Available online at
http://journale-academiauitmt.uitm.edu.my
e-Academia Journal

e-Academia Journal of UiTM Cawangan Terengganu 14 (1) 57-72, June 2025

Exploring Sentiment Trends in TikTok Comments Using GPT for Influencer Content Strategy

Muhammad Akmal Hakim Ahmad Asmawi[1], *Pradeep Isawasan[1], Lalitha Shamugam[2], Khairulliza Ahmad Salleh[1] & K.S. Savita [3]

[1]Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA, Perak Branch, Malaysia.

[2]Department of Business Analytics, Sunway Business School,
Sunway University, Malaysia.

[3]Department of Computer and Information Sciences, Positive Computing Center,
Universiti Teknologi PETRONAS, Malaysia.

*Corresponding author's email: pradeep@uitm.edu.my

Submission date: 7 March 2025 Accepted date: 28 May 2025 Published date: 29 June 2025

To cite this article: Asmawi, M. A. H. A., Isawasan, P., Shanmugam, L., Salleh, K. A., Savita, K.S. (2025). Exploring Sentiment Trends in TikTok Comments Using GPT for Influencer Content Strategy. e-Academia Journal of UiTM Cawangan Terengganu, 14 (1) 57-72, June 2025.

ABSTRACT

The growth of TikTok has reshaped social media marketing, with influencer-driven content playing a crucial role in consumer engagement and purchasing decisions. In the Malaysian Beauty and Personal Care category, TikTok comments serve as direct consumer feedback, yet extracting meaningful insights from this highly expressive data remains challenging. Traditional sentiment analysis methods struggle with multilingual text, slang, abbreviations, and emojis, limiting their effectiveness in interpreting user sentiment. This study addresses these challenges by leveraging GPT-based sentiment analysis to analyze TikTok comments, examining sentiment trends, linguistic patterns, and their correlation with influencer revenue. The study focuses on the top 20 highest-revenue influencers within the Malaysian Beauty and Personal Care category, collecting 34,597 comments from 3,912 videos using Apify's TikTok scraping API. The dataset was preprocessed using GPT-based text normalization, slang resolution, and emoji-to-text conversion, ensuring consistency in sentiment classification. It categorized comments into Positive, Neutral, or Negative, followed by a detailed examination of frequently used words and sentiment patterns across different influencers. Results indicate strong positive sentiment overall, highlighting consumers' enthusiasm. However, neutral comments typically product-related questions also revealed significant engagement. Interestingly, influencers receiving more neutral inquiries still achieved high revenue, suggesting that active audience interaction, rather than purely positive sentiment, is critical for financial success. This underscores the importance of aligning content and engagement strategies rather than relying solely on sentiment analysis. The findings offer practical guidance for brands and influencers, emphasizing the integration of sentiment insights into broader marketing strategies.

Keywords: *Business Intelligence, Social Media Analytics, TikTok, Sentiment Analysis, User Generated Content*

1.0 INTRODUCTION

The rapid growth of TikTok as a leading social media platform (Harriger et al., 2023) has transformed digital marketing, e-commerce, and brand engagement. With over 1 billion monthly active users (Kemp, 2025), TikTok has become a key channel for influencer-driven marketing, where brands leverage short-form videos and viral trends to drive consumer engagement (Salminen et al., 2024). In particular, the Beauty and Personal Care industry has thrived on TikTok, as influencers play a crucial role in shaping product perception, creating brand awareness, and influencing purchasing decisions (Ahmad Asmawi & Isawasan, 2024; Darmatama & Erdiansyah, 2021). Unlike traditional advertising, TikTok interactions are largely user-generated, with consumers actively commenting, reacting, and sharing product experiences. These comments which are rich in sentiment, serve as direct consumer feedback, providing valuable insights into public perception, brand reception, and market trends (Farkas & Schwartz, 2018; Shien et al., 2023). These behavioral expressions are often influenced by underlying psychological traits. A study by (Ali, 2021) found that Malaysian social media users with higher openness to experience and extraversion showed stronger purchase intentions, suggesting that expressive engagement on platforms like TikTok may reflect a user's predisposition toward product interest and buying behavior.

However, analyzing sentiment in TikTok comments presents several challenges. Unlike structured product reviews on e-commerce platforms, TikTok comments are highly unstructured, multilingual, and often filled with slang, abbreviations, and emojis (Cheng & Li, 2024). The informal nature of TikTok frequently presents challenges for traditional sentiment analysis techniques as rule-based and lexicon-based methods often struggle to reliably capture nuances such as localized slang, mixed languages, abbreviations, and emotional expressions through emojis. (Ray & Chakrabarti, 2022; Thangavel & Lourdusamy, 2023). Conventional machine learning methods find it difficult to properly categorize sentiment with mixture of words, symbols, and visual elements to describe emotions (Cho, 2024). Consequently, this inability to accurately interpret user-generated sentiment impedes brands capability to effectively align content strategies with consumer expectations and behaviors, potentially limiting influencer marketing performance and revenue generation. Given these challenges, advanced NLP techniques like GPT-based sentiment analysis offer a more effective approach. By leveraging large-scale language models trained on diverse datasets, GPT can better interpret context, normalize informal language, and improve sentiment classification accuracy in TikTok comments (Zhang et al., 2023).

Unlike previous studies focused on structured sentiment analysis in e-commerce, this research explores TikTok's unique conversational patterns, providing insights into how sentiment appears in short-form user interactions. Specifically, it aims to uncover sentiment trends in influencer-driven engagement, analyze linguistic and emoji-based sentiment patterns, and examine the relationship between audience sentiment and influencer revenue. The study is guided by the following objectives:

- To analyze the sentiment distribution of TikTok comments in the category and identifying dominant sentiment trends.
- To identify common word and emoji-based patterns in consumer sentiment, examining how users express emotions and product perceptions.
- To examine sentiment variations across top influencers by analyzing the distribution of sentiments in their comment sections and correlating sentiment trends with influencer revenue.

