European Journal of Clinical Pharmacy

Print ISSN: 2385-409X **Online ISSN**: 3105-0409
Website: https://farmclin.com/

Original Research Article

Sentiment Analysis for YouTube Comments and Video using AI

Article History:

Name of Author:
[1]Kamjula Lokesh Reddy, [2]Kaliseti Dhanush, [3]Appanaboina Charan Teja, [4]Bhogyam Manohar and [5]Ms.C. Priyatharsini (Ass.t professor)

Affiliation: *[1,2,3,4,5]CSE, Mahendra Engineering College.*

Corresponding Author:
Kamjula Lokesh Reddy

Received: 15-10-2025
Revised: 30-10-2025
Accepted: 18-11-2025
Published: 30-12-2025

Abstract: The determination of the feeling of people towards content produced by users on video sharing sites has become a significant instrument in understanding the feeling of many people. With the help of state-of-the-art NLP and DL technologies, a figuring out mood of YouTube comments and video transcripts is not demonstrated using an automated system. Textual data is collected with the help of the YouTube Data API and YouTube Transcript API, and thus, a deep sentiment analysis of a large variety of different material can be conducted. In order to obtain improved feature representation, the text obtained is systematically processed and involves normalization, lemmatization, tokenization, and the removal of stop words. Context aware transformer based and recurrent neural network models are applied in estimating sentiments. BERT is employed in sentiment classification at the comment-level, and RoBERTa, Long Short-Term Memory, and Gated Recurrent Unit networks are used to analyze video transcripts. In order to ensure that the analysis is reliable, model performance is assessed according to typical tools of evaluation, such as precision, recall, accuracy, and F1-score. According to available experimental evidence, the DL models perform significantly better in comparison with the baseline methods. RoBERTa was most accurate with an accuracy of 99% and LSTM and GRU models were second with above 97 percent. The studied feelings are displayed in an interactive web interface thus it is less difficult to comprehend their sense. The system allows you to access real-time and scalable sentiment information to make data-driven decisions on analytics of digital media and audience engagement metrics.

Keywords: Sentiment analysis, YouTube comments, video transcript analysis, natural language processing, deep learning, BERT, RoBERTa, recurrent neural networks.

INTRODUCTION

It has become much easier to figure out what people think since sentiment analysis of online content is possible, particularly on websites such as YouTube where individuals post their opinions and emotions through comments and interaction with the videos. The quantity and diversity of user-generated content complicate the process of sentiment evaluation by humans, rendering the process slow and sometimes even impossible. YouTube is increasingly gaining popularity in providing information, entertainment and advertising. Due to this fact we require the automated systems that are able to measure the correct emotions towards the content of big datasets. NLP and DL strategies provide a solid foundation towards addressing such issues. They allow us to extract useful information on text on the internet and consider the minor variations of meaning and sequential dependencies inherent in human language. Automated mood analysis systems classify written data as positive, negative, or neutral using complex models such as BERT, RoBERTa, LSTM and GRU. The design of BERT is effective in terms of comprehension of text context due to its transformer nature. This suits it well in the comment-level sentiment analysis [3, 4]. RoBERTa does it even better as it enhances training approaches and makes the model more sustainable in large-scale text classification tasks [2]. It is good to apply sequential models such as LSTM and GRU to video transcripts since they can determine how the meaning of speech-derived text evolves with time and depending on the context, and as a result, it is possible to determine the mood of words or phrases in sequence [1], [7].

The APIs of YouTube can be employed to gather data on mood analysis with ease and in a short time. YouTube Data API allows retrieving user comments, whereas the YouTube Transcript API allows retrieving what people say in videos. Stopword removal, normalization, tokenization and lemmatization are some preprocessing techniques that improve text input and enable better performance of DL models [5], [6]. With these models placed within an online platform, it is possible to watch the way people feel about something in real time using basic tools of graphics such as pie charts and bar graphs. This assists the stakeholders in getting the data in a fast manner and assists them to make decisions.

The method is aimed at assisting content creators in several ways, including assisting marketers in determining what people think and researchers in examining how feelings alter over time regarding various subjects. The overall aim is to make it automatic to determine the emotions of the comments and video transcripts. This ought to be in a manner that is right, scalable and capable of operating in real time via an interface that is user friendly. This will enhance decision making process and means of involving the audience [8], [9], [10].

