YouTube Sentiment Analysis

A

Project Report

Submitted for the partial fulfilment

Of B.Tech Degree

in

INFORMATION TECHNOLOGY

by

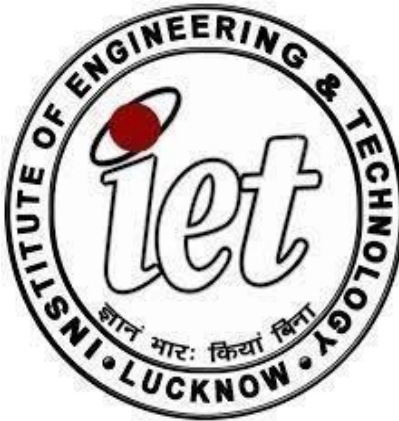**Abhishek Verma (1805213004)**

Harsh Jalan (1805213022)

Sanjay Gautam (1805213049)

Under the supervision of

Dr. Natthan Singh

Mr. Mahima Shanker Pandey

Department of Computer Science and Engineering

Institute of Engineering and Technology

Dr. A.P.J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh

CONTENTS

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our belief and knowledge, it contains no material previously published or written by another person or material which to a substantial error has been accepted for the award of any degree or diploma of university or other institute of higher learning, except where the acknowledgement has been made in the text. The project has not been submitted by us at any other institute for requirement of any other degree.

Submitted by: -   Date:

(1) Name: Abhishek Verma

Roll No.: 1805213004

Branch: Information Technology

Signature:

(2) Name: Harsh Jalan

Roll No.: 1805213022

Branch: Information Technology

Signature:

(3) Name: Sanjay Gautam

Roll No.: 1805213049

Branch: Information Technology

Signature:

CERTIFICATE

This is to certify that the project report entitled "YouTube Sentiment Analysis" presented by Abhishek Verma, Harsh Jalan and Sanjay Gautam in the partial fulfilment for the award of Bachelor of Technology in Information and Technology, is a record of work carried out by them under my supervision and guidance at the Department of Computer Science and Engineering at Institute of Engineering and Technology, Lucknow.

It is also certified that this project has not been submitted at any other Institute for the award of any other degrees to the best of my knowledge.

Dr. Natthan Singh

Mr. Mahima Shanker Pandey

Department of Computer Science and Engineering
Institute of Engineering and Technology, Lucknow

ACKNOWLEDGEMENT

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 hurts other individual sentiments or is neutral.

LIST OF FIGURES

CHAPTER – 1

1.1 INTRODUCTION

Social media websites includes 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 videorecord 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 information is often used to help and grow a business.

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

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.

1.2 MOTIVATION

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.

1.3 OBJECTIVES

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.

CHAPTER – 2

LITERATURE REVIEW

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 [1]. 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 [2].

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 [3]

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..
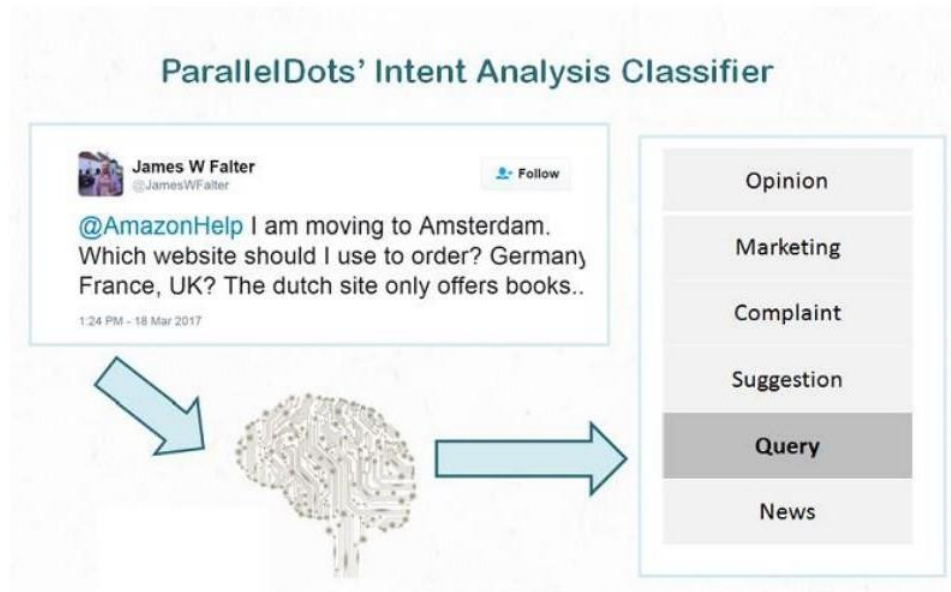
Visual representation of Intent classification as done by ParallelDots API