Based on the objectives outlined, this study is guided by the following research questions, RQ1: What is the overall sentiment distribution (Positive, Neutral, Negative) of TikTok comments in the Beauty and Personal Care category, RQ2: What linguistic patterns, including words and emojis, are commonly associated with each sentiment category in TikTok comments, and RQ3: How does sentiment distribution vary across top influencers, and how is it related to their revenue performance. The significance of this study is through the practical applications in business intelligence. Brands can gain insights into consumer preferences, product reception, and influencer impact, helping them refine marketing strategies, optimize influencer partnerships, and improve audience targeting. This is particularly valuable for small businesses and low-income entrepreneurs, such as B40 mumpreneurs, who often lack formal marketing knowledge, digital tools, and consistent promotional strategies—factors that significantly limit their marketing effectiveness (Mohamad Fuzi & Mohd Noor, 2024). Additionally, by evaluating how sentiment correlates with influencer revenue, businesses can gain a data-driven perspective on the financial impact of audience engagement. Beyond its business relevance, this research contributes to AI-driven sentiment analysis by demonstrating how GPT-based models handle unstructured and informal social media text more effectively than traditional sentiment analysis methods. As social media platforms continue to shape consumer behavior, advancements in sentiment analysis will be essential for extracting meaningful insights from large-scale user-generated content.

2.0 RELATED WORK

Sentiment analysis has advanced from traditional machine learning techniques to deep learning and transformer-based models, greatly improving how sentiment is interpreted in social media text. Earlier methods relied on classifiers like Naïve Bayes (NB), Support Vector Machines (SVM), and logistic regression, which required manual feature engineering and struggled to capture contextual meaning (Pang et al., 2002). The rise of deep learning in the 2010s introduced Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, allowing models to automatically learn feature representations and better understand sequential relationships in text (Gandhi et al., 2021; Kim, 2014). Bidirectional LSTMs (BiLSTMs) further enhanced sentiment classification by integrating both forward and backward context. The emergence of transformer-based models, particularly BERT (Devlin et al., 2018) and RoBERTa (Liu et al., 2019), brought a major breakthrough with self-attention mechanisms, enabling models to analyze entire sentence structures more effectively. BERT's bidirectional contextual embeddings and pretraining on large text corpora helped surpass earlier CNN and LSTM approaches in sentiment classification. However, BERT-based models still struggle with informal, short-form, and highly contextual social media text, where sentiment depends on linguistic variations, sarcasm, and emoji interpretations (Nguyen et al., 2020). Generative models such as GPT-3 and GPT-4, though primarily designed for text generation, have demonstrated competitive performance in sentiment classification through fine-tuning or prompt engineering (Zhang et al., 2023). Comparative studies confirm that GPT-based classifiers outperform traditional models on social media sentiment tasks, particularly in handling informal language and nuanced expressions (Kheiri & Karimi, 2023).

Social media platforms, particularly Twitter, Facebook, and Instagram, have been the primary focus of sentiment analysis research due to their high volume of user-generated content (Tariq et al., 2025). Early studies used lexicon-based and supervised learning methods, with SVM outperforming Naïve Bayes in tweet classification (Saif et al., 2012). The adoption of deep learning, particularly CNNs and LSTMs, improved accuracy, while transformer-based models like BERTweet which pretrained on 850 million tweets, outperformed general-purpose transformers in Twitter sentiment tasks (Nguyen et al., 2020). However, sentiment analysis on social media is challenging due to informality, brevity, sarcasm, and heavy emoji use, which complicate traditional NLP pipelines. While lexicon-based models like VADER (Hutto & Gilbert, 2014) addressed some of these issues, recent advancements favor pretrained transformers that incorporate domain-specific linguistic patterns. Multilingual transformers such as XLM-R have further improved sentiment classification in diverse linguistic contexts. Unlike Twitter, TikTok remains underexplored in sentiment analysis, with existing studies relying on small datasets and classical classifiers such as Naïve Bayes and SVM. TikTok comments pose unique challenges, including extreme brevity, informal and evolving slang, heavy emoji usage, and multilingual content. Context dependency is another critical issue, as sentiment often relates to video content, making standalone text analysis less reliable (Cho, 2024). Few studies have applied deep learning or transformer models to TikTok sentiment. Existing research has primarily employed Random Forests or LSTMs with moderate success (Ahmed Khan et al., 2024; Guia et al., 2019), but no large-scale fine-tuning of GPT-based models has been conducted. Given TikTok's linguistic complexity, transformer models could provide significant improvements over traditional classifiers.

TikTok presents unique linguistic challenges that differ from platforms like Twitter or Facebook, particularly due to its informal, audio-visual, and emoji-heavy comment culture. Recent studies have started to explore these challenges. For instance, (Cho, 2024) analyzed sentiment in TikTok fad diet content and highlighted the platform's high use of sarcasm, informal expressions, and mixed-code language. Similarly, (Cheng & Li, 2024) noted that TikTok comments often include performative humor and emojis that carry context-dependent sentiment cues. These characteristics complicate standard sentiment classification and call for advanced models that can handle multimodal and context-aware interpretations. However, research on sarcasm detection and emoji sentiment interpretation specific to TikTok remains limited, suggesting a growing but underdeveloped niche within NLP research. Addressing these gaps is essential for accurately analyzing user sentiment on TikTok, particularly in influencer-driven contexts where engagement is highly expressive and nuanced. Beyond technical challenges, emerging research in influencer marketing has highlighted that user engagement is shaped less by perceived influencer credibility and more by interactive factors like value co-creation and social media marketing activities. (Xuan et al., 2023) found that these factors are stronger predictors of Brand Engagement in Self-Concept (BESC), especially when product involvement is high. Their findings suggest that influencer campaigns must focus on contextually relevant content and interactive engagement to resonate with audiences, reinforcing the need to analyze sentiment trends as a reflection of deeper consumer-brand alignment.