RELATED WORK

Sentiment analysis of YouTube comments has received a lot of research due to the sheer number of users on the site and the daily number of videos, songs, and other content that is uploaded there. The last research has been dedicated to applying sophisticated AI and DL techniques to sort the emotions of comments with high speed. The idea is to provide valuable information to content makers, marketers and researchers. A complete account of the various ways of mood analysis that can be applied to YouTube comments was provided by Meghana [11]. She demonstrated the application of the combination of old NLP methods and novel DL models to assist in accuracy and context deciphering. The research emphasizes that such preparations as tokenization, lemmatization, and stopword removal are critical to achieving improved outcomes with those models that are trained on user-written text.

According to Rathod et al. [12], a system named Comment Compass was proposed that scans YouTube comments with the help of web scraping, sentiment analysis and generative AI. This approach demonstrates that it is possible to extract something practical out of large collections of comments and provide something practical to individuals with a vested interest in the results. The system facilitates easier comprehension of the changing trends in opinions using generative AI models and assists marketers and content creators make content marketing and optimization decisions. Similarly, Pavithra et al. [13] examined the application of LSTM models to determine the emotions of people towards YouTube remarks. Their experiment demonstrated that the model was excellent in comprehending contextual disparity in informal language and managing serial reliance. It was also very successful in detecting positive, negative and neutral affects.

Bindhumol et al. [14] investigated comments on YouTube with the help of sentiment analysis that is based on DL. This demonstrated that neural networks designs can be used with large datasets. The experiment revealed that both the LSTM and GRU models outperform the conventional ML models in terms of classifying the mood particularly when the text sequences are of varied lengths and difficulty. These findings were supported by Bharathvaj and Mageshkumar [15] who employed LSTM-based models to estimate the mood of YouTube comments and demonstrated that deep recurrent networks are stable and can be used in real-time to analyze the material.

Mironela et al. [16] examined NLP-related techniques of analyzing comments on YouTube videos, and how grammatical and meaning-based factors can be combined to assist in, locating the overall mood of a comment. Their experiment demonstrated that classification is significantly improved when lexical resources are combined with environmental embeddings as compared to when rules are used. Liyih et al. [17] applied DL models to examine the evolution of feelings in a highly sensitive YouTube video on different subjects such as the Hamas-Israel conflict. This demonstrates the numerous places and situations of high stakes where sentiment analysis can be applied.

Chinnasamy et al. [18] developed a sentiment analyzer, which was based on NLP to analyze comments at the comment level; it is easy to automate evaluating sentiment through a combination of preprocessing, feature extraction, and deep learning classification. A sentiment analysis performed on transcripts of YouTube videos by Shabrina et al. [19] demonstrated that DL techniques can be applied to effectively determine how individuals perceive video content. This complements analysis by comment to provide a complete picture of responses by people. It can get hectic due to spam and numerous irrelevant content in YouTube comments as Xiao and Liang [20] described. ML techniques helped them to eliminate the noise thereby improving the accuracy and reliability of sentiment classification systems.

MATERIALS AND METHODS

The proposed system will desire to enhance the analysis of YouTube content by incorporating the LSTM, GRU, and RoBERTa models to analyze video transcripts and enhance the already existing BERT model to better classify the mood of a comment. This combination of the two techniques relies on the advantages of sequential neural networks to comprehend the impact of context on meaning in video transcripts. Meanwhile, models based on the use of a transformer ensure that the comment text is clearly perceived. The system is designed to obtain information related to comments based on the YouTube Data API and video information based on the YouTube Transcript API. Then, the preprocessing, such as tokenization, lemmatization, and stopword removal, are performed to ensure that the model performs better. By combining these models, the system provides wholesome sentiment analysis on both comments and video content. The system also has a simplified web interface that can be used to view sentiment distributions in real time. This provides content creators and advertisers with valuable information to be implemented. The approach enhances the scalability and accuracy of existing solutions [21], [22], [23].

