# YouTube Comment Sentiment Analysis

P NandieswarReddy[1]
*Dept of Computer Science and Engineering,*
*Amrita School of Computing,*
*Amrita Vishwa Vidyapeetham*
Bengaluru, India
BL.EN.U4AIE20046@bl.students.amrita.edu

Rithvika Alapati[1]
*Dept of Computer Science and Engineering*
*Amrita School of Computing,*
*Amrita Vishwa Vidyapeetham*
Bengaluru, India
BL.EN.U4AIE20054@bl.students.amrita.edu

Sai Aswath S[1]
*Dept of Computer Science and Engineering ,*
*Amrita School of Computing,*
*Amrita Vishwa Vidyapeetham*
Bengaluru, India
BL.EN.U4AIE20056@bl.students.amrita.edu

Dr. Beena B.M[2]
*Department of Computer Science and Engineering,*
*Amrita School of Computing*, Bengaluru,
*Amrita Vishwa Vidyapeetham*, India

***Abstract*—Sentimental analysis is critical in understanding the user's reaction toward the content on social media platforms. YouTube is one of the most used social media platforms in the current era. Understanding the user's reaction towards the content posted on such platforms is important in improving the content. A sentimental model using LSTM and NLP techniques is built and trained using the IMDB dataset and deployed using Amazon web services (AWS). 85% accuracy persisted and detailed the model's performance in categorizing comments as positive and negative. The interactive dashboard is built using stream-lit.**

**Keywords – AWS, Sentimental Analysis, LSTM, NLP, Hyperparameters.**

## I. INTRODUCTION

In today's virtual surroundings, user-generated content material on systems together with YouTube has turned out to be a quintessential part of online communique. The number of comments and the kind of remarks consisting of video reflect customers' rich feelings. Emotion evaluation, a developing practice in herbal language processing, affords a method of decoding the underlying emotional tones that underlie those troubles. This application is pushed using the want to apply sentiment evaluation to YouTube content, unpacking the emotions expressed by customers and supplying actionable insights to content material creators and platform managers. The main moto for stepping into YouTube comment sentiment analysis lies in its capability to transform content strategy, community engagement, and platform dynamics, and a video posted on YouTube will have millions of comments and it is very hard for the creator to go through all of them and understand the user requirements. Exploring audience emotion in their content material can help content creators customize their content to resonate more with the target content, leading to elevated engagement and viewership. Also, it enables corporations by way of read purchaser sentiment towards their products or services and may offer precious comments for improving product development, advertising techniques, and customer service. Research scholars can examine public sentiment on numerous subjects and may contribute to investigations in areas that consist of psychology, politics, and advertising and marketing. As motion pictures acquire thoughts and memories, information and the emotions expressed in that feedback are essential for content creators trying to align their content material with target audience options, and platform managers trying to personal experience may be superior. As a part of cloud computing, Amazon Web Services (AWS) offers multiple responses that combine nicely with our sentiment assessment framework in YouTube contexts. It translates into advanced overall performance, scalability, and global get-right of access. The versatility of AWS is in particular obvious in its scalability competencies, which is a high problem given the dynamic nature of YouTube content. With the capability to dynamically scale assets primarily based actually on calls, AWS ensures that our sentiment assessment model remains responsive, even at some point of durations of increased times whilst clients are connected. Efficiency in version learning is a cornerstone to carrying out our task, and AWS allows it fairly in this aspect. With parallel processing abilities AWS hurries up the training of our sentiment assessment model, which is a specially complicated venture dealing with large facts units which include IMDb now this parallelization no longer reduces schooling issues but supports iterative refinement technique, if we suppose it is suitable and tremendous Additionally, AWS' managed offerings, mixed with Sage Maker, play a key feature in streamlining our tool studying workflow. By abstracting the complexity of infrastructure management, AWS shall us see the evolution and optimization of our sentiment analysis version. Scalability and price-effectiveness allow us to optimize using AWS content material, and adapt sensitivity evaluation obligations to the right computational goals. Given YouTube's global goal market, global reach is paramount. AWS's international community of listing services, blended with Content Delivery Network services sures the reach of our sentiment analysis software program. This global strategy contributes to continuing consumer liberty, irrespective of the geographical place of customers interacting with our application. Additionally, AWS offers dependable and robust surroundings for our sentiment analysis software program. The chosen approach involves training a sentiment analysis version using the IMDb dataset, which is a complete film analysis repository. Using the abilities of Amazon Web Services (AWS), the model is skilled in using sentiment analysis concepts from movie reviews to diverse kinds of content determined in YouTube content material This fact's structure and platform choice guarantees a foundation of tough for the version, growing its flexibility and performance.