GPT models have demonstrated strong performance in sentiment classification, rivaling fine-tuned BERT models in several benchmark tasks (Araci, 2019; Mughal et al., 2024). Their main strength is the ability to handle informal, unstructured text without requiring extensive preprocessing. Research shows that GPT-3 generalizes well to unseen data, outperforming traditional models like Naïve Bayes and even LSTMs in sentiment classification (Elmitwalli & Mehegan, 2024). However, GPT models have limitations, particularly in detecting sarcasm, handling aspect-based sentiment analysis, and managing biases from their training data. Additionally, their high computational cost makes real-time, large-scale applications challenging. That said, GPT's ability to process informal language, including slang and emojis, makes it well-suited for TikTok sentiment analysis. Overcoming its limitations would require fine-tuning on TikTok-specific datasets and incorporating context-aware AI techniques to improve sarcasm detection and emoji-based sentiment interpretation. Despite the advancements in transformer-based sentiment analysis, TikTok remains largely understudied, especially in the context of GPT-powered sentiment classification. Most existing research still relies on traditional methods, missing the opportunity to leverage recent NLP breakthroughs. This study fills that gap by applying GPT models to TikTok sentiment analysis, assessing how well they handle slang, emojis, multilingual content, and context-dependent sentiment. Unlike previous studies that focus only on classification, this research goes further by examining the link between sentiment distribution and influencer revenue, offering a fresh perspective on how sentiment-driven engagement impacts business outcomes. The findings contribute both to AI-driven sentiment analysis research and practical applications in influencer marketing and brand analytics.

3.0 METHODOLOGY

This study employs the Data Science Trajectories (DST) model (Martinez-Plumed et al., 2021) as a structured methodology for analyzing TikTok comments (Figure 1). Table 1 summarizes each phase of the DST model and its implementation in this research.

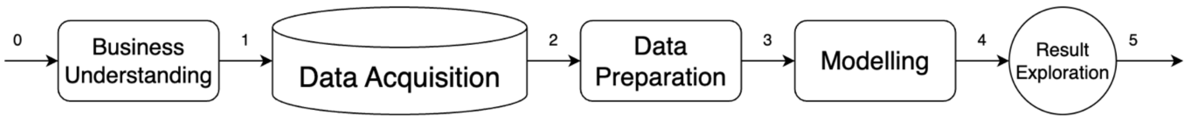

Figure 1: The research methodology

Table 1: DST Methodology and Implementation

DST Phase	Details	Action Takens
Business Understanding	Establish research goals and align them with key digital marketing metrics identified from literature.	• Reviewed existing literature. • Set clear research objectives relevant to brand and influencer marketing.
Data Acquisition	Collect user-generated TikTok comments for sentiment analysis.	• Used Apify, a web scraping tool. • Extracted comprehensive TikTok comment datasets.
Data Preparation	Clean and preprocess comments to handle informal language, slang, emojis, and multilingual text.	• Removed irrelevant data (noise). • Normalized text (expanded slang, handled emojis).
Modelling	Perform sentiment classification using GPT-based sentiment analysis.	• Used Python with GPT-4o API. • Classified comments as Positive, Neutral, or Negative.
Result Exploration	Interpret sentiment results to find correlations with influencer revenue and derive insights useful for TikTok e-commerce strategies	• Analyzed sentiment distribution across influencers. • Explored connections between sentiment patterns and revenue.

3.1 DATA ACQUISITION

This study explores audience sentiment in TikTok's Beauty and Personal Care category, which generated the highest revenue during the data collection period (RM195.39 million). To conduct a large-scale sentiment analysis, comments were gathered from the top 20 highest-earning individual influencers. The first step involved retrieving TikTok videos from the identified influencer profiles. The selection of profiles was based on data from Kalodata, which identified the highest-revenue individual accounts within the Beauty and Personal Care sector. To ensure a systematic and scalable approach, two application programming interfaces (APIs) from Apify were utilized to collect the data. Each account identified was then scraped using TikTok Profile Scraper API, configured to extract up to 300 videos per influencer, prioritizing recent content to capture current audience interactions. Due to differences in content production frequency among influencers, the actual number of videos collected per profile varied. Ultimately, 3,912 videos were extracted, covering content posted from January 1, 2024, onwards. The second step focused on extracting user comments, as they serve as the primary data source for sentiment analysis. The TikTok Comments Scraper API was deployed to collect up to 100 comments per video, prioritizing the most recent and highly engaged comments as determined by TikTok's ranking algorithm. In total, 34,597 comments were retrieved from the collected videos. This sampling approach which is systematic influencer selection, recent content prioritization, and engagement-ranked comment extraction, balances scale, relevance, and visibility, enhancing the validity and contextual relevance of the sentiment analysis. Table 2 shows the summary of data collection details. Following data collection, text preprocessing techniques were applied to normalize the comment text, remove linguistic noise, and convert emojis into textual representations to enhance sentiment classification accuracy.

Table 2: Summary of data collection details

Parameter	Details
Subject	Top 20 influencers with the highest total revenue in the Beauty and Personal Care category
Number of videos per influencer	Up to 300 videos per influencer (actual count varies based on content production)
Number of comments per video	Up to 100 most recent comments per video, ranked by engagement

Data collection date	28th August 2024
Period of the collected videos	1st January 2024 – 28th August 2024
Total number of videos collected	3,912
Total number of comments collected	34,597
Number of comments after grouped by video	2,016

3.2 DATA PREPARATION

Before conducting sentiment analysis, a comprehensive text preprocessing pipeline was implemented to ensure that TikTok comments dataset was clean and standardized. Preprocessing steps played a key role in handling spelling variations, informal text, and multilingual expressions to maintain accurate and meaningful sentiment classification (Hickman et al., 2022). GPT-based text normalization was applied specifically to address these challenges, including,

- Spelling correction and standardization, handling informal contractions and spelling variations.
- Expansion of localized slang and abbreviations common among Malay-speaking users (e.g., expanding “nk bli” to “nak beli”).
- Conversion of emojis to text equivalents, preserving their emotional meaning for sentiment interpretation.

Empty comments or those containing only special characters without clear meaning were removed. Emojis were carefully preserved because of their importance as emotional cues to ensure the sentiment models could accurately interpret intended emotions (Gupta et al., 2023). Table 3 illustrates examples of original TikTok comments compared to their cleaned versions, demonstrating how preprocessing enhanced data quality without distorting the original user sentiment.

Table 3: Example of pre-processed TikTok comments

Original Comments	Cleaned Comments
dunno la knape xleh nk check out	saya tidak tahu kenapa tidak boleh check out
how u know 😂😂😂	macam mana kamu tahu :face_with_tears_of_joy
beli skrg plsss	beli sekarang tolong
😍first😍;aku suke tgok rambut dy smpai aku…	:smiling_face_with_heart-eyes:pertama:smiling_...
result please 🥺;Wow;Result!...	keputusan tolong :pleading_face:;Wow;Keputusan...