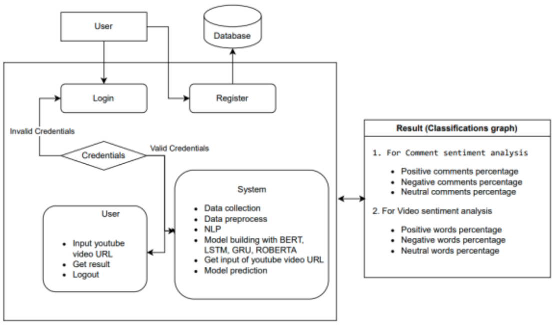

Fig.1 Proposed Architecture

Fig. 1 is a system design of YouTube sentiment analysis. After a system has validated a person, he provides the system with a video URL. This engine is built on such models as BERT and LSTM to retrieve data collection, preprocessing, and NLP. The result will give the Classifications graph indicating the proportion of the positive, negative, and neutral responses to the movie and comments.

i) Dataset Collection:

The information in the system consists of the YouTube comments and the transcripts of the videos. This provides a general perspective of the feeling of the people towards the videos. They can be pulled in with the YouTube Data API (enabling users to respond automatically) and with pre-collected datasets, such as the GBcomments of Kaggle. YouTube Transcript API obtains video texts by eliminating subtitles in videos. This allows the sentiment analysis of the video content to be done. Stopword removal, lemmatization, and tokenization are performed to these datasets before they are used to enhance the quality of text and model performance. By joining comments and transcripts, you can have a complete view of the way the audience is feeling about both written and spoken data [24].

ii) Pre-processing:

Mood analysis involves data pre-processing which ensures that the YouTube comments and video transcripts are clean, consistent and ML and DL models are useful. The initial one is the tokenization. In this step the text is divided into small segments such as words or phrases which are valuable. The idea of tokenization facilitates the process of locating the special portions of the text, and models are able to capture meaning relationships amid tokens in an excellent manner. The text is cleaned up after the process of tokenization to dispose of such items as special characters, URLs, emojis, HTML tags, and additional spaces. The step ensures that the model is able to learn without being distracted by needless symbols. It also improves the data in general.

Lemmatization comes next. This process transforms words into their root forms or base. As an example, a running will be run and better will be good. This allows models to consider various kinds of words to be a single meaning. This reduces the vocabulary size and increases the capability of being able to comprehend the circumstance. Removal of stopwords is also performed and unwanted words such as is, the, and and are removed since they do not contribute much to the sentiment analysis task. The exclusion of these words is a method of concentrating on the words that are emotive and that is why the model is more precise.

GRU, LSTM and transformer based models require the input length to be of equal length, so truncation and padding ensure that all sequences are of equal length. Special tokens are inserted in short sequences and long sequences are reduced to a predetermined maximum length. This action ensures that the inputs are all of the same size and this enables quick processing of a batch and the performance of the model remains consistent. These preprocessing steps applied to the textual information in the YouTube comments and video transcripts put it in a structure, and it is now prepared to be successfully classified into moods. This forms a good foundation to further analysis and visualization.

iii) Train & Test:

The training part of the preprocessed dataset is what we use to train our sentiment analysis models and the testing part of the dataset is what we use to evaluate the effectiveness of our sentiment analysis on new data. The training set enables the models to learn patterns of data to be able to determine contextual relationships, sentiment hints, and sequential relationships in YouTube comments and video transcripts. The testing set is also an unknown data, which is applied to determine the validity of the trained models in regard to their ability to adapt to new, real-world data. This division ensures that the models do not fit the training data too well and can rely on adequately predicting new material. Ensuring that the model is efficient and that the sentiment classification is robust, it should be noted that the data should be divided properly and that there should be care when working with the two groups: the training and the testing one.

iv) Algorithms:

BERT: BERT or Bidirectional Encoder Representations of Transformers is a transformer model that attempts to learn the relationships of words within a sentence by reading the text in both directions. It traces the hard relationships among words in order to comprehend all aspects of a language, meaning and sentence structure. BERT works excellent with NLP problems such as sentiment analysis, question answering, and named object recognition as it forms contextual embeddings on the each token of the input text. It is reversible allowing it to examine words before and after the target word, enhancing the correct meaning interpretation. In sentiment analysis, BERT is effective at determining whether a comment or transcript is positive, negative or neutral since it is capable of isolating minor clues to mood and accommodates variations in casual language.