Image Source : blogs.paralleldots.com

Fig. 1 [4]

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 displays

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.

It also aims to perform sentiment analysis on YouTube videos & perform categorization

2.1 Sequence Diagram

Sequence diagram is a collaborative diagram with details of how tasks are performed - what messages are sent and when. They are sorted by time.

Functional items are listed from left to right depending on when they are participating in the message sequence.
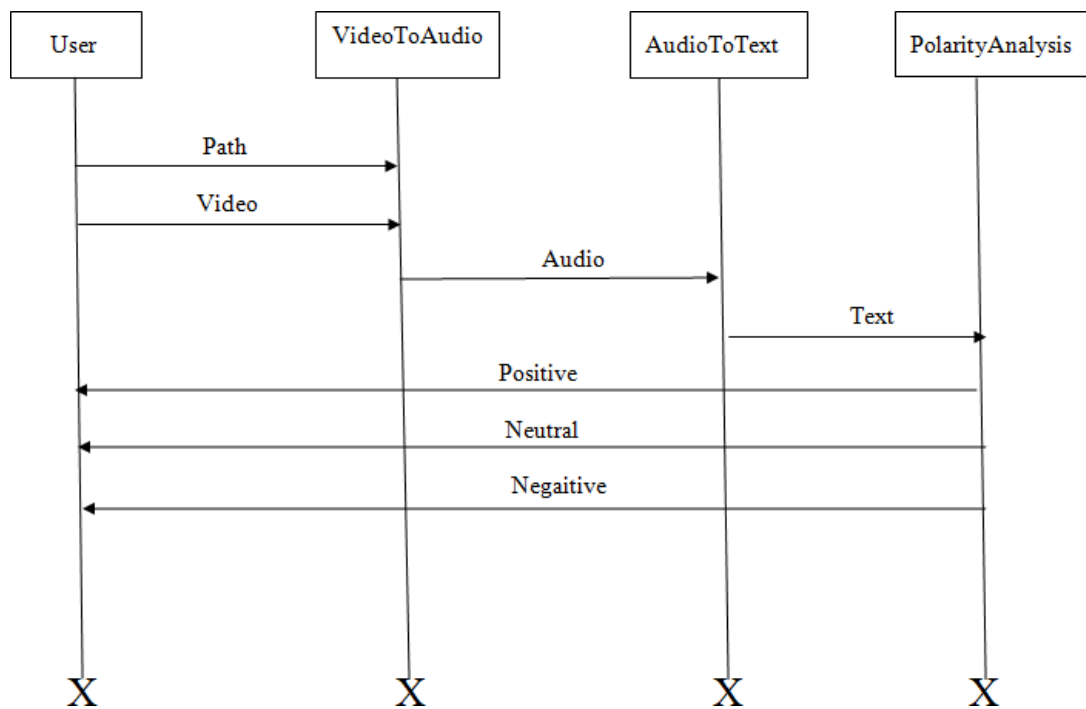
Fig. 2 Sequence Diagram

2.2 Data Flow Diagram

The Data Flow (DFD) diagram shows the flow of information of any process or process. Uses defined descriptions such as rectangles, circles and arrows, as well as short text labels, to indicate inputs, outputs, archives and routes between each destination.
Attached below is a Data Flow Diagram to depict the flow of information as well as associated processes that deals the complete flow of sentiment analysis of YouTube Video.

