Decoding Viewer Reactions: Sentiment and Emoji Analysis on YouTube

Satyendra Sharma [1*] Makhan Kumbhakar [2]
Vandit Hedau[3] and Vijay Baboo Gupta[4]

[*1]Scholar, DAVV, Indore India, [2, 3,4]Professor, DAVV, Indore, India
[1*]satyendra.sage@gmail.com, [2]kumbhkar010385@gmail.com, [3]vanditfsp@gmail.com
[4]vbgupta.davv@gmail.com

Abstract. In the modern digital era, YouTube comments are not only words, but they show the sentiments of the viewers with lots of insights. This paper uses YouTube API to analyze comments and extract valued information about viewer sentiments. The present study helps content creators, marketers, and decision-makers understand the audience's reaction to the specific video(s). The exploratory data analysis has been conducted to reveal the user engagement pattern, emphasizing sentiment analysis, usage of emojis, and trending tags. The dataset used for the research focused on a motivational video that highlights the emotional influence of content and the part of emojis in the engagement of viewers. The emoji expressions and trending tag analysis have been standardized to have an inclusive view of the dynamic environment of YouTube. This paper bridges the gap between content creators, promoters, and researchers, and also makes them capable of strategy optimization to improve user experiences depending on decoded sentiments and user interactions on this comprehensive platform.
Keywords: YouTube Comments, Sentiment analysis, user engagement, emoji analysis, trending tags analysis.

1 Introduction

In the modern age, when digital content is supreme, YouTube is one of the peerless giants that provides a virtual platform. Here, creators, viewers, and advertisers join together. YouTube is not merely a platform, as more than one billion hrs. of videos are viewed daily. It is now like a cultural incident, an international forum that affects trends and nurtures societies. When someone tries to understand the complex matrix of the user's reactions, sentiments, and continuous connectivity, the person finds their own at the intersection point of psychology, digital anthropology, and data science.

Let us first understand the immensity of YouTube. It has reached around 50% of all net users including miscellaneous interests/demographics for knowledge, fun, or motivation. Mobile handsets account for 37 % of all the net traffic, and YouTube is at the frontier among them. The content available on YouTube varies from pet antics to cutting-edge technology applications, which is shared by creators, influencers, teachers, entertainers, and devotees. Whether someone is learning salad cutting or

S. Bhalerao et al. (eds.), *Proceedings of the International Conference on Recent Advancement and Modernization in Sustainable Intelligent Technologies & Applications (RAMSITA-2025),* Advances in Intelligent Systems Research 192,
https://doi.org/10.2991/978-94-6463-716-8_5

understanding the secrets of the universe, YouTube provides a personal experience. Under each video, there is a comment section for viewers to give their sentimental reactions, and have chats with others. Such comments are valuable as they capture real emotions and honest opinions, and create connections.

Understanding the comment is very important for content creators, it helps them know the viewer's interest, and decide the length of the video and the kind of thumbnail to be used. The sentiment analysis focused on the audience's reactions guides creators to update the content to better connect with the viewers. A reaction with a thumbs-up emoji means liked by someone, whereas a fire emoji shows excitement. The same is the case with marketers and advertisers, they can decide how to make ads for improved emotional connections. The analysis helps them build a strong image in the market. An emoji with tears may show empathy, on the other hand, a laughing emoji shows fun. Researchers keep their eyes on YouTube to find the user's behavior, current market trends, and the influence of online communities. The sentiment analysis helps explore how people feel in groups and how the influencers affect them. A heart-eyed emoji and a facepalm emoji show admiration and frustration respectively.

In this paper, the authors explore YouTube comments by sentiment analysis and understand the meanings of various emojis and trending tags. The main aim is to connect creators, viewers, and decision-makers by knowing what an emoji and comment means. Each comment has a story, each emoji has an emotion, and each view shows a connection. It's a journey of understanding and searching for what makes digital interactions so meaningful.

