YOUTUBE SENTIMENTS ANALYSIS USING NATURAL LANGUAGE PROCESSING

Guide Name

Prof. Aparna Patil

Pune

aparnapatil16@gmail.com

Puja Pravin Kasar

Student

Pune

pujakasar15@gmail.com

Shivam Eknath Khandare

Student

Pune

khandareshivam28@gmail.com

Nikita Pradip Erande

Student

Pune

erandenikita26@gmail.com

Nikita Subhash Bankar

Student

Pune

nikitabankar92000@gmail.com

ABSTRACT

We live in the era where majority of the population resorts to social media platforms to get their daily digest of news, entertainment, livelihoods, etc. YouTube is one such social media platform that hosts majority of the population. YouTube attracts lot of eyes when it comes to provide entertainment, news, etc. The issue with this platform is that the majority of the population is unconsciously consuming biased media content and it only gets worse when we know the reality that there is no mechanism in existence that addresses this problem.

Through this project we aim to address this issue. Our focus is to build a mechanism that

leverages the power of modern computing advances i.e., Machine Learning to come up with a model that performs a sentiment analysis of the content. Sentiment analysis refers to the field of analyzing and evaluation of - opinions, sentiments, attitudes, and emotions of users via scrutinizing media content posted on social media. The analysis is performed in real time so that any user who is consuming some content on YouTube gets an idea that whether the content being played is biased or not, whether the content can hurt other individual sentiments or is neutral.

1. INTRODUCTION

Social media websites include blogging, micro-blogging, streaming applications like Twitter, WhatsApp, YouTube, etc. These platforms are popularly used by individuals to express their sentiment and opinion on a variety of topics. For many of these applications the main medium of communication is via text.

YouTube as a platform relies on video i.e., it uses video as main medium for information sharing. For example, consider the case of unboxing of a newly bought product, here a user videorecords the complete procedure of unboxing of a product, during this process the user also shares his/her thoughts on the quality of the product bought, passes reviews and other metadata that will be potentially used by hundreds of other customers to decide whether to make the purchase or not. It is very clear from above example how important it is to have a checking mechanism that prompts or alert a user that the content they are watching is biased or unbiased because based on the content they are consuming; they will later manifest their opinion on a certain topic.

Emotional analysis can help to alleviate this hatred on a large scale. Sensitive systems that can scratch these information resources can help create popular emotions. Such
There exists number of techniques that perform text-based sentiment analysis. Through this model we had shown that by using a combination of NLP (Natural Language Processing) & ASR (Automatic Speech Recognition), it is possible to have audio sentiment extraction of good accuracy. In a very brief, we followed following steps to complete the task:

1. Extraction of audio from video content.
2. Conversion of this audio to text using ASR information is often used to help and grow a business.
3. Polarity analysis on the text to predict the sentiment
The text-based emotional system used a method of marking parts of speech to automatically extract text features, which were then used in a large-scale entropy segmentation system to predict emotional variability.

2. LITERATURE SURVEY

Sentiment analysis is processing/churning of text to identify and extracts subjective information in source material to help understand the social sentiment of the media while monitoring online content Analysis of social media streams is usually restricted to just basic sentiment analysis and count based metrics. This is similar to just scratching the surface and missing out on those high value insights that are waiting to be discovered.
Creative use of advanced Machine Learning techniques can be an effective tool for doing in- depth research.
We believe it is important to classify incoming customer conversation about a brand based on
following lines:
1. Key aspects of a brand's product and service that customers care about.
2. Users' underlying intentions and reactions concerning those aspects. These basic concepts when used in combination, become a very important tool for analyzing millions of brand conversations with human level accuracy. The age of getting meaningful insights from social media data has now arrived with the advance in technology. The Uber case study gives you a glimpse of the power of Contextual Semantic Search

It is time for organizations to move beyond emotions and calculate supported metrics. Companies have been using the power of data lately, but to get the depth of knowledge, you have to use the power of AI, in-depth learning and clever categories such as Contextual Semantic Search and emotion analysis.

3. OBJECTIVES OF SYSTEM

Sentiment analysis is a method of natural language processing (NLP) used to predict whether data is fair, negative or neutral. Emotional analysis is often done on text data to monitor product and product sentiment in customer feedback, and to understand customer needs.

When we do a YouTube search for videos, more often than not, it happens that the caption displayed at the bottom of video player doesn't match appropriately with the content displayed. Our project aims to analyze the caption displayed and match it with the video content for the appropriateness, & in case of faulty or wrong caption, suggest a correct caption sequence.

4. EXISITING SYSTEM

Sentiment analysis is a process that discovers the user opinions and views against any service or a product. YouTube is one of the most popular videos sharing platforms obtaining millions of views.
In YouTube video it contains valuable information that helps in improving the knowledge related uploaded video content. The content is utilized by using natural language processing techniques. There are many attempts had been proposed scholarly with three (positive, negative or neutral) classes. It is use to perform sentiment analysis on YouTube Content in identifying the polarity as well.