Fig. 3 Data Flow Diagram

- Above diagram is Data flow diagram of our project
- Data flow diagram depicts the stream of data in this project
- Each path is entitled with some processing and tools used in each processing

CHAPTER - 3

METHODOLOGY

YouTube is overall a great platform, which explains how it has been popular for so long. However it does have its own share of flaws.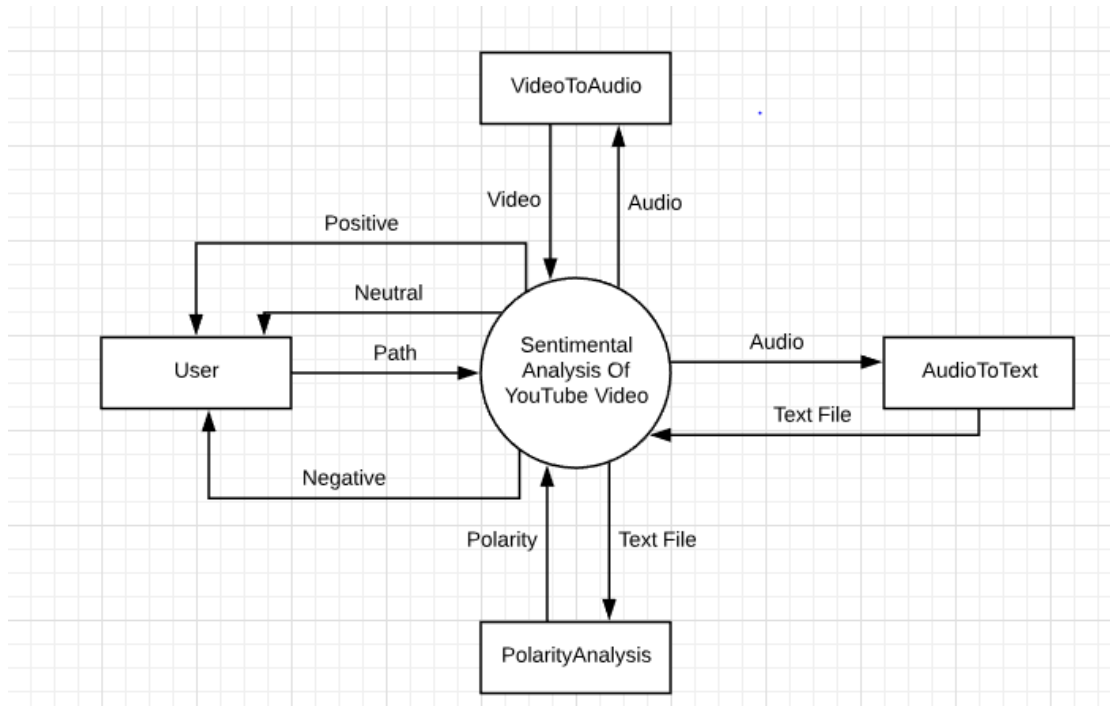 Two of the biggest problems most users face are bad recommendation videos, and spam comments. It could help in better subtitles generation.

3.1 Modules

Our project is composed of 5 modules, each of these modules have to be completed in a sequential order for completion of the project.

We have discussed our approach for building interface of the application, what libraries to use to perform video to audio and then audio text conversion and the algorithm used for performing sentiment / polarity analysis.

3.1.1 Interface of the Application

Tk/Tcl [4] has long been an important tool in Python. It provides a easy and platform independent windowing toolkit, using the tkinter package, and its extension- tkinter tix and the tkinter.ttk modules. The tkinter package is a thin object-orientedlayer on top of Tcl/Tk. To use tkinter you don't need to write Tcl code, but you will need to consult the Tk documentation, and occasionally the Tcl documentation. Tkinter is a set of wrappersthat implement the Tk widgets as Python classes. In addition, the internal module tkinter providesa threadsafe mechanism which allows Python and Tcl to interact.