Stopwords which is common words with minimal impact were removed to streamline the text. However, sentiment terms like “tidak” (not) and other negations were intentionally kept to preserve the sentiment interpretation (Gupta et al., 2023; Hickman et al., 2022). Combining GPT-based text transformation with standard NLP preprocessing techniques allowed for thorough data cleaning and structuring to ensure the dataset was well-prepared for accurate sentiment classification.

3.3 MODELLING

The sentiment analysis in this study was conducted using OpenAI’s GPT-4o model, accessed via the official OpenAI API, as it effectively interprets informal text, slang, and emojis common in TikTok comments (Wang et al., 2023). Compared to earlier transformer models like BERT or RoBERTa, GPT-4o requires less domain-specific fine-tuning and demonstrates better contextual understanding in low-resource, noisy-text environments (Ahmed Khan et al., 2024; Zhang et al., 2023). Its ability to generate context-aware responses through prompt engineering makes it more flexible and accurate for short-form user-generated content, especially when slang, mixed languages, and emojis are involved. While models like BERT may perform well on structured datasets, GPT-4o offers superior adaptability in unstructured, real-world social media data scenarios such as TikTok. As the objective was business insight generation, no accuracy validation was performed, and GPT-4o was used directly via prompt engineering without a labeled validation dataset. Figure 2 illustrates the sentiment classification approach, detailing each step from prompt engineering to final manual verification to ensure reliable and context-aware sentiment analysis of TikTok comments.

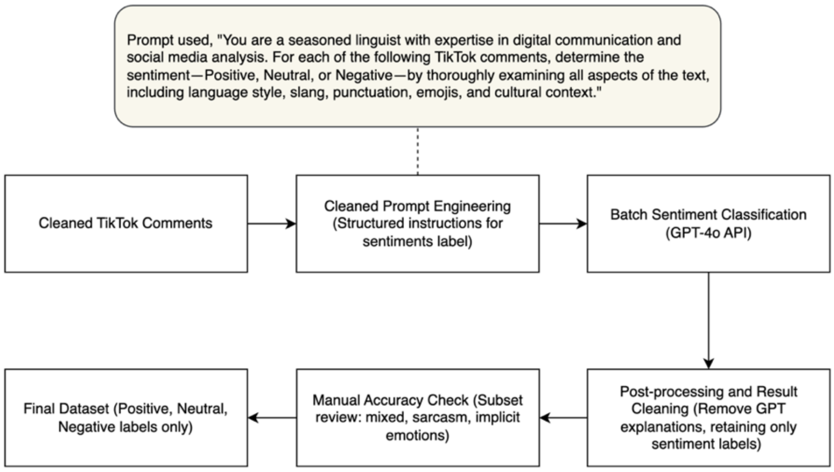

Figure 2: GPT-Based Sentiment Analysis Process

4.0 RESULT AND DISCUSSION

4.1 SENTIMENT DISTRIBUTION IN TIKTOK COMMENTS

The sentiment analysis reveals that TikTok comments about Beauty and Personal Care products are mostly positive (see Figure 3). Out of 2,016 comments, 1,455 (72.18%) are positive, 490 (24.28%) are neutral, and 71 (3.52%) are negative. The insights from each sentiment is described in Table 4 below.

Table 4: Insights from each sentiment

Sentiments	Insights
Positive	• Reflects viewer enthusiasm driven by engaging influencer content, such as beauty tutorials and product reviews. • Positive reactions typically peak during product launches or collaborations, indicating audience excitement.
Neutral	• Consists of factual questions or informational requests about products. • Comments suggest that many consumers seek additional details before making purchasing decisions. • Neutral reactions highlight opportunities for influencers and brands to provide clearer and more comprehensive product information
Negative	• Highlights consumer dissatisfaction regarding product quality or unmet expectations. • Negative comments usually address broader issues like misleading advertisements or ingredient transparency.

	• The low volume might indicate generally positive experiences overall, but it could also reflect users' hesitance to openly criticize products publicly.

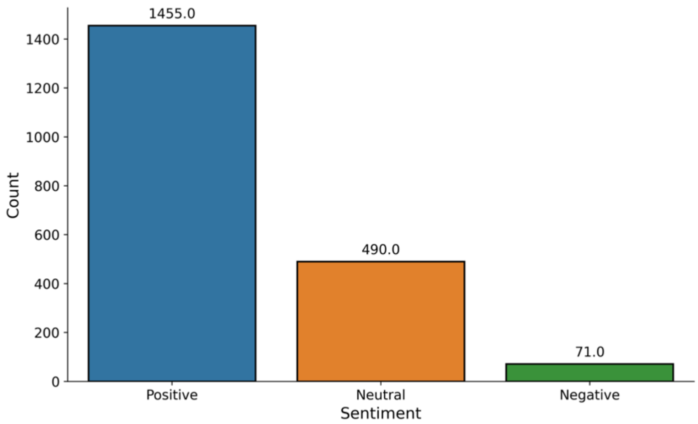

Figure 3: The sentiment distribution

Brands and influencers should make use of the strong positive engagement by continuing to create interactive and engaging content. At the same time, addressing neutral comments by offering clear and detailed product information can drive more confident purchasing decisions. Negative comments provide valuable feedback that should be managed transparently to provide greater trust among followers. In conclusion, understanding sentiment trends allows brands and influencers to better align their content strategies with audience expectations, thus strengthening consumer relationships.

4.2 FREQUENTLY USED WORDS IN TIKTOK COMMENTS

The analysis of frequently used words in TikTok comments in figure 4 reveals how users naturally discuss beauty products and share their opinions and experiences. In positive comments, words like *'saya'* (I) and *'nak'* (want) are common, highlighting users' personal interest in products. The frequent use of 'pakai' (use) indicates active conversations about product usage, application methods, or personal experiences. Emojis such as 'smiling_face_with_hearts' and 'face_with_tears_of_joy' regularly appear, clearly showing excitement and admiration. These expressions suggest strong personal connections users have with beauty products featured by influencers.