2 Literature Review

2.1 YouTube-Specific Sentiment Analysis

The sentiment analysis measures how strong the emotions are, called sentiment score [1]. Such an analysis on YouTube helps in understanding how text comment signals and emojis impact user engagement and community interactions. Sentiment analysis is one of the popular fields of research in natural language processing (NLP) because it has many uses, like finding opinions, extracting emotions, and predicting trends on social media [2]. Nowadays, emojis are used by viewers in a variety of fields. This is a unique way of expressing emotions effectively in a limited space [3]. Every YouTube channel allows users to express their reactions by comments. The comments may be positive or negative showing support or disliking respectively [4]. The YouTube comments can be analyzed for polarization and categorized as positive or negative [5]. Through the opinions, as received from comments one can judge the quality of the concerned product or service, and also the areas of improvement [6].

Nowadays YouTube has turned into one of the main sources of income for creators/influencers. The earnings are based on Likes, Dislikes, Subscribers, and Views, all these aspects also makes the channel popular. The comments make a significant impact on the video, so it is crucial to dissect the comments on videos posted on YouTube to evaluate viewers' reactions. Scholars have discusses methodology to analyze comments – especially to conduct the sentiment analysis for the better understanding of the viewers' attitudes and, therefore, to enhance the strategies of broadcasting. For instance, Cunha et al. (2019) employed a deep neural network for sentiment analysis to detect comments' positivity, negativity, or neutrality on YouTube, and views' sentiment Analysis [7]. Jelodar et al. (2020) developed a content analysis model based on fuzzy logic to measure the positivity of comments on Oscar nominated movie trailers [8]. Sarakit et al employed decision trees, one of the machine learning techniques for defining anger, joy and sad in Thai media comments [9].

Such other works are the application of NLP to address difficulties and strategies in recent investigations on sentiment analysis. Islam et al. (2024) discuss these challenges, and Isac et al. (2021) employ sentiment analysis in predicting like ratios in YouTube comments taking into account the neutral feedback [10]. Ritika et al. also worked on the analysis of predictions in the comments section of YouTube with using different classifiers [11]. Since some comments are ambiguous or contain mistakes, Rhitabrat et al. proposed the classification of this kind of comments to help creators find valuable suggestions for their channel expansion [12].

3 Methodology

3.1 Data Collection

The research it was conducted using two main datasets retrieved from Kaggle. The first dataset, "Sentiment 'YouTube Comment Analysis,' has fields such as video identifier, comment, likes,and replies. It records reactions, by viewers, concerning different YouTube videos associating each comment to a specific video. This affords an insight into how the viewers engage and show some reactions. emotions are attached to texts, likes and replies. The second dataset, "Trending "YouTube Video Statistics," offers information on each video: video identification number, name,channel name, category, tags, total views, likes, dislikes, total comments, thumbnail link and date. This dataset represents the metrics of every video posted on the channel which include the view, likes, and dislikes. Including those from its tags and categories as well as its comments.

By doing so, researchers can analyze the two different aspects of the data, namely, content and frequency of videos in relation with the sentiments represented in comments provided by viewers. This combined approach enables further analysis of the relationships between video material, and how much audiences interact with it, how they convey their emotions, thus becoming the foundation for any research.

The investigation focuses on comparing positive and negative sentiment and the frequency of emojis in the comments sections of YouTube videos.

3.2 Data Preprocessing

Therefore, specifying the functions of TextBlob, a rather simple and efficient tool for text processing, sentiment analysis was carried out on the comments section of the targeted YouTube videos. TextBlob is based on the Natural Language Toolkit (NLTK) and gives access through API for easy use for simple NLP activities like tagging, finding nouns, phrases, and sentiments. In Figure 1 below, we demonstrate how the sentiment analysis is done using the TextBlob.

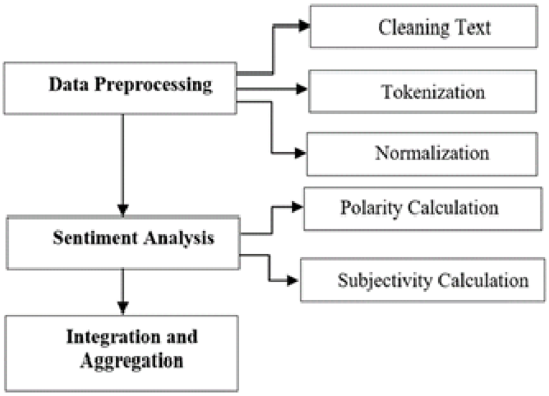

Fig.1. Sentiment Analysis Process

3.3 Emoji Analysis

Emoticons are now an important characteristic of electronic communication, which aim at conveying emotions and feelings in the form of pictures. Bot says emojis are useful in exhibiting the sentiments of the viewers in the YouTube comment section besides sentiment analysis. This section captures how emojis are processed out of the comments through frequencies and their interpretation. First, the text data are preprocessed to clear away any characters or symbols that might pose a problem during analysis of the comments. The next step involves structuring the cleaned data in a manner that will be convenient for analysis of text. The process of emoji analysis is illustrated in Figure 2 below.