Tkinter's chief virtues are fast, and that it usually comes bundled with Python. Althoughits standard documentation is weak, good material is available, which includes: references, tutorials, a book and others. tkinter is also famous for having an outdated look and feel, which has been vastly improved in Tk 8.5 (standard GUI).

In Tk, there is a utility command, interacting with the window manager. Options to the command allow you to control things like titles, placement, icon bitmaps, and the like. In tkinter, these commands have been implemented as methods on the class [5]. Toplevel widgets are subclassed from the class, and so can call the methods directly.

To get at the top-level window that contains a given format, often just refer to the widget'smaster. if the widget has been packed inside of a frame, the master won't represent a top-level window. To get at the top-level window directly call the root()method [9].

3.1.2 Video To Audio

MoviePy : MoviePy is an open source software written by Zulko [3] and released underthe MIT licence. It works on Windows, Mac, and Linux, with Python 2 or Python 3. There is alsoa MoviePy forum on Reddit .It is a Python module for video editing, which can be used for basic operations (cuts, concatenations, insertions), video compositing (non-linear editing), video processing, or to create advanced effects.

1. You have many videos to process or to compose in a complicated way.
2. You want to automatize the creation of videos on a web server.
3. You want to automatize tedious tasks, like insertions tracking objects, cutting scenes, making end credits, subtitles.
4. You want to code your own video effects to do something no existing video editor can.
5. You want to create animations from images generated by another python library (Matplotlib, Gizeh, scikit-images)

3.1.3 Audio To Text

Speech recognition API : Speech recognition is the process of converting spoken words to text. Python supports many speech recognitio and APIs, including Google Speech Engine, Google C cloud Speech API, Microsoft Bing Voice Recognition and IBM Speech to Text.

1. Speech must be converted fromphysical sound to an electrical signal with a microphone,

2. and then to digital data with an analog-to-digital converter.

3. Once digitized, several models can be used to transcribe the audio to text.

This method works by assuming that the speech signal, when measured at a time short enough to say, ten milliseconds, can be measured logically as a static procedure- the mathematical features do not change over time. Most modern speech recognition systems rely on the so-called Hidden Markov Model [6].

3.1.4 Speech-to-Text API recognition

The Speech-to-Text API Compatibility Application is an easy way to capture attention to speech audio data. Text-to-Text can process 1 minute of audio audio data sent with a consistent request. After speech-to-text processes and detects all sounds, it returns feedback.

Sync request is blocked. Text-to-Text usually processes audio faster than real time, processing 30 seconds of audio in 15 seconds on average. The first part of speech recognition is, of course, speech. Speech should be converted from portable audio to a portable electronic signal, and then to digital data via an analog to digital converter. Once digitized, a few models can be used to record audio into text.

In low audio quality cases, your alert request may take longer. All compatible Speech-to-Text API applications should include speech recognition field.In normal HMM, speech signal is separated by 10-millisecond pieces. The energy spectrum of each episode, which is a signal strength structure as a frequency function, is mapped to a vector map of real numbers known as cepstral coefficients. The vector size is usually small — sometimes as low as 10, although more accurate systems may be 32 or more in size. The final result of HMM is the sequence of these vectors.

All of the magic in Speech Recognition happens with the 'Recognizer' class.

The primary purpose of a Recognizer instance is, of course, to recognize speech. Each instance comes with a variety of settings and functionality for recognizing speech from an audio source

3.1.5 Text To Polarity Analysis

In this module we have used **VADER Sentiment Analysis.**
VADER (Valence Aware Dictionary and sEntiment Reasoner) is a legal-based emotion analysis tool focused on emotions expressed on social media.