## II. Literature Survey

[1] Analysing viewers' sentiments toward YouTube films is a critical manner to measure video enjoyment, in step with such. This paper makes use of emotional scores to expect chance by inspecting the emotional tone of the response. The YouTube API is used to retrieve comments, which are then pre-processed to get rid of inappropriate facts and standardize the captions. Comments are then categorized as typical, terrible, or unbiased, and sensitivity is analyzed through the use of language processing gear inclusive of TextBlob or NLTK The equal type data have been derived from 100 tremendous comments distribution as compared to the entire. This approach offers treasured insights for content material creators and marketers, supporting the sentiment of the target audience and consequently offering content material is powerful.

[2] Sentiment evaluation includes classifying text into superb, terrible, or impartial sentiment. YouTube is a common platform for expressing thoughts on different subjects. The huge improvement in Bali province for the duration of the Jokowi era, consisting of the development of the Mengwi-Gilimanuk dual carriageway, provoked quite a few network responses due to the extensive use of agricultural land. The Naïve Bayes algorithm is generally used, and analytical metrics inclusive of accuracy, precision, and recollect are used. Analysis of 18 YouTube video links with 701 feedback of 50.64% positive emotions, 7.70% bad emotions, and 39.23% impartial feelings.

[3] This journal examines and analyzes YouTube content generated via the Kompas TV station that's investigating public opinion on capability candidates in the Indonesian presidential election in 2024. The statistics series uses Koberi to download content material, that's then analyzed In Python using the Pandas library, matplotlib, word cloud, and textblob The number one goal is to have a look at public reactions to capability presidential applicants inside the well-known election. The findings display that the general public has fine sentiments in the direction of 3 wonderful applicants: Ganjar Pranowo, Anis Baswedan, and Prabowo Subianto.

[4] NLP techniques are proposed in this study to categorize Arabic comments as either positive or negative. With 4212 tagged comments, the model is trained and achieved a Kappa score of 0.818. The comparison of six classifiers: Naïve Bayes, SVM, Random Forest, KNN, Decision Tree, and Logistic Regression is done. Naïve Bayes achieves a good 94.62% accuracy and 91.46% MCC score, and the corresponding precision, recall, and F1-measure are 94.64%, 94.64%, and 94.62%. This study offers insightful information to content creators who want to improve viewership and their content and audience engagement.

[5] This evaluates the accuracy of emotion category with the use of Random Forest and Word2Vec Skip-gram function extraction. Word2Vec proves effective in representational translation, increasing emotional type accuracy. This facts set includes 31,947 remarks received from a YouTube channel viewing the 2019 presidential debate, with 23,612 advantageous feedback and 8,335 terrible feedback Oversampling is used to stabilize high quality and terrible facts to lessen bias. Skip-gram is used to extract components of phrases, imparting functions across the content of each phrase. The sensitivity type model is built the use of a random forest, in which distinctive periods and window dimensions are used for evaluation. Cross-validation is used to evaluate the performance of each model. Experiments with exclusive ages yield sampling accuracies ranging from 90.1% to 91%. However, the check effects display an accuracy starting from 88.77% to 89.05%. Although the accuracy of the version decreases barely, the difference isn't statistically substantial. In subsequent experiments, it's miles advocated to use more than twenty durations and larger window sizes, respectively, to look a significant increase in accuracy The number of durations and the larger window size in Skip Village considerably have an effect on accuracy, and emphasize the significance of cautious consideration and use of those parameters of representation

[6] Because of its functions, Amazon Web Services (AWS) is becoming increasingly popular among man or woman customers and massive corporations. This observation investigates the application of sentiment analysis inside the AWS Elastic Compute Cloud (EC2) the use of Twitter records. In the case of EC2, information management is facilitated by using elastic load balancing. The amassed records are preprocessed to enhance purity, and then a device gaining knowledge of the-based logistic regression model is hired to distinguish first-rate and terrible sensory lessons. In terms of overall performance, the advised version obtains an extremely good accuracy of 94.17%, outperforming the existing techniques. This examines a way to contribute to the expertise of sentiment analysis on AWS and shows the effectiveness of advanced gadget analyzing strategies.