For neutral comments, commonly used words include 'saya' (I), 'nak' (want), 'tak' (no/not), and 'boleh' (can), indicating that many users are asking questions or making factual statements rather than expressing strong emotions. The presence of 'smiling_face_with_hearts' and 'face_with_tears_of_joy' in neutral sentiment suggests that even neutral comments may carry some emotional undertones, possibly reflecting mild enthusiasm or sarcasm. Many of these words indicate product inquiries or clarification requests, showing that users rely on TikTok as a platform for discovering and understanding beauty products. Negative comments are relatively rare but important. Common words like 'dia' (he/she) or 'ada' (have/exist) imply that criticism may target specific influencers or products rather than general dissatisfaction. Interestingly, positive emojis sometimes appear even in negative comments, suggesting sarcasm or nuanced emotional expression rather than straightforward negativity.

Overall, this analysis highlights the conversational and multifaceted nature of TikTok interactions. Users freely blend personal experience, humor, and casual language, providing a layered view of audience attitudes toward beauty content. Brands and influencers should use these nuanced insights to improve their communication strategies, directly addressing user questions and carefully responding to criticism, ultimately strengthening their connection with TikTok audiences.

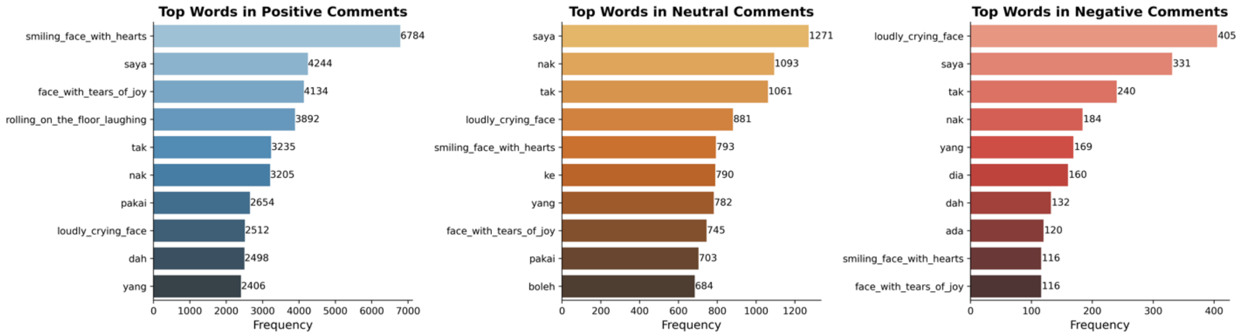

Figure 4: The top 10 words for each sentiments

4.3 SENTIMENT TRENDS IN TIKTOK AUDIENCE ENGAGEMENT

The word cloud visualization of TikTok comments in figure 5 highlights how users express emotions and interact with beauty-related content. Emojis play a prominent role, reflecting strong emotional engagement, while product-focused language underscores active user interest. The identified key themes and their respective insights are illustrated in Table 5 below.

Table 5: Themes identified

Themes	Insights
Emoji-driven expressions of emotion	• Emojis such as "smiling_face_with_heart", "face_with_tears_of_joy", and "rolling_on_the_floor_laughing" are among the most frequent. • This suggests a playful, enthusiastic interaction style typical on TikTok.
Product-focused vocabulary	• Common words like "nak" (want) and "beli" (buy) show that users openly express interest in products, share intentions to purchase, or ask product-related questions. • This indicates direct consumer intent and personal engagement with beauty content.
Language and cultural insights	• Frequent Malay words ("saya", "nak", "pakai", "beli") indicate a predominantly Malay-speaking audience. • The mixture of Malay and English suggests content creators target bilingual viewers, highlighting the value of localized content strategies.
User engagement and content interaction	• Comments often relate directly to popular video content like GRWM, skincare routines, reviews • Users frequently ask questions or express personal experiences, showing active engagement rather than passive consumption

These findings underline TikTok's role as a key platform for beauty marketing, especially among younger Malaysian audiences. Brands and influencers should adapt by creating culturally relevant and emotionally engaging content, such as tutorials, product demonstrations, and reviews, which resonate strongly with local audiences. Understanding these nuances can help brands tailor their marketing approach, fostering deeper connections and improving consumer relationships.

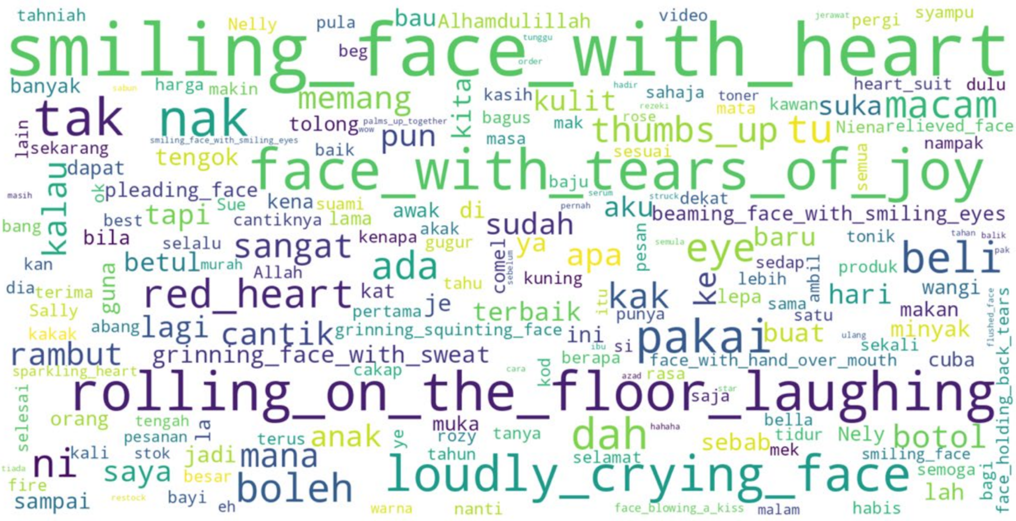

Figure 5: The word cloud of TikTok comments

4.4 SENTIMENT BREAKDOWN ACROSS INFLUENCERS

Figure 6 below shows the sentiment distribution between each influencers. Table 6 below illustrate the key observations from the analysis.