Using the A sentiment dictionary combination is a list of dictionary features (e.g., words) that are usually labeled according to their semantic structure such as good or bad.

VADER not only tells us about Positivity and Negativity score but also tells us how good or bad it feels..

It is fully open-sourced under the MIT License.

The simplest way is to use the command line to do an installation from [PyPI] using pip.

>> pip install vaderSentiment
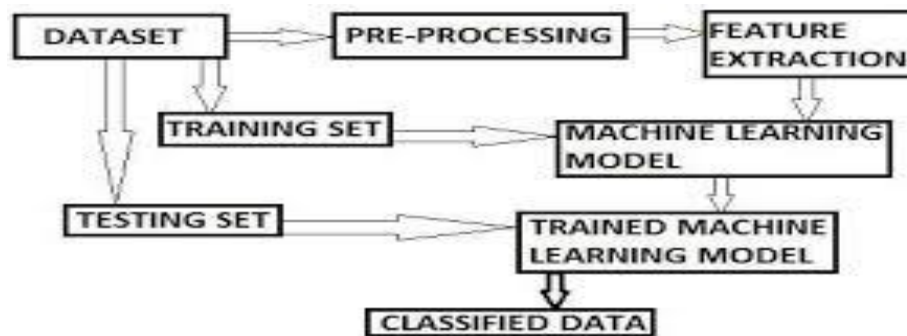
Fig 4 Flow chart for Vader Analysis [4]

- The dataset is of form of lexicon-based matter.it contain GIFs,emojis,text
- Valence Aware dictionary for sEntimenal Analysis matched every word with its dictionary called as aware dictionary and gives each word score which lies between -4 to +4
- It is supervised learning NLP model which is inbuilt as library in python
- It matches the word with and decides a score positive (from 0 to 4),neutral(0),negative(-4 to 0)

Table 1
Sentiment Score Table

Sentiment Metric	Score
Positive	0.674
Neutral	0.326
Negative	0.0
Compound	0.735

- The scores represented by the 'Positive', 'Negative' and 'Neutral' labels corresponds to the proportion of text falling in these categories. Here entered sentence was rated as 67% Positive, 33% Neutral and 0% Negative.

- The Compound score is a metric that calculates the sum of all the lexicon ratings which have been normalized between -1(most extreme negative) and +1 (most extreme positive). In the case above, lexicon ratings for some words are 2.9 and 1.3. The compound score will turn out to be 0.75. This represents very high positive sentiment.

- Positive sentiment: (compound score >=0.05)

- Neutral sentiment:(compound score >-0.05) and (compound score <0.05)

- Negative sentiment :(compound score <=-0.05)

CHAPTER – 4

EXPERIMENTAL RESULTS

Fig 5 PM Narendra Modi addressing the 14th Seoul Peace Prize Award Ceremony

- Above image is taken from a video where Prime Minister of India, Mr. Narendra Modi has been conferred with 14th Seoul Peace Prize award ceremony.
- Our project does a sentiment analysis on Prime Minister Narendra Modi commencement to the audience.
- The video is entered as the input field in the GUI.
- Every second of video is analysis using Vader sentiment
- The polarity is depicted
- Based on the metric score, polarity of the video will be determined as to whether the video is positive, negative, neutral.
- Configuration of video used is