## III. METHODOLOGY

The methodology in the development of the sentiment analysis or reaction analysis model for YouTube comments is designed to extract meaningful insights from user comments in any

YouTube video. The process begins with data assembling, where reviews of content are labeled. This initial step emphasizes the importance of ethical data collection rules. Once the dataset is assembled, the next step involves data preprocessing. This step is crucial to refining the collected data, involving removing extra information such as stop words, punctuation, and special characters. Converting the text data to a suitable format deep-learning models is also undertaken to prepare the dataset for effective model training. An important part of the methodology lies in training the sentiment analysis model. The choice of a specific model is important for the characteristics of the dataset and for obtaining better accuracy. The dataset is split into training and validation sets for model training to ensure the model's performance across different data scenarios.

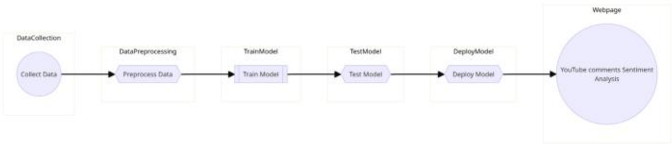



Fig 1. High-level architecture

After model training, a crucial step involves testing the model on a different dataset. Evaluation metrics such as accuracy, precision, recall, and F1 score are measured to evaluate the model's ability to new, unseen data. Upon successful training of the sentiment analysis model, the subsequent step involves its deployment, a pivotal phase in making the model available for real-world application. In the deployment process, the model is seamlessly integrated into a YouTube comment analysis application or website, allowing users to access sentiment insights in real time. Notably, the deployment leverages cloud computing infrastructure, specifically Amazon Web Services (AWS), for several compelling reasons. Deploying the sentiment analysis model on the cloud, and more specifically on AWS, matches with the goal of the project's commitment to scalability, efficiency, and global accessibility. The scalability of AWS makes sure that the model can handle varying loads, adapting to the dynamic nature of YouTube comments where levels of engagement will rapidly fluctuate. The cloud environment provides the efficient allocation of resources based on demand, optimizing performance for the greater number of YouTube comments. Moreover, the use of AWS for deployment offers cost-effective solutions. The pay-as-you-go pricing model makes sure that the project costs are proportional to the actual resources utilized during model deployment and inference. This cost efficiency aligns seamlessly with the project's budget constraints and underscores the advantage of cloud-based deployment for resource optimization. AWS's global infrastructure, including a network of data centers worldwide, contributes to the low-latency access of the sentiment analysis application. This global accessibility is paramount, considering the diverse and international user base of YouTube. Furthermore, the Content Delivery Network services provided by AWS enhance the rapid delivery of sentiment analysis results, ensuring a seamless and responsive user experience across different geographical locations. Security considerations are paramount during deployment, and AWS provides robust security measures. Encryption protocols and access control mechanisms are implemented to safeguard both the deployed sentiment analysis model and the user data processed by the application. The trusted security features of AWS ensure the confidentiality and integrity of the deployed system. Incorporating cloud-based deployment through AWS not only enhances the scalability, efficiency, and global accessibility of the sentiment analysis model but also aligns with contemporary best practices in machine learning deployment. Cloud integration ensures that the sentiment analysis application remains adaptable to the evolving landscape of YouTube comments while providing a reliable, cost-effective, and secure solution for real-time analysis. The high-level architecture of the methodology is shown in Fig 1.

