. <https://doi.org/10.54254/2755-2721/76/20240599>
- Winstedt, R. O. (1947). The Suma Oriental of Tomé Pires and the Book of Francisco Rodrigues. Translated From the Portuguese MS. In the Bibliothèque De La Chambre Des Députés, Paris, and Edited by Cortesão Armando. 2 Vols. Pp. I + Xciv, 1-578. Hakluyt Society, London, 1944. *Bulletin of the School of Oriental and African Studies*. <https://doi.org/10.1017/s0041977x00080095>
- Wu, Y., Jin, Z., Shi, C., Liang, P., & Zhan, T. (2024). Research on the Application of Deep Learning-based BERT Model in Sentiment Analysis. *ArXiv*, abs/2403.0. <https://api.semanticscholar.org/CorpusID:268379403>