  - Type -MP4
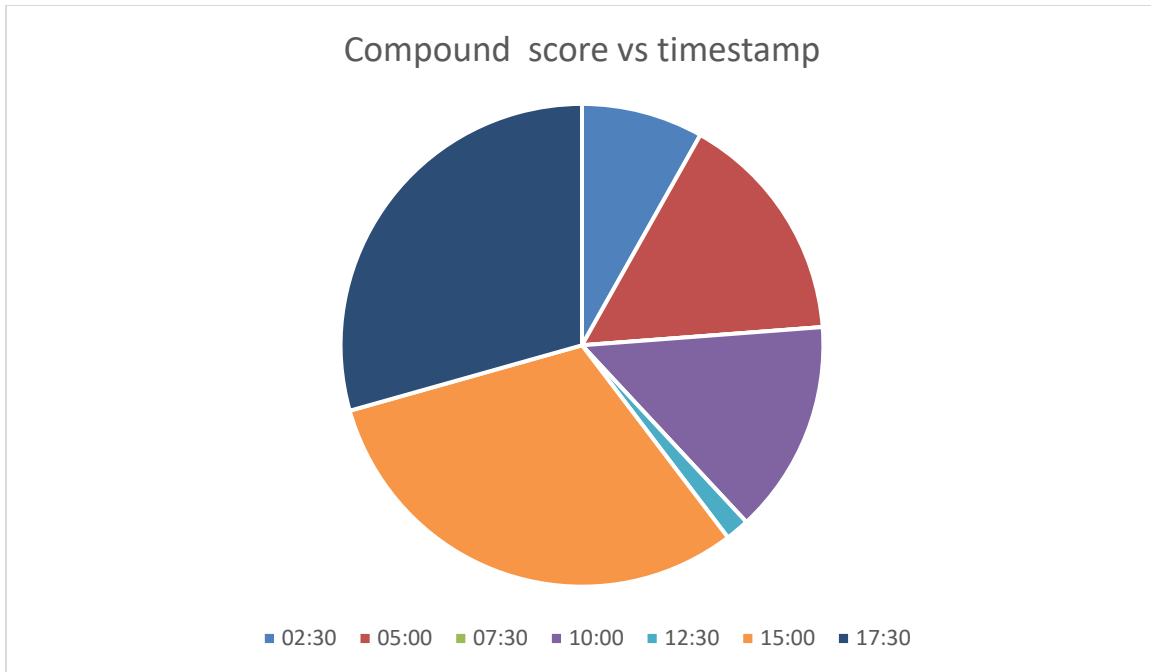 - Size- 56 MB
  - Time duration-19:35 min

Fig 6 Computational graph

- After the successfully running the model in the console the computational graph in the pie chart is drawn which depicts the metric score with time score
- After interval of 2 min 30 sec compound score of text is measured using computational technique and the graph is plotted with score and time
- The graph also depicts the overall polarity of video ,whether it is positive, neutral or negative.
- The time frame of 2min30sec is set in moviePie editor and everytime the video is converted to text format then using vader sentimental polarity is calculated.
- The code dynamically after every 2:30min and its score is calculated all way
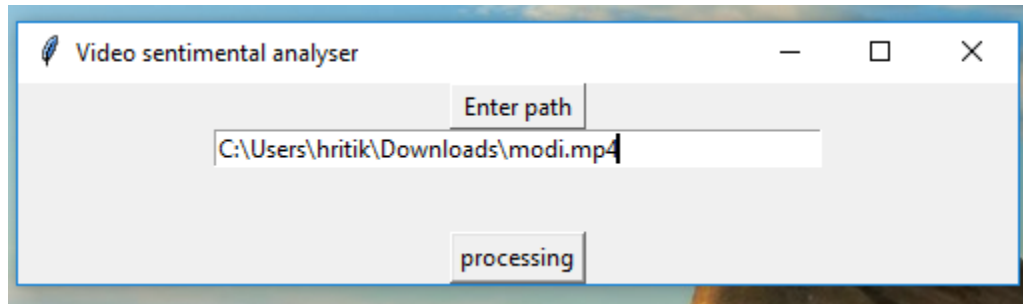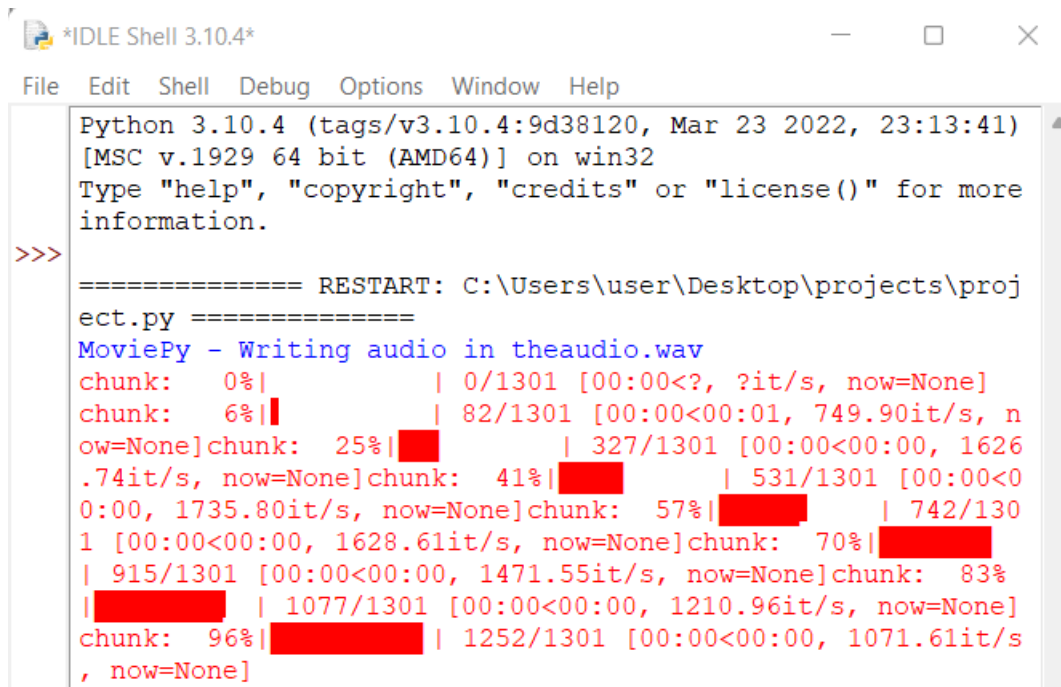
4.1 INPUT