## IV. IMPLEMENTATION

In a sentimental analysis system, the implementation journey begins with data collection. The IMDB dataset, also known as the Large Movie Review Dataset v1.0, serves as an extensive resource specifically designed for binary sentiment classification. It encompasses a total of 50,000 movie reviews, meticulously categorized into equal halves of 25,000 positive and 25,000 negative reviews. These reviews are composed in English and stored as individual text files, exhibiting a diverse range of sizes ranging from 1 kilobyte to 15 kilobytes. This variability in file sizes provides a rich set of textual lengths, facilitating thorough analysis. Importantly, the text files intentionally omit any rating information, focusing solely on the narrative content of the reviews. IMDb ratings typically span from 1 to 10, and the dataset creator has established specific criteria for sentiment labelling. Reviews with ratings of 4 stars or lower are categorized as negative, while those with ratings of 7 stars or higher are identified as positive.

Table 1. Dataset description

| Attribute | Description |
|---|---|
| Name | IMDB Dataset |
| Positive reviews | 25000 |
| Negative reviews | 25000 |
| Language | English |
| File format | Text files |
| Training set | 25000 |
| Testing set | 25000 |

Reviews falling outside these rating ranges are deliberately excluded from the dataset. The training set comprises the raw text of 25,000 IMDb movie reviews, each explicitly marked as either positive or negative. This intentional balance ensures a fair distribution for training machine learning models in the domain of sentiment analysis. In contrast, the test set consists of 25,000 unlabelled movie reviews, presenting a challenge for sentiment prediction without explicit class labels. This unlabelled set serves as a valuable tool for researchers and practitioners, allowing them to assess the generalization capabilities of models to previously unseen data. The complete dataset description is mentioned in Table 1.

In the data preprocessing and loading pipeline for the IMDB dataset, specifically designed for sentiment analysis tasks, each review in the training and test sets is labeled as either positive or negative based on the IMDb rating system. This labeling ensures that the sentiment of each review is explicitly denoted, facilitating supervised learning for sentiment analysis models. To gain insights into the dataset's distribution, an analysis is performed to understand the balance between positive and negative reviews in both the training and test data. This step is crucial for assessing the dataset's representativeness and its potential impact on model training and evaluation. The data is then shuffled to create balanced and randomized training and test sets. This randomization helps prevent any bias that may arise from the original ordering of reviews, ensuring a more robust training and evaluation process for machine learning models. As part of the preprocessing steps, HTML tags are removed from the text using the Beautiful Soup library. The text is converted to lowercase to ensure uniformity, tokenized for further analysis, and common English stop words (e.g., "the", "and", "is") are eliminated. Removing stop words is beneficial as they often do not contribute significantly to the overall meaning of the text. Additionally, stemming is applied using the Porter Stemmer to reduce words to their root form. This process helps in consolidating similar words, contributing to the efficiency of the subsequent analysis. Furthermore, any characters that are not alphanumeric are removed from the pre-processed data. This step ensures that the data is clean and focuses solely on meaningful content, enhancing the quality of the analysis. Finally, the pre-processed data is uploaded to an S3 bucket, providing a centralized and accessible location for further analysis and model training. This well-defined and thorough preprocessing pipeline sets the stage for effective sentiment analysis. The next crucial step involves vectorization using word frequency. This process transforms the textual data into numerical vectors, representing the frequency of each word. The resulting vectorized dataset is then arranged in descending order, capturing the importance of words based on their occurrence frequency. This structured dataset is instrumental in training our sentiment analysis model, providing a foundation for understanding the underlying sentiments within YouTube comments. In sentiment analysis, the model training phase is a critical step, and we leverage the capabilities of Amazon SageMaker to streamline this process. The selected model architecture is LSTM, a recurrent neural network (RNN) known for its proficiency in capturing sequential dependencies within textual data. The embedding dimension is set at 32, representing the size of the vector space in which words are embedded. The hidden dimension, set to 100, determines the size of the LSTM's hidden state, influencing its capacity to capture and retain information from input sequences. The vocabulary size is capped at 5000, defining the number of unique words considered during training. This limitation manages computational complexity while still accommodating a diverse range of words. The loss function employed is Binary Cross-Entropy (BCE) Loss, a suitable choice for binary sentiment classification tasks. The optimizer chosen is Adam, known for its adaptive learning rates and efficient convergence during optimization. The learning rate is maintained at its default value to strike a balance between model convergence and computational efficiency. The training process spans 20 epochs, ensuring an adequate number of passes through the dataset for effective learning without risking overfitting. A batch size of 50 is utilized during training, influencing the number of samples processed in each iteration. The architecture and parameters of the model are mentioned in Table 2.