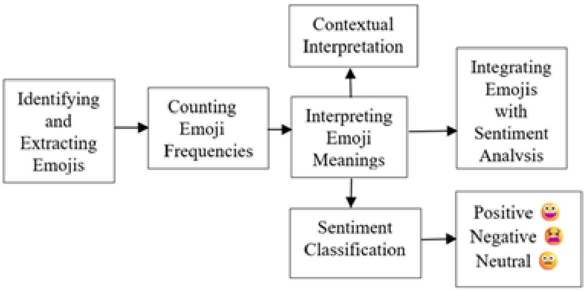

Fig.2. Emoji Analysis Steps

3.4 Tag Analysis Subtitles on YouTube serve as tags that enable the organization of related videos and their search become easier. They give the insight of the kind of issues and trends that are frequent in the forum. As it will be discussed in this research, tags from YouTube videos were used to consider trends, as well as how they affect viewers' attitude, popular themes and topics on the platform. In this research, tags from YouTube videos were extracted to study trends and see how they relate to viewer sentiment. This section illustrates how the method to extract and analyses the tags was applied (Figure. 3.), with focus on the content trend and audience engagement.

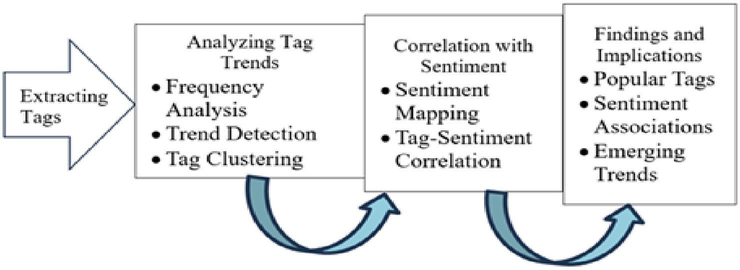

Fig.3. Tag Analysis Methodology

4 Exploratory Data Analysis (EDA)

4.1 EDA for Positive Sentences

Analyzing EDA of positive sentences of YouTube comments revealed the insights into what viewers like. Studying the positive feedbacks, the existence of what attracts the audience can be distinguished. This section explains the procedure and analysis of the EDA process in terms of positive sentences with special reference to the frequently used positive words and phrases. The word cloud visualizations of positive words for the dataset are shown in Figures 4.

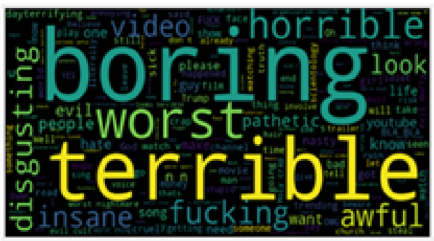

Fig.4. Word Cloud of Positive sentences **Fig.5.** Word Cloud of Negative Sentences

4.2 EDA for Negative Sentences

The EDA of negative sentences in YouTube comments uncovers significant areas where viewers feel dissatisfied or critical. This analysis helps identify aspects of content that may need improvement, highlighting common themes in viewer concerns. The visualization of common words and phrases through a word cloud provides a compelling snapshot of negative viewer sentiment, highlighting the aspects of content that drive negative engagement. The results of negative word cloud visualizations implemented on the dataset are shown below in Figure 5.

4.3 Analyzing Tags column, what are trending tags on YouTube

The tags for each video in the dataset were extracted and aggregated into a single text corpus. A word cloud was generated to visualize the most frequently used tags, highlighting the trends and themes prevalent on YouTube. Figure 6 shows the word cloud of tags analysis.