GRU: GRU, also known as Gated Recurrent Unit, is a form of RNN, which is well positioned to handle rapidly changing data, which comes in a sequence. It addresses the vanishing gradient issue of the normal RNNs with gating mechanisms such as reset and update gates regulating the flow of information over time steps. GRU is exceptionally appropriate in areas that have temporal constraints, such as text or speech processing and sentiment classification. GRU identifies patterns in sequences and makes accurate guesses of text sequences of varying length through retention of useful information and discarding the rest. It is effective in determining how individuals perceive video clips and other sequential information in which the context and sequence of words play a significant influence on the way feeling can be perceived.

LSTM: Long Short-Term Memory or LSTM is a special form of recurring neural network, which learns long-term dependencies in a better manner as compared to normal RNNs. It stores information in a sequence-directed manner using a sequence of gating techniques (input, forget and output gates) and memory cells to store or forget information within memory. Modeling of sequences is often applied to LSTM models; it involves activities such as determining what one is talking about, what is coming next in a sentence, and reading a text to understand how definite or negative the mood is. They are excellent at establishing time and context relationships between text sequences and therefore it becomes possible to sort the mood of both comments and video transcripts accurately. With LSTMs, it is easy to maintain long-term dependencies; therefore, LSTMs can work very well with long sequences of text in which vital hints to the meaning are not concentrated in a few lines or words.

RoBERTa: RoBERTa or robustly optimized BERT pretraining approach is an enhanced variant of BERT that applies pretraining strategies such as larger batching, prolonged training, and dynamically covering tokens as a way of enhancing performance. Like BERT, it is a transformer-based design that builds contextual representations that demonstrate the relationship between words both in meaning and grammar. RoBERTa is quite effective in natural language understanding tasks such as sentiment detection of a text, text classification, and question answering. It is more effective than BERT on various types of more complex textual data, hence it can be used to analyze comments and video transcripts. RoBERTa is able to learn tiny indicators of emotion and classify them to a reasonable degree of positive or negative or neutral with a high degree of accuracy.

RESULTS AND DISCUSSIONS

Accuracy: The test accuracy is determined by the ability of the test to distinguish between the sick and healthy. In order to have a clue about the test accuracy we should determine the percentage of true positives and true negative among all cases that were tested. In math, this is written as