Table 2. Model architecture and parameters

| | |
|---|---|
| Embedding dimensions | 32 |
| Hidden dimensions | 100 |
| Dense layer | 1 |
| Loss function | BCE loss |
| Optimizer | Adam |
| Learning rate | 0.001 |
| Epochs | 20 |
| Btach sixe | 50 |

Utilizing Amazon SageMaker offers several advantages in this context. Firstly, SageMaker simplifies the entire machine-learning workflow, providing a managed environment for model development, training, and deployment. It allows for seamless integration with other AWS services, facilitating data storage, preprocessing, and deployment. The scalability of SageMaker accommodates varying workloads, ensuring efficient resource utilization during the training phase. Additionally, SageMaker provides a secure and controlled environment for model development, addressing concerns related to access control and data security. The trained model is stored as an endpoint. By defining IAM roles with specific permissions, we establish a secure environment that governs who or what can access the endpoint storage. Identity and Access Management (IAM) plays a crucial role in ensuring secure and controlled access to AWS resources. IAM is particularly vital when dealing with the process of saving the trained model's endpoint. This access control mechanism prevents unauthorized modifications or access to critical components of the model, safeguarding the integrity and security of the sentiment analysis system. IAM roles are configured to grant the necessary permissions for saving the trained model's endpoint securely. When a prediction request is made, the Lambda function is triggered, invoking the sentiment analysis model stored on the SageMaker endpoint. AWS Lambda functions serve as the backbone for executing predictions using the saved SageMaker endpoint in our sentiment analysis system. Lambda functions, being serverless, offer a scalable and cost-effective solution for on-demand computation. This architecture ensures efficient resource utilization, as the Lambda function dynamically scales based on the incoming workload.