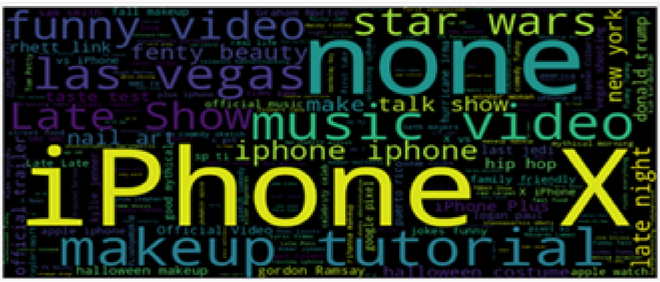

Fig.6. Word Cloud of tags column

5 Results and Discussion

5.1 Relation Among Continuous Variables

Discuss the relationships found between continuous variables and their implications for content strategy and user engagement.

i. Regression Plot for Views & Likes:

The connection between the number of views and the number of likes on YouTube videos is shown in Figure 7. From the plot, it is clear that as the number of views increases, the number of likes also tends to increase. This means that the regression line indicates a positive correlation between views and likes.

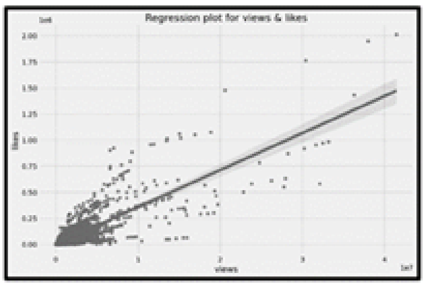

Fig.7. Regression plot for Views/Likes

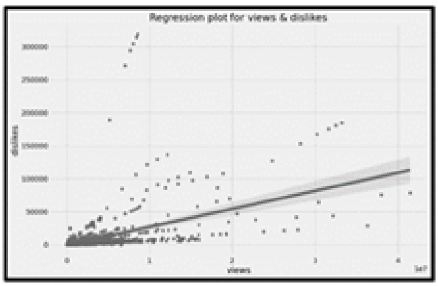

Fig.8. Regression plot for Views/Dislikes

ii. **Regression Plot for Views & Dislikes:** The Regression Plot between Views & Dislikes is shown in Figure 8. Compared to the views & likes relationship, the number of dislikes has somewhat less relation to views.

iii. **Correlation matrix:** There is a moderate positive correlation between likes and dislikes. This might suggest that videos with more likes also tend to attract more dislikes, but the relationship is less strong compared to views and likes. A correlation coefficient of 0.50 indicates a moderate relationship. The correlation matrix is shown in Table 1. Figure 9 depicts the heat map.

Table :1 Correlation matrix

	Views	Likes	Dislikes
Views	1.000000	0.832844	0.541955
Likes	0.832844	1.000000	0.497439
Dislikes	0.541955	0.497439	1.000000

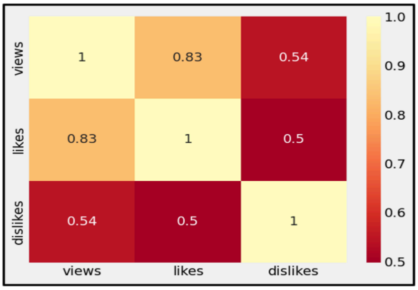

Figure: 9 – Heat Map

iv. **Role of Emojis in Comments:** The analysis of emoji usage in comments reveals significant insights into user engagement

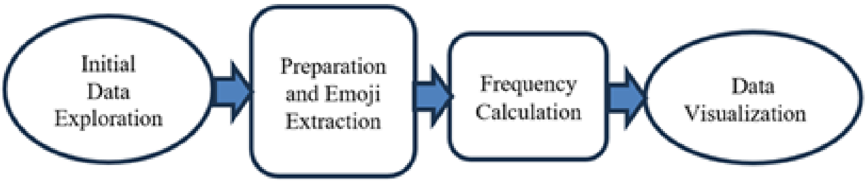

Fig.10. Emoji Analysis Steps

and sentiment. Emojis are a powerful tool for expressing emotions, and their frequency can provide valuable information about the overall tone and reception of content. The visualization effectively communicates the most popular emojis and their respective counts, highlighting the dominant expressions within the comments. Figure 10 illustrates the emoji Analysis Steps.