Table 6: Key observations of sentiments breakdown

Sentiments	Insights
Positive Sentiments	• Influencer 09 has the highest total positive comments and no negative feedback, indicating highly successful content or strong audience alignment. • Influencers 14 and 16 also receive notable positive reactions, likely due to engaging content such as viral challenges, collaborations, or appealing product promotions.
Mixed Sentiments	• Influencer 08 displays a balanced mix of positive, neutral, and negative comments, suggesting their content lacks clear audience resonance, pointing to potential areas for improvement.
Negative Sentiments	• Influencer 06 experiences the largest number of negative comments, which could be due to controversies, unmet expectations, or product dissatisfaction.
Minimal Sentiments	• Influencer 13 receives almost no engagement, suggesting they may be less active or have limited visibility, highlighting an opportunity to increase audience interaction.

These findings underline how audience sentiment can significantly impact influencer success within the Beauty and Personal Care category on TikTok. Influencers experiencing mostly positive reactions such as Influencer 09, 14, and 16, should capitalize on their current strengths by maintaining interactive and engaging content. In contrast, influencers dealing with mixed or negative sentiment, notably Influencer 06 and Influencer 08, should carefully reassess their strategies, respond to criticism, and adjust content to better meet audience expectations. Additionally, influencers with minimal engagement, such as Influencer 13, have opportunities to boost visibility and interaction through more interactive formats like live sessions or direct audience engagement. By closely aligning content strategies with audience perceptions and sentiment trends, influencers and brands can effectively enhance their relationships with followers, optimize content performance, and achieve sustained success on TikTok.

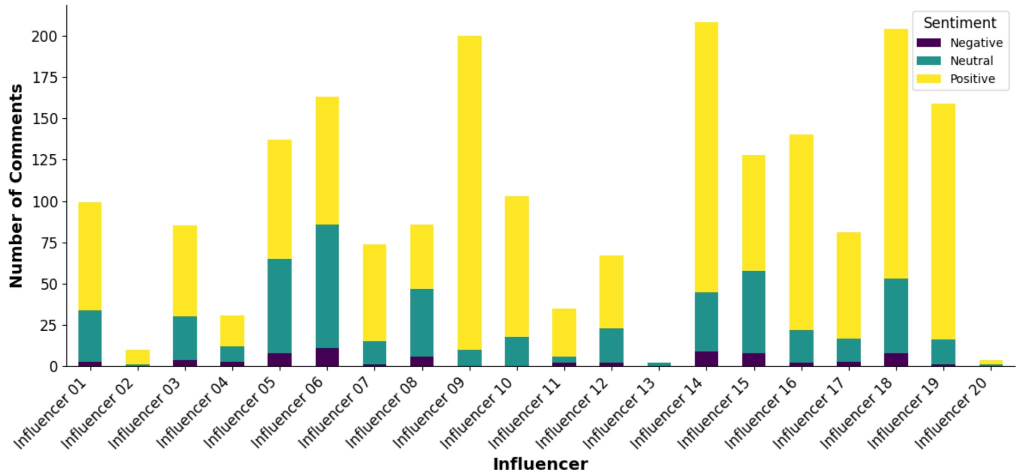

Figure 6: The distribution of sentiments across influencers

4.5 RELATIONSHIP BETWEEN SENTIMENT AND INFLUENCER REVENUE

The data in table 7 shows how influencer sentiment and revenue are related in the beauty and personal care category on TikTok. Most influencers receive more positive comments than negative ones, which suggests that their audience generally reacts well to their content. However, having more positive comments does not always mean that an influencer will earn more revenue. Some influencers with fewer overall comments still generate higher earnings, showing that other factors also play an important role. Revenue differences between influencers are noticeable. Influencer 02 has the highest revenue (RM1,540,000) but very few comments, which suggests that their earnings may come from exclusive brand partnerships or high-value sponsorships, rather than audience engagement. In contrast, Influencer 19 earns RM593,260 while receiving a high number of positive comments, showing that organic engagement and strong audience relationships can also contribute to financial success. However, Influencer 09, who received 190 positive comments and no negative ones, has a much lower revenue of RM322,980. This suggests that even with strong audience support, an influencer may not achieve high earnings unless they effectively monetize their content through promotions, collaborations, or product endorsements.

Some influencers receive a mix of positive, neutral, and negative comments, leading to moderate earnings. Influencers like 05, 15, and 17 fall into this category, showing that consistent engagement can support stable revenue over time. They likely focus on long-term audience trust and steady content quality, which helps them maintain a balance between engagement and earnings. Influencer 13 and 20, for example, have very few comments but still generate considerable earnings. This could be due to targeted brand sponsorships or niche marketing strategies, where they earn from direct promotions rather than relying on audience interaction. Their earnings may come from product collaborations or influencer marketing campaigns that prioritize brand deals over organic engagement.

Several real-world factors could explain these trends. Some influencers may have gained revenue from viral beauty trends, short-term promotional campaigns, or trending challenges, which helped increase product visibility without necessarily increasing audience interaction. Consumer behavior may also play a role, as audiences today tend to favor influencers who are seen as authentic and trustworthy, often engaging more with those who share genuine product reviews and relatable content. In summary, having positive sentiment does not always lead to higher revenue. While strong audience support can help an influencer's brand, earnings often depend on business strategies, partnerships, and content monetization efforts. For brands and influencers looking to improve their financial success, understanding these patterns can help them refine their content approach and make better decisions inTikTok beauty industry.

Table 7: Influencer's distribution of sentiment and their corresponding revenue

Influencer	Negative	Neutral	Positive	Revenue
Influencer01	3	31	65	302240
Influencer02	0	1	9	1540000
Influencer03	4	26	55	429780
Influencer04	3	9	19	330480
Influencer05	8	57	72	305310
Influencer06	11	75	77	311150
Influencer07	1	14	59	545570
Influencer08	6	41	39	451530
Influencer09	0	10	190	322980
Influencer10	0	18	85	346310
Influencer11	2	4	29	361520
Influencer12	2	21	44	348890
Influencer13	0	2	0	328490
Influencer14	9	36	163	446480
Influencer15	8	50	70	485200
Influencer16	2	20	118	418690
Influencer17	3	14	64	548190
Influencer18	8	45	151	577870
Influencer19	1	15	143	593260
Influencer20	0	1	3	304490