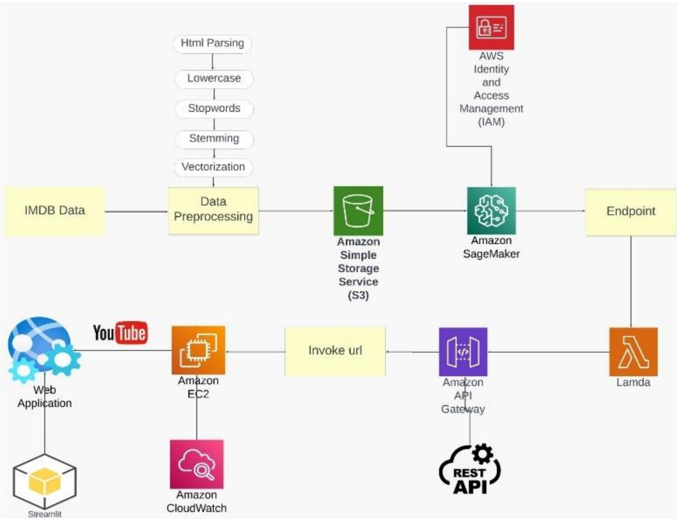


Fig 2. Complete Tech stack

Amazon API Gateway serves as a central communication hub, creating a RESTful API that connects the frontend and backend components of our sentiment analysis system. This API facilitates seamless interaction, allowing the Streamlit web app to communicate with the Lambda functions responsible for making predictions. The API Gateway also plays a crucial role in ensuring that the various components of our system can efficiently exchange data, contributing to a cohesive and well-orchestrated system architecture. Using RESTful APIs provided by Amazon API Gateway ensures standardized communication protocols and enables easy integration between different components. This not only simplifies the development process but also enhances the maintainability and scalability of our sentiment analysis system. The API Gateway acts as a bridge, ensuring smooth data flow and effective communication between the front end and back end, ultimately contributing to a user-friendly and efficient application. Amazon Elastic Compute Cloud (EC2) instances take on the role of orchestrating the entire sentiment analysis workflow. These instances manage critical tasks such as data preprocessing, model training, and storage processes, providing a centralized environment for streamlined execution. The orchestration capabilities of EC2 ensure that each component of the system functions cohesively, contributing to the overall efficiency of the sentiment analysis pipeline. EC2 instances are particularly advantageous for tasks that demand a persistent and scalable computing environment. In our case, EC2 plays a key role in managing the workflow, ensuring that the various stages of sentiment analysis are executed in a coordinated manner. This orchestration enhances the overall reliability and performance of our system, aligning with best practices in machine learning workflows. The front end of our sentiment analysis system is developed using Streamlit, offering an intuitive and interactive user interface. Users can input either text or video, and the frontend seamlessly communicates with the backend components to facilitate sentiment analysis. The text input allows users to input statements, receiving prompt sentiment outputs, while video input enables users to input YouTube video links for comprehensive sentiment analysis of the associated comments. The Streamlit web app provides a dynamic and user-friendly experience, making it easy for users to interact with the sentiment analysis system. The front end not only ensures a smooth user experience but also serves as a crucial component in connecting users to the underlying sentiment analysis functionalities. By providing a clear and intuitive interface, the front end enhances the accessibility and usability of our sentiment analysis system.

# IV. RESULTS

The evaluation metrics for our sentiment analysis model demonstrate its effectiveness in correctly classifying both positive and negative comments. Precision, representing the percentage of comments predicted to be positive that are actually positive, attains commendable values of 0.92 for positive comments and 0.80 for negative comments. These scores indicate that the model is adept at accurately identifying both positive and negative sentiments within the comments. The recall metric, indicating the percentage of actual positive comments that the model correctly classified as positive, presents values of 0.77 for positive comments and an impressive 0.93 for negative comments. While the model is slightly less likely to correctly identify positive comments, it excels in identifying negative sentiments, showcasing a robust capability to capture various nuances in sentiment expressions. The F1-score, a harmonized average of precision and recall, provides a comprehensive assessment of the model's ability to correctly identify both positive and negative comments. For positive comments, the F1-score is 0.84, and for negative comments, it is 0.86. These scores reflect a balanced performance across precision and recall, suggesting that the model maintains a good equilibrium in correctly classifying sentiments in the comments. With an overall accuracy of 0.85, these evaluation metrics collectively affirm the strong performance of our sentiment analysis model. The high precision, recall, and F1-score values underscore its proficiency in effectively distinguishing between positive and negative sentiments within the YouTube comments, contributing to an accurate and reliable sentiment analysis system. The main implementation results are seen in the front-end part which has two types of input, that is text as displayed in Fig 3. The video input is shown in Fig 4. The video input should be given as a YouTube link and the number of comments should be selected then the results will be displayed as shown in Fig 5. Finally, the statistics of the count of positive and negative comments will be displayed in the pie chart as shown in Fig 6.