Results and Insights:

The total number of comments: was 718,424, and the total number of extracted emojis: was 333,278. Figure 11, highlights the distribution of the top 20 emojis, with 😂 (laughing face) being the most frequent, followed by ❤ (heart) and 😍 (heart eyes). Emojis like 😂, ❤, and 😍 indicate high engagement and emotional responses from users.

The prevalence of positive emojis (hearts, laughing faces, and smiles) suggests that comments are generally positive or express strong emotions. The presence of different emojis (e.g., 👏, 🔥, 👍) shows a variety of reactions, from appreciation to excitement. Analyze the frequency and types of emojis used, correlating them with sentiment and engagement metrics.

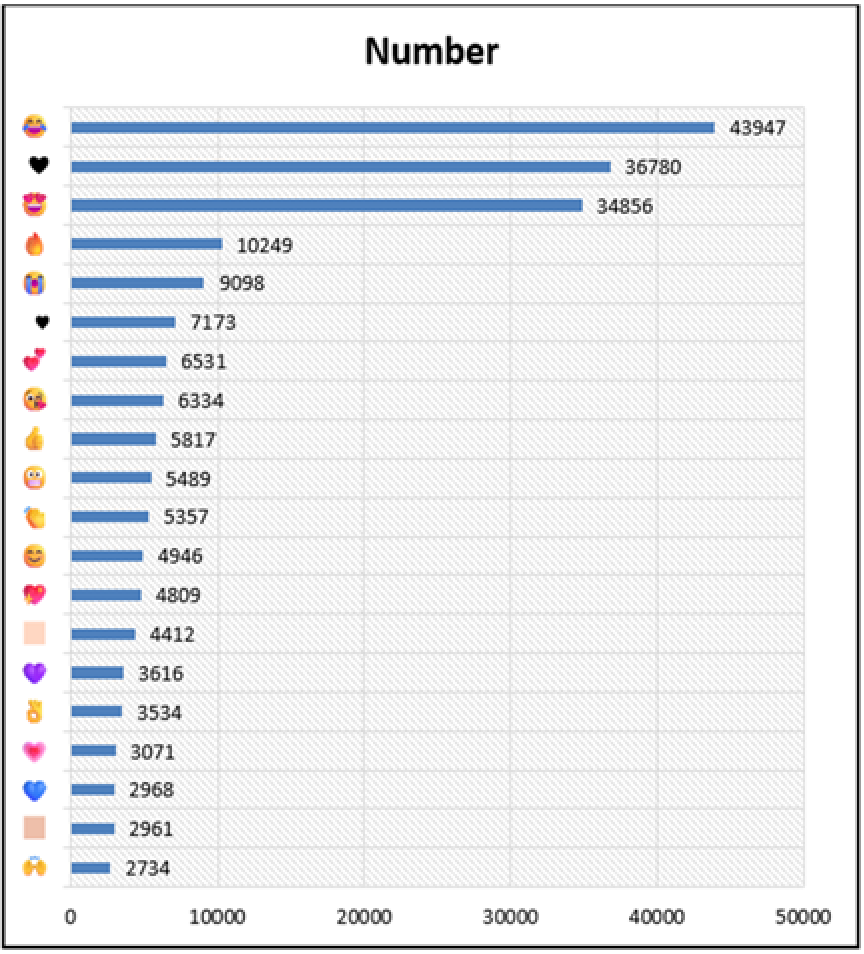

Fig.11.Top 20 Most frequent Emojis in Comments

6 Conclusion

The analysis of YouTube comments and video metrics provided several key insights. A regression analysis between views and likes revealed a strong positive correlation, indicating that videos with higher views tend to receive more likes, reflecting user engagement. The correlation matrix further supported this, showing a high positive correlation between views and likes (0.832844), and moderate positive correlations between views and dislikes (0.541955) as well as likes and dislikes (0.497439). Additionally, an extensive emoji analysis of 718,424 comments identified that emojis

are commonly used to express emotions, with a total of 333,278 emojis extracted. The most commonly used emojis 😂, ❤, and 😍 indicate that the general sentiment in the comments is positive, with viewers frequently showing joy and affection.
Cognitively, the results reaffirm the importance of emotional appeal in web-based interactions and provide directives for how content providers can foster enhanced levels of interaction with their audiences through their video content.