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \quad (1)$$

Table.1 Performance Evaluation

Model	**Accuracy**
BERT	74 %
GRU	97.78 %
LSTM	97.98 %

ROBERTA	99%

Table desc

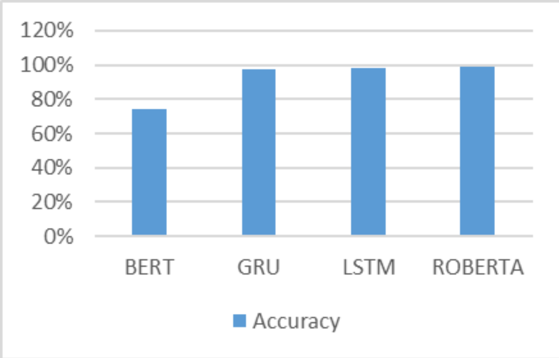

Fig.2 Comparison Graph

Graph desc

Fig.3 Home Page

Fig.4 About Page

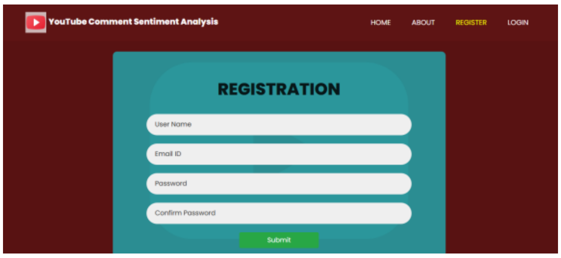

Fig.5 Registration Page

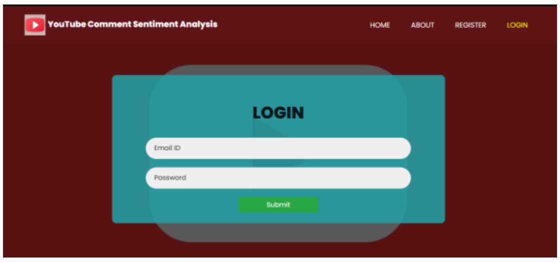

Fig.6 Login Page

Fig.7 Main Page

Fig.8 Enter URL

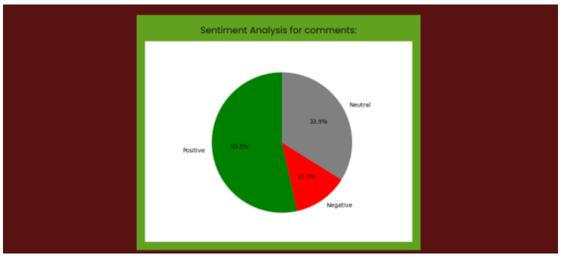

Fig.9 Predict Result

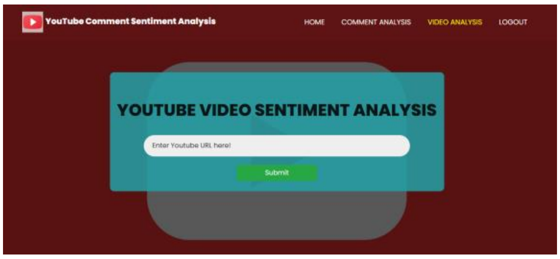

Fig.10 Enter YouTube URL

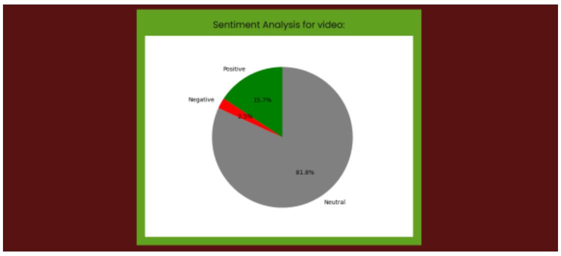

Fig.11 Final Outcome

CONCLUSION

YouTube Comments and Videos Sentiment Analysis method automatically and correctly determines the feeling of people towards comments and video transcripts. It provides valuable data to the content creators, marketers, and researchers. Advanced natural language processing algorithms and DL architectures enable the system to handle large volumes of text data, which ordinarily should be handled by a human being. The high-performance models like RoBERTa could successfully and frequently guess the way people were feeling accurate to 99 percent. In the case of sequential analysis of transcripts, LSTM and GRU models also performed well as the accuracy rates are 97.98% and 97.78, respectively. These percentage rates also guaranteed that the mood interpretation of video contents was high. With the help of such tools as pie charts and bar graphs, one can grasp the way people feel about something easier. The modular design enables the possibility of adding features, real-time analysis of the data, and changes made to apply to several languages and cross-platform interaction. All in all, the system demonstrates a reliable, fast, and very precise method of sentiment analysis. This assists the stakeholders in making decisions that are informed by data and engage with their audiences at the digital content landscape better.

It is expected that someday the YouTube Comments and Videos Sentiment Analysis system will be revised to permit multilingual sentiment analysis using models such as mBERT or XLM-R. This will render it applicable globally. Having access to Twitter, Instagram, Facebook, and other social media platforms, one will be able to study content created by users more thoroughly. It will also be helpful to add multimodal sentiment analysis which will search audio, facial expressions, and movements on videos to understand how the audience feels. These are aimed at ensuring that content makers, marketers as well as researchers can scale up, obtain more precise results, and obtain more comprehensive results.