Fig 7 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 has to 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

- Tkinter is the object-based technique to designed the GUI

- There are 15 widget in tkinter here we have used only 4 widget.

- The dimension of the box used in the model is 500x100x300

- Three methods are used by tkinter for widget:- pack(), grid(), place()

- In this module we have used pack() method to specify the geometry of widgets

- Tkinter is platform independent compiled on every supporting devices

- For the installing tkinter the syntax is import tkinter.tk()

4.2 PROCESSING

Fig 8 Text mining process

The above figure states the following

After the path of video is entered and the processing button pressed as shown in fig 6 the processing starts

- Firstly with the help of MoviePie 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.

4.3 OUTPUT

but I am truly honoured and humbled by this opportunity to speak to you I stand here at one of you are representative of the people I came to you with the greeting of 1.25 billion people of a nation that is linked to the Australia by the great Indian ocean by are connected history in our many sad inheritance and even more by our deeply interlink West Indies and today I come unite in the spirit I'll be there once by Geography the spirit that is felt by many stories of human success and sacrifice this morning Prime Minister and I honoured our soldiers who hundred years ago made the supreme sacrifice together in the battle of gallipoli the man who designed this beautiful capital of giffen Lies greed in old city of Lucknow in India 150 years ago Australian and lawyer journalling for the legal battle for a brave Indian Freedom Fighter the queen of Jhansi Jhansi Rani Lakshmi Bai against the British East India company in India's first war of independence
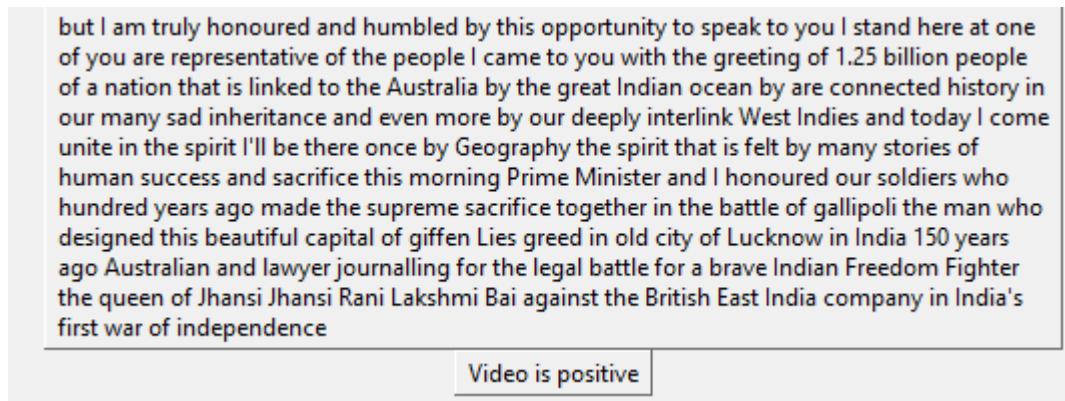
Video is positive