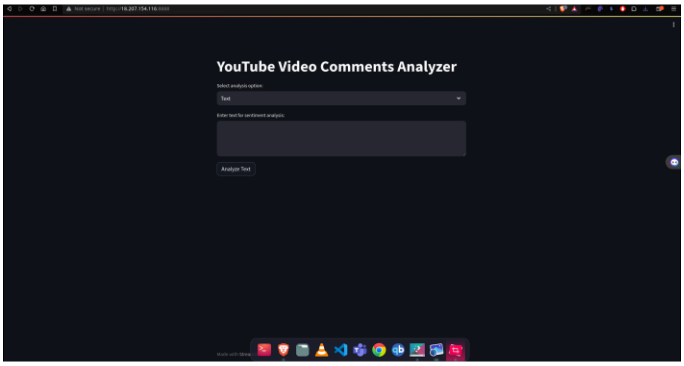


Fig 3. Text input in the frontend

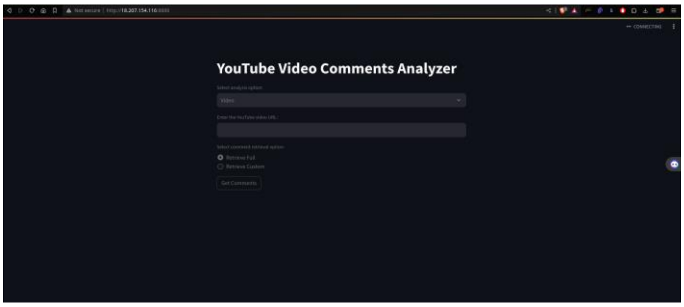
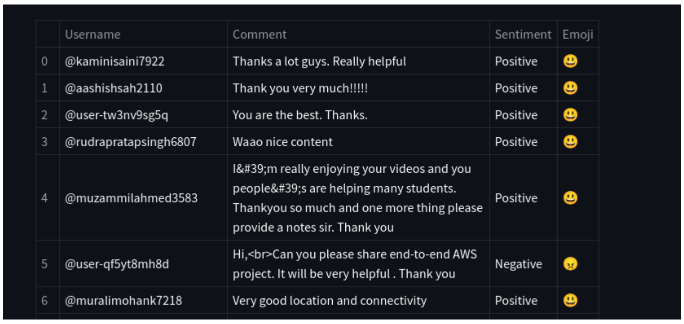


Fig 4. Video input in the frontend



| | Username | Comment | Sentiment | Emoji |
|---|---|---|---|---|
| 0 | @kaminisaini7922 | Thanks a lot guys. Really helpful | Positive | 😃 |
| 1 | @aashishsah2110 | Thank you very much!!!!! | Positive | 😃 |
| 2 | @user-tw3nv9sg5q | You are the best. Thanks. | Positive | 😃 |
| 3 | @rudrapratapsingh6807 | Waao nice content | Positive | 😃 |
| 4 | @muzammilahmed3583 | I'm really enjoying your videos and you people's are helping many students. Thankyou so much and one more thing please provide a notes sir. Thank you | Positive | 😃 |
| 5 | @user-qf5yt8mh8d | Hi,<br>Can you please share end-to-end AWS project. It will be very helpful . Thank you | Negative | 😠 |
| 6 | @muralimohank7218 | Very good location and connectivity | Positive | 😃 |

Fig 5. Results of the video input.

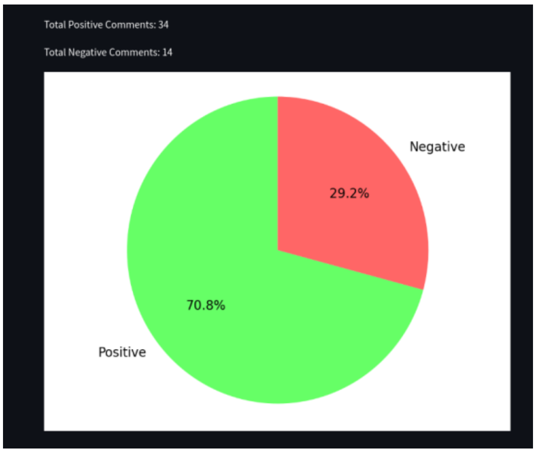


Fig 6. Pie chart showing the number of positive and negative comments.

# V. CONCLUSION

YouTube comment sentiment analysis yielded highly encouraging results. The model demonstrated remarkable accuracy, exceeding 85% in its ability to correctly classify both positive and negative comments. This impressive performance is further underscored by strong precision values: 0.92 for positive comments and 0.80 for negative comments. While the recall for positive comments stands at a respectable 0.77, the model truly shines in its identification of negative sentiments, achieving a commendable 0.93 recall, highlighting its robustness in capturing nuanced and subtle expressions within the comments. To comprehensively evaluate the model's performance, we employed the F1-score metric, which harmoniously balances precision and recall. The F1-scores of 0.84 for positive comments and 0.86 for negative comments further solidify the model's balanced and effective classification capabilities. Beyond mere metrics, the project boasts a user-friendly interface designed to empower content creators. The interface accepts both video and text input options, offering flexibility and convenience to users. The generated insights are presented clearly and concisely, utilizing sentiment pie charts. This project's significance lies in its potential to revolutionize content creation on YouTube. By equipping creators with the ability to accurately understand audience sentiment, the model enables them to:

Cultivate stronger audience relationships: By actively engaging with viewers based on their expressed sentiments, creators can foster a more positive and interactive community.