REFERENCES

1. Athindran, N. S., S. Manikandaraj, and R. Kamaleshwar. "Sentiment Analysis from YouTube Video Using Bi-LSTM-GRU Classification." *SpringerLink*, 2018.
2. Tan, K. L., C. P. Lee, and K. M. Lim. "RoBERTa-GRU: A Hybrid Deep Learning Model for Enhanced Sentiment Analysis." *Applied Sciences*, vol. 13, 2023, p. 3915.
3. Murfi, H., et al. "BERT-Based Combination of Convolutional and Recurrent Neural Network for Indonesian Sentiment Analysis." *arXiv Preprint*, arXiv:2211.05273, 2022.
4. Khan, A. H., U. Qamar, and S. Bashir. "Integrated BERT Embeddings, BiLSTM-BiGRU, and 1-D CNN Model for Sentiment Classification Analysis of Movie Reviews." *Multimedia Tools and Applications*, vol. 80, 2019, pp. 11443–11458.
5. Liu, B., G. Cao, and J. Yin. "Bi-Level Attention Model for Sentiment Analysis of Short Texts." *IEEE Access*, vol. 7, 2019, pp. 13–22.
6. Li, W., et al. "The Automatic Text Classification Method Based on BERT and Feature Union." *IEEE Parallel and Distributed Systems*, vol. 25, 2019, pp. 774–777.
7. Minaee, S., E. Azimi, and A. A. Abdolrashidi. "Deep-Sentiment: Sentiment Analysis Using an Ensemble of CNN and Bi-LSTM Models." *arXiv Preprint*, arXiv:1904.04206, 2019.
8. Le, Q., and T. Mikolov. "Distributed Representations of Sentences and Documents." *Proceedings of the International Conference on Machine Learning*, 2014, pp. 1188–1196.
9. Liu, W., and E. Cambria. "Targeted Aspect-Based Sentiment Analysis via Embedding Commonsense Knowledge into an Attentive LSTM." *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.
10. Talaat, A. S. "Sentiment Analysis Classification System Using Hybrid BERT Models." *Journal of Big Data*, vol. 10, 2023, p. 110.
11. Meghana, K. "Artificial Intelligence and Sentiment Analysis in YouTube Comments: A Comprehensive Overview." *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, IEEE, 2024, pp. 1565–1572.
12. Rathod, M., et al. "Comment Compass: An Approach to Analyze YouTube Comments through Web Scraping, Sentiment Analysis, and Generative AI for Actionable Insights." *2024 First International Conference on Pioneering Developments in Computer Science & Digital Technologies (IC2SDT)*, IEEE, 2024, pp. 1–5.
13. Pavithra, D., et al. "YouTube Comment Analysis Using LSTM Model." *Generative Artificial Intelligence: Concepts and Applications*, 2025, pp. 265–281.
14. Bindhumol, M., T. Singh, and P. Patra. "Sentiment Analysis Using YouTube Comments." *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, 2024, pp. 1–7.
15. Bharathvaj, T., and M. Mageshkumar. "YouTube Comment Analysis Using LSTM Model." *Generative Artificial Intelligence: Concepts and Applications*, 2025, p. 265.
16. Mironela, P., et al. "Analysis of YouTube Video Comments with NLP Methods." *2024 16th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, IEEE, 2024, pp. 1–6.
17. Liyih, A., et al. "Sentiment Analysis of the Hamas-Israel War on YouTube Comments Using Deep Learning." *Scientific Reports*, vol. 14, no. 1, 2024, p. 13647.
18. Chinnasamy, P., et al. "Comment Analyzer by Sentimental Analysis through Natural Language Processing." *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, IEEE, 2024, pp. 1123–1128.
19. Shabrina, U. I., et al. "Sentiment Analysis of Presidential Candidate Debates from YouTube Videos." *2024 IEEE International Conference on Artificial Intelligence and Mechatronics Systems (AIMS)*, IEEE, 2024, pp. 1–6.
20. Xiao, A. S., and Q. Liang. "Spam Detection for YouTube Video Comments Using Machine Learning Approaches." *Machine Learning with Applications*, vol. 16, 2024, p. 100550.
21. Jamil, M., H. Hadiyanto, and R. Sanjaya. "Sentiment Analysis: Classifying Public Comments on YouTube in Disaster Management

Simulation in Indonesia Using Naïve Bayes and Support Vector Machine." *Ingénierie des Systèmes d'Information*, vol. 29, no. 2, 2024, p. 437.
22. Raza, A., et al. "An Improved Deep Convolutional Neural Network-Based YouTube Video Classification Using Textual Features." *Heliyon*, vol. 10, no. 16, 2024.
23. Möller, A. M., S. A. Vermeer, and S. E. Baumgartner. "Cutting through the Comment Chaos: A Supervised Machine Learning Approach to Identifying Relevant YouTube Comments." *Social Science Computer Review*, vol. 42, no. 1, 2024, pp. 162–185.
24. Gurung, M. I., et al. "Decoding YouTube's Recommendation System: A Comparative Study of Metadata and GPT-4 Extracted Narratives." *Companion Proceedings of the ACM Web Conference 2024*, 2024, pp. 1468–1472.
25. Singh, A., D. Sharma, and V. K. Singh. "Misogynistic Attitude Detection in YouTube Comments and Replies: A High-Quality Dataset and Algorithmic Models." *Computer Speech & Language*, vol. 89, 2025, p. 101682.