5. IMPLEMENTATION
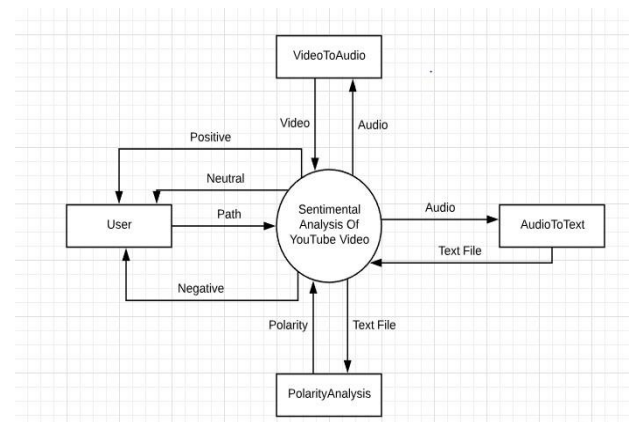
6. EXPERIMENTAL RESULTS
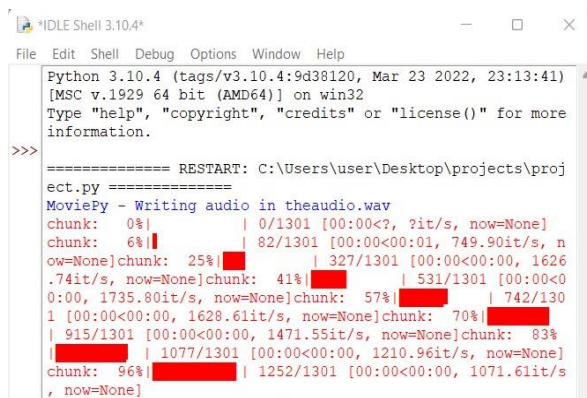
❖ Input
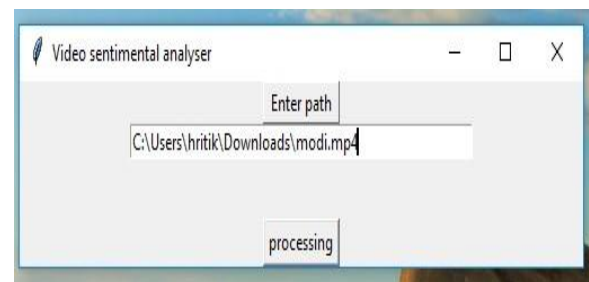
This figure depicts:

- The GUI used in this project is Tkinter

- Here we designed a dialogue box where input field is designed to enter the address of video on the processing must be implemented.

- In the input field the path of video is entered as an input option and from there processing begin

- Part of video is fetched and converted into training model for deciding polarity and predicting the result.
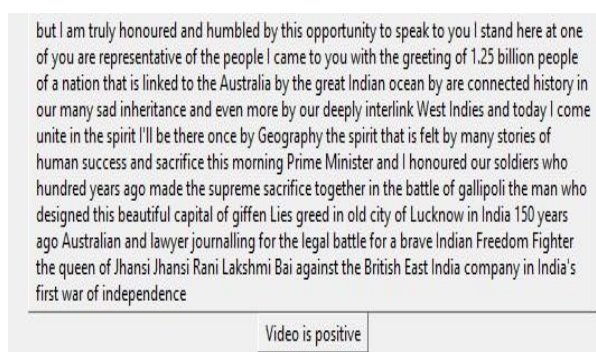
❖ **Processing**

The above figure states the following

- After the path of video is entered and the processing button pressed as shown in fig the processing starts

- Firstly, with the help of Movie Pie editor library of python, the video is cropped within some time frame

- Then that snippet of video is converted into text format using google API Speech Recognition

- The text format act as dataset for trained model of NLP and polarity is detected within that time instance

- Above processing is being showed where chunk of word in aware dictionary is matched with our text.

❖ **Output**

The output generated after processing chunk completed 100%. In the input field we have taken the snip of video of India's PM fig 4, who was awarded for peace. All his voice is displayed in this GUI in words. At the end of content, a result is displayed depicting the polarity of all the above textual matter.

- The overall verdict of the video i.e., 'Positive', 'Negative' or 'Neutral'

- Converted text file for the audio present in the video

- Selected portion of the text file upon which the verdict is heavily dependent

- Measures the polarity of the video (ranges from -4 to 4). This measurement is done by Vader

- Vader sentimental analysis is powerful tool of NLP algorithm which calculate the polarity of each word and gives the compound score of word, sentence, or paragraph
  Compound score is basically average of neutral, positive, and negative score.

7. CONCLUSION

At the end of our investigation, we came to the conclusion that gas leaks in homes and businesses cause risks to people's lives and property. So, our objective will provide a technique to prevent such mishaps by sending notifications to pre-programmed mobile numbers and turning on the LCD display and buzzer. In essence, it is a good project.

8. REFERENCES

[1] Ayushka Tiwari and Ritika Singh. "YouTube Comments Sentiment Analysis". Date: May – 2021 International Journal of Scientific Research in Engineering and Management (IJSREM). ISSN: 2582-3930

[2] Mehta, Munish, Kanhav Gupta, and Shubhangi Tiwari. "A Review on Sentiment Analysis of Text, Image and Audio Data." 2021 5th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2021.

[3] H. Jelodar et al., "A NLP framework based on meaningful latent-topic detection and sentiment analysis via fuzzy lattice reasoning on youtube comments", pp. 41554181, 2021.

[4] Chen H, Xie W, Vedaldi A, Zisserman A. Vggsound: A large-scale audio-visual dataset. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. 2020;721-725.

[5] Wang J, Yu LC, Lai KR, Zhang X. Tree-structured regional CNN-LSTM model for dimensional sentiment analysis. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2019 ; 28:581-91

[6] M. H. A. El-jawad and R. Hodhod, "Sentiment Analysis of Social Media Networks Using Machine Learning", 2018 14th Int. Comput. Eng. Conf, pp. 174-176, 2017.

[7] H. Bhuiyan, J. Ara, R. Bardhan and M. R. Islam, "Retrieving YouTube video by sentiment analysis on user comment", Proc. 2017 IEEE Int. Conf. Signal Image Process. Appl. ICSIPA 2017, pp. 474-478, 2017.

[8] F. I. Tanesab, I. Sembiring and H. D. Purnomo, "Sentiment Analysis Model Based on Youtube Comment Using Support Vector Machine", vol. 6, no. 8, pp. 180-185, 2017.