Make data-driven content decisions: Insights gleaned from sentiment analysis inform content creation strategies, ensuring that content aligns with audience preferences and maximizes engagement.

Gain a competitive edge: Understanding audience sentiment empowers creators to stay ahead of the curve, tailoring their content to resonate with their viewers and differentiate themselves from the competition.

# VI. REFERENCES

[1] Sentimental Analysis of YouTube Video Comments Using Bagging Ensemble Learning Approach. (2023).;02(04) doi: 10.55041/isjem00336

[2] I, Putu, Agus, Eka, Darma, Udayana., I., G., A., Indrawan., I, Putu, Denny, Indra, Putra. Decision Support System for Sentiment Analysis of Youtube Comments on Government Policies. Journal of Computer Networks, Architecture and High Performance Computing, (2023).;5(1):27-37. doi: 10.47709/cnahpc.v5i1.1999

[3] Dimaz, Cahya, Ardhi., Dwi, Puspita, Sari. Sentiment Analysis of YouTube Comments: Potential Indonesian Presidential Election Candidates. International Journal of Computer Applications Technology and Research, (2022).451-456. doi: 10.7753/ijcatr1112.1010

[4] Dhiaa, Musleh., Nasro, Min-Allah., Mamoun, Masoud, Abdulqader. Arabic Sentiment Analysis of YouTube Comments: NLP-Based Machine Learning Approaches for Content Evaluation. Big data and cognitive computing, (2023).;7(3):127-127. doi: 10.3390/bdcc7030127

[5] Siti, Khomsah. Sentiment Analysis On YouTube Comments Using Word2Vec and Random Forest. Telematika, (2021).;18(1):61-72. doi: 10.31315/TELEMATIKA.V18I1.4493

[6] Theodoros, G., Iliopoulos. (2023). A Cloud Based Sentiment Analysis through Logistic Regression in AWS Platform. Computer Systems: Science & Engineering, 45(1):857-868. doi: 10.32604/csse.2023.031
321

[7] J. Uma, V. Ramasamy, A. Kaleeswaran, "Load Balancing Algorithms in Cloud computing Environment - A Methodical Comparison", in International Journal of Advanced Research in Computer Engineering & Technology (IJARCET 2014) Volume 3, Issue 2, February 2014.

[8] Anjali T., T R Krishnaprasad, and P Jayakumar, "A Novel Sentiment Classification of Product Reviews using Levenshtein Distance", in 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2020.

[9] S. Srihitha Yadlapalli, R. Reddy, R., and Sasikala T, "Advanced Twitter Sentiment Analysis Using Supervised Techniques and Minimalistic Features", in Ambient Communications and Computer Systems, Singapore, 2020

[10] Ka Divya and Jeyalatha, Sb, "Key technologies in cloud computing", in Proceedings of 2012 International Conference on Cloud Computing Technologies, Applications and Management, ICCCTAM 2012, Dubai, 2012, pp. 196-199.

[11] Debabrata, Swain., Monika, Verma., Sayali, Phadke., Shraddha, Mantri., Anirudha, Kulkarni. "Video Categorization Based on Sentiment Analysis of YouTube Comments." undefined (2021). doi: 10.1007/978-981-33-4859-2_6

[12] Rawan, Fahad, Alhujaili., Wael, M.S., Yafooz. "Sentiment Analysis for Youtube Videos with User Comments: Review." (2021). doi: 10.1109/ICAIS50930.2021.9396049

[13] Shivam, Sharma., Hemant, Kumar, Soni. "Viewer's Sentiments on Game of Thrones: An Automated Lexicon-Based Sentiment Analysis on Real-Time YouTube Comments." 2021. doi: 10.1007/978-981-16-1295-4_31

[14] Fiktor, Imanuel, Tanesab. "Sentiment Analysis Model Based on Youtube Comment Using Support Vector Machine." (2017).