Fig 9 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 .

CHAPTER – 5

5.1 CONCLUSION

Through this project we have demonstrated how organizations can leverage modern computing augmented with Machine Learning to perform a Sentiment Analysis on a YouTube Video.

The process starts with first converting the Video/Audio file to text file. This textual format of media serves as input data which are used to train Machine Learning Model. We used text-to-speech API to generate the text (input data).

This input data then underwent a series of processing via VADER analysis. VADER analysis generated a sentiment score metric. The sentiment score for the entire input file was generated and if the average value was found above a certain threshold, video was termed as positive else if score turned out to be zero, then neutral, else negative.

5.2 RELATED WORKS

To have a comparative study between our model and other models made earlier which also perform sentiment analysis on various dimension of YouTube content, we have listed a table which mentions the domain on which sentiment analysis has been performed as well as the type of Machine Learning model, algorithm used to create the model.

Author Name	Sentimental Analysis	Methodology Approach	Description
Ayushka Tiwari ISSN 2582-3930 [1]	YouTube comments	Naïve basis, SVM, decision tree	Sentimental analysis of public comments by using an annotated corpus consists of citation sentences. The corpus is made up of 1500 citation sentences. The corpus is annotated using some rules to assign the polarity to citation sentences.
H. Bhuiyan, J. Ara, R. Bardhan and M. R. Islam, Proc. 2017 IEEE Int. Conf. Signal	YouTube reviews, comments	Unsupervised lexicon-based approach STRUCT-SVM	They adopted a knowledge driven approach and ready a social media list of terms and phrases expressing 6 the user sentiment and opinion. But their results also showed that recall of negative sentiment is poorer compared to the positives, which can flow from to the wide linguistic variation utilized in

Image Process [7]			expressing frustration and dissatisfaction
M. H. A. El-Jawad and R. Hodhod, 2018 14th Int. Comput. Eng. Conf, pp. 174-176, 2017 [6]	Social Media Networks-twitter	Artificial Neural Networks	They examined the hypothesis that sentiment analysis are often reated as a special case of topic-based text classification. Their work depicted that standard machine learning techniques such as Naive Bayes or Support Vector Machines (SVMs) outperform manual classification techniques that involve human intervention.
F. I. Tanesab, I. Sembiring and H. D. Purnomo, vol. 6, no. 8, pp. 180-185,2017 [8]	Sentiment Analysis Model Based On Youtube Comment	SVM, K-clustering	The process of taking a certain comment from the text had used a snipping method. The data is taken randomly based on recent video results that appears in YouTube.
Our Model	Sentiment Analysis of YouTube Videos	NLP-Vader, Speech recognition , moviepy	Here lexicon-based approach is used to identify the polarity of video played on the social media platform like YouTube

5.3 FUTURE WORKS

I.   Making the model using real time values and work in real time.

II. Then calculating the accuracy of new model and comparing it with this model and see if there is little or no difference.

III. Training the model for recognition of other languages other than English.

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. 4155-4181, 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.