5.0 ACTIONABLE INSIGHTS

5.1 OBJECTIVE 1 - ANALYZING THE SENTIMENT DISTRIBUTION OF TIKTOK COMMENTS

The sentiment analysis reveals that Positive sentiment is dominant, while Neutral comments play a significant role in audience engagement, and Negative sentiment is minimal. The high frequency of purchase-related words in Positive comments suggests that consumer excitement and admiration for products and influencers are key drivers of engagement. Neutral sentiment primarily consists of product inquiries, demonstrating that users engage not only emotionally but also for informational purposes. For brands and marketers, the prevalence of Neutral sentiment highlights the importance of engaging with potential consumers through informative responses. Businesses should optimize their content strategies by incorporating interactive Q&A sessions, influencer-led product demonstrations, and comment engagement initiatives to convert Neutral discussions into purchases. Furthermore, the minimal presence of Negative sentiment suggests that TikTok is a favorable platform for brand perception, but businesses should remain vigilant in monitoring consumer concerns and addressing dissatisfaction promptly to maintain credibility.

5.2 OBJECTIVE 2 - IDENTIFYING LINGUISTIC AND EMOJI-BASED PATTERNS IN CONSUMER SENTIMENT

The frequent use of emojis as sentiment indicators suggests that non-verbal expressions are crucial in interpreting TikTok user sentiment. Positive sentiment is characterized by joyful emojis and purchase-related terms, while Neutral sentiment is often expressed through factual words related to product inquiries and decision-making processes. Negative sentiment includes dissatisfaction-related words and crying-face emojis, often linked to unmet expectations or product concerns. For sentiment analysis researchers and AI developers, these findings indicate that future sentiment models should integrate multimodal understanding, incorporating emoji interpretation alongside text processing. Traditional sentiment classifiers that ignore emojis may misinterpret the intended sentiment. For businesses and influencers, leveraging emojis in their own responses and marketing materials could enhance emotional engagement and resonate better with TikTok's informal communication style.

5.3 OBJECTIVE 3 - EXAMINING SENTIMENT VARIATIONS ACROSS INFLUENCERS AND THEIR CORRELATION WITH REVENUE

While Positive sentiment dominates across all influencers, Neutral sentiment varies, particularly among those focused on educational content, product reviews, or tutorials. Influencers with more Neutral comments facilitate discussions and product inquiries, rather than just emotional reactions. Conversely, those with both Positive and Negative sentiment often engage in aggressive marketing or controversial promotions. While high Positive sentiment is linked to engagement, revenue data shows sentiment alone does not determine earnings. Some influencers with high Positive sentiment earn less, while others with fewer sentiment-driven interactions generate more revenue, likely due to live selling, brand collaborations, and sponsorships. Neutral sentiment plays a key role in consumer decision-making, signaling an engaged audience seeking product insights. Influencers should leverage this by fostering discussions and addressing queries to boost trust and conversions. Those facing higher Negative sentiment should assess audience perception and brand alignment to maintain credibility. For marketers, these findings suggest sentiment alone is not a reliable revenue predictor. Brands should analyze Neutral sentiment as an indicator of purchase intent and prioritize influencers who effectively convert engagement into sales rather than just generating high sentiment-driven interactions.

6.0 FUTURE WORK

Future research should extend this study by applying sentiment and revenue analysis across different TikTok product categories, such as fashion, electronics, or food, to determine whether the observed patterns in the Beauty and Personal Care segment hold in other domains. Investigating real-time sentiment monitoring during live sessions could offer deeper insights into how immediate audience reactions influence content adjustments and sales outcomes. Additionally, exploring how influencer engagement with Neutral comments such as replying to product inquiries or clarifying details that impacts trust and conversion rates would provide practical implications for content strategy. Further refinement of sentiment models is also necessary, particularly in enhancing their ability to detect sarcasm, interpret mixed sentiment, and accurately process emoji-rich or context-heavy informal language. These future directions will help develop more adaptive and accurate sentiment analysis tools while broadening the practical value of GPT-based methods in social media marketing and business intelligence.

7.0 CONCLUSION

This study examined GPT-based sentiment analysis of TikTok comments in the Beauty and Personal Care category, focusing on sentiment trends, linguistic patterns, and their correlation with influencer revenue. Positive sentiment was the most prevalent, often indicating excitement and purchase intent. However, Neutral sentiment emerged as a critical engagement type, frequently reflecting product-related inquiries and decision-making behaviors. A key contribution of this study lies in its emphasis on Neutral sentiment as a meaningful form of consumer interaction. Unlike traditional sentiment analysis that focuses primarily on polarity extremes, Neutral comments often signal active interest and intent to purchase. Addressing these comments can foster trust, clarify consumer doubts, and enhance conversion potential.

Importantly, the findings show that sentiment volume alone does not reliably predict influencer revenue. Factors such as content monetization strategies, live selling formats, and brand collaborations play a more significant role in financial outcomes. This highlights the need for brands to combine sentiment analysis with performance metrics to evaluate influencer impact more effectively. By offering insight into the nuances of TikTok comment sentiment, this study provides practical guidance for optimizing influencer marketing strategies and contributes to the growing literature on AI-driven sentiment analysis in informal, multilingual social media contexts.

Declaration of Generative AI and AI-assisted technologies

This work was prepared with assistance from generative AI (ChatGPT) for language and readability improvements. The author(s) oversaw all edits and accept full responsibility for the final content.

REFERENCES

Ahmad Asmawi, M. A. H., & Isawasan, P. (2024). *Top 20 TikTok Beauty & Personal Care Influencers*. Kaggle. https://doi.org/10.34740/KAGGLE/DS/6111672

Ahmed Khan, T., Sadiq, R., Shahid, Z., Alam, M. M., & Mohd Su'ud, M. (2024). Sentiment Analysis using Support Vector Machine and Random Forest. *Journal of Informatics and Web Engineering*, *3*(1), 67–75. https://doi.org/10.33093/jiwe.2024.3.1.5

Ali, A. (2021). Understanding Purchase Intention for Different Personality Traits in Social Networking Services. *E-Academia Journal of UiTM Cawangan Terengganu*, *2*(10), 44–58. https://e-ajuitmct.uitm.edu.my/v3/e

Araci, D. (2019). *FinBERT: Financial Sentiment Analysis with Pre-trained Language Models*.