Future Work

The future studies can enrich these findings by investigating the emoji usage, which reflects various emotional responses that particular content usually elicits. Analyzing the time of using emojis could also complement the information about online viewers' engagement chronology. It is possible that deeper analysis of the referenced content based on the sentiment analysis approach might assist in identifying a more subtle emotional profile of the commenting audience. Examining the timing of emoji usage could also reveal patterns in viewer engagement over time. A more detailed sentiment analysis might help uncover the subtle emotional states of commenters. It expands the list of videos and comments in multiple languages that need to be put into the selection that would enhance the applicability of the findings. Second, it can be expanded to the inclusion of data from other social networking sites, which can enable between-site comparison based on concerns such as use activity and level of emotional display.

7 References

1. Islam, M.S., Kabir, M.N., Ghani, N.A. et al. "Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach". Journal: Artificial Intelligence Review, Volume: 57, Article Number: 62, Publication, March 2024, DOI: 10.1007/s10462-023-10651-9
2. Nafis Irtiza Tripto, Mohammed Eunus Ali, 'Detecting Multilabel Sentiment and Emotions from Bangla YouTube Comments', International Conference On Bangla Speech And Language Processing(Icbslp), 21-22 September, 2018.
3. Bai Q, Dan Q, Mu Z and Yang M(2019) A Systematic Review of Emoji: Current Research and Future Perspectives, Frontiers in Psychology, vol. 10, Article Number – 2221, October 2019.
4. K.M.Kavitha, Asha Shetty, Bryan Abreo, Adline D'Souza, Akarsha Kondana, 'Analysis and Classification of User Comments on YouTube Videos', International Workshop on Artificial Intelligence for Natural Language Processing (IA&NLP2020), Science Direct Procedia Computer Science 177 (2020) 593–598.
5. Wang L, Meng Z (2022) Multichannel two-dimensional convolutional neural network based on interactive features and group strategy for Chinese sentiment analysis, Journal - Sensors, Volume: 22, Issue: 3, Page: 714, January 2022
6. Pavitha N, Pungliya V, Raut A, Bhonsle R, Purohit A, Patel A, Shashidhar R (2022) Movie recommendation and sentiment analysis using machine learning. Global Trans Proc 3(1):279–284. Journal: Global Transitions Proceedings, Volume: 3, Issue: 1, Pages: 279–284, 2022

7. Alexandre Ashade Lassance Cunha(B) ,Melissa Carvalho Costa, and Marco Aur´ elio C. Pacheco, 'Sentiment Analysis of YouTube Video Comments Using Deep Neural Networks', Springer Nature Switzerland AG 2019 L. Rutkowski et al. (Eds.): ICAISC 2019, LNAI 11508, pp. 561–570, 2019, DOI: 10.1007/978-3-030-20912-4_51
8. Hamed Jelodar, Yongli Wang, Mahdi Rabbani, Sajjad Bagheri Baba Ahmadi Lynda Boukela Ruxin Zhao, Raja Sohail Ahmed Larik, 'A NLP frame work based on meaning ful latent-topic detection and sentiment analysis via fuzzy lattice reasoning on you tube comments', Journal: Multimedia Tools and Applications, Volume 80, Early online date: 28 Sept 2020, published by Springer Science+Business Media, LLC, part of Springer Nature 2020, https://doi.org/10.1007/s11042-020-09755-z,
9. Phakhawat Sarakit, Thanaruk Theeramunkong, Choochart Haruechaiyasak, Manabu Okumura, 'Classifying Emotion in Thai YouTube Comments', 6th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES), 2015, ISBN: 978-1-4799-8565-4/15, IEEE.
10. Isac Lorentz & Gurjiwan Singh, 'Sentiment Analysis on Youtube Comments to Predict Youtube Video Like Proportions', Degree Project Technology, First Cycle, 15 Credits, Stockholm Sweden, 2021.
11. Ritika Singh, Ayushka Tiwari, 'Youtube Comments Sentiment Analysis', International Journal of Scientific Research in Engineering and Management (IJSREM) Volume: 05 Issue: 05, May - 2021 ISSN: 2582-3930.
12. Rhitabrat Pokharel and Dixit Bhatta. 2021. Classifying YouTube Comments Based on Sentiment and Type of Sentence. Publication: arXiv publication. Portland, OR, USA.