Cheng, Z., & Li, Y. (2024). Like, Comment, and Share on TikTok: Exploring the Effect of Sentiment and Second-Person View on the User Engagement with TikTok News Videos. *Social Science Computer Review*, *42*(1), 201–223. https://doi.org/10.1177/08944393231178603

Cho, B. (2024). A Sentiment Analysis to investigate the sentiment of users on fad diet content on Tik Tok. In *Journal of High School Science* (Vol. 8, Issue 4).

Darmatama, M., & Erdiansyah, R. (2021). *The Influence of Advertising in Tiktok Social Media and Beauty Product Image on Consumer Purchase Decisions*.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 4171–4186. http://arxiv.org/abs/1810.04805

Elmitwalli, S., & Mehegan, J. (2024). Sentiment analysis of COP9-related tweets: a comparative study of pre-trained models and traditional techniques. *Frontiers in Big Data*, *7*. https://doi.org/10.3389/fdata.2024.1357926

Farkas, J., & Schwartz, S. A. (2018). Please like, comment and share our campaign!: How social media managers for danish political parties perceive user-generated content. In *Nordicom Review* (Vol. 39, Issue 2, pp. 19–33). Sciendo. https://doi.org/10.2478/nor-2018-0008

Gandhi, U. D., Malarvizhi Kumar, P., Chandra Babu, G., & Karthick, G. (2021). Sentiment Analysis on Twitter Data by Using Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM). *Wireless Personal Communications*. https://doi.org/10.1007/s11277-021-08580-3

Guia, M., Silva, R., & Bernardino, J. (2019). Comparison of Naïve Bayes, Support Vector Machine, Decision Trees and Random Forest on Sentiment Analysis. *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 525–531. https://doi.org/10.5220/0008364105250531

Gupta, S., Singh, A., & Kumar, V. (2023). Emoji, Text, and Sentiment Polarity Detection Using Natural Language Processing. *Information*, *14*(4), 222. https://doi.org/10.3390/info14040222

Harriger, J. A., Thompson, J. K., & Tiggemann, M. (2023). TikTok, TikTok, the time is now: Future directions in social media and body image. *Body Image*, *44*, 222–226. https://doi.org/10.1016/j.bodyim.2023.01.005

Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. (2022). Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. *Organizational Research Methods*, *25*(1), 114–146. https://doi.org/10.1177/1094428120971683

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Kemp, S. (2025, February 5). *Digital 2025: Global Overview Report*.

Kheiri, K., & Karimi, H. (2023). *SentimentGPT: Exploiting GPT for Advanced Sentiment Analysis and its Departure from Current Machine Learning*.

Kim, Y. (2014). *Convolutional Neural Networks for Sentence Classification*. http://arxiv.org/abs/1408.5882

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. http://arxiv.org/abs/1907.11692

Mohamad Fuzi, A., & Mohd Noor, N. H. (2024). Application of the marketing mix by small business: A focus group analysis of B40 Mumpreneurs. *E-Academia Journal of UiTM Cawangan Terengganu*, *2*, 115–127. https://www.researchgate.net/publication/386220116

Mughal, N., Mujtaba, G., Shaikh, S., Kumar, A., & Daudpota, S. M. (2024). Comparative Analysis of Deep Natural Networks and Large Language Models for Aspect-Based Sentiment Analysis. *IEEE Access*, *12*, 60943–60959. https://doi.org/10.1109/ACCESS.2024.3386969

Nguyen, D. Q., Vu, T., & Nguyen, A. T. (2020). *BERTweet: A pre-trained language model for English Tweets*. http://arxiv.org/abs/2005.10200

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning. *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - EMNLP '02*, *10*, 79–86. https://doi.org/10.3115/1118693.1118704

Ray, P., & Chakrabarti, A. (2022). A Mixed approach of Deep Learning method and Rule-Based method to improve Aspect Level Sentiment Analysis. *Applied Computing and Informatics*, *18*(1/2), 163–178. https://doi.org/10.1016/j.aci.2019.02.002

Saif, H., He, Y., & Alani, H. (2012). *Semantic Sentiment Analysis of Twitter* (pp. 508–524). https://doi.org/10.1007/978-3-642-35176-1_32

Salminen, J., Wahid, R., Yang, Y., & Jansen, B. (2024). Engagement Patterns in TikTok: An Analysis of Short Video Ads. *Proceedings of the 35th ACM Conference on Hypertext and Social Media*, 323–329. https://doi.org/10.1145/3648188.3677048

Shien, O. Y., Huei, N. S., & Yan, N. L. (2023). The Impact of Social Media Marketing on Young Consumers' Purchase Intention in Malaysia: The Mediating Role of Consumer Engagement. *International Journal of Academic Research in Business and Social Sciences*, *13*(1). https://doi.org/10.6007/IJARBSS/v13-i1/15806

Tariq, R., Isawasan, P., Shamugam, L., Ahmad Asmawi, M. A. H., Nor Azman, N. A., & Zolkepli, I. A. (2025). Exploring Social Media Research Trends in Malaysia using Bibliometric Analysis and Topic Modelling. *ICST Transactions on Scalable Information Systems*, *12*. https://doi.org/10.4108/eetsis.7003

Thangavel, P., & Lourdusamy, R. (2023). A lexicon-based approach for sentiment analysis of multimodal content in tweets. *Multimedia Tools and Applications*, *82*(16), 24203–24226. https://doi.org/10.1007/s11042-023-14411-3

Wang, Z., Xie, Q., Feng, Y., Ding, Z., Yang, Z., & Xia, R. (2023). *Is ChatGPT a Good Sentiment Analyzer? A Preliminary Study*. http://arxiv.org/abs/2304.04339

Xuan, C. J., Boon Liat, C., Chalke, A., & Idris, I. (2023). Brand Engagement in Self-Concept for Influencer Marketing: The Role of Product Involvement as A Moderator. *E-Academia Journal of Universiti Teknologi MARA Cawangan Terengganu*, *12*(1), 90–104. http://journale-academiauitmt.uitm.edu.mye

Zhang, W., Deng, Y., Liu, B., Pan, S. J., & Bing, L. (2023). *Sentiment Analysis in the Era of Large Language Models: A Reality Check*. http://arxiv.org/abs/2305.15005