JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Sentiment Analysis on Multimedia

[1]Prof. Reena Deshmukh, [2]Faraz Baig [3]Shubham Chavan, [4]Prachi PatilAssistant Professor, Student, Student, Student

[1,2,3]Department of Computer Engineering, University of Mumbai, Shivajirao S. Jondhale College of Engineering, Dombivli, (E.)

Abstract: *Sentiment analysis often referred to as opinion mining, has gained immense importance and relevance in the context of multimedia content, such as images, videos, and audio, due to the ever-increasing volume and diversity of multimedia data being generated and shared across various platforms. This form of analysis goes beyond traditional text-based sentiment analysis by extracting and understanding the emotions and opinions expressed in non-textual media. In an era dominated by user-generated content on social media, sentiment analysis on multimedia serves several crucial purposes. It enables businesses to gauge customer reactions to their products through image & video posts, helping them refine marketing strategies and improve products. Furthermore, as we move towards more immersive technologies like virtual and augmented reality, sentiment analysis on multimedia content can enhance user experiences personalizing content and adapting it to individual emotional response.*

Keywords:- Natural Language Processing, Sentiment Analysis, Neural Network, Image Processing.

I. INTRODUCTION

Sentiment analysis is a powerful computational technique used to identify and categorize the opinions and emotions expressed within various forms of communication, such as text, speech, and multimedia content. It delves into understanding people's emotional responses, attitudes, and perspectives towards different topics, events, or conversations. The essence of sentiment analysis lies in deciphering human emotions and attitudes, which holds immense value across diverse applications. One notable application involves imbuing computers with the ability to comprehend and respond to human non-verbal cues, including emotions. By discerning emotions expressed in conversations, machines can adapt and tailor experiences to suit individual preferences and needs.

Furthermore, sentiment analysis extends beyond textual content to encompass audio and video streams. In this project, sentiment analysis is conducted on both audio and video data streams through a series of intermediate processes. These processes encompass techniques such as speech recognition, lemmatization, tokenization, regular expression parsing, and neural network classification. For stored video streams, convolutional neural network (CNN) models are employed to extract sentiment insights.

The project's ultimate goal is to classify sentiments into a broader spectrum of emotional states, ranging across six distinct classes: happy, sad, angry, fear, surprise, and neutral. By leveraging multimedia sentiment analysis, this project aims to provide deeper insights into human emotional responses across various forms of communication, thereby enhancing the understanding and interaction bet ween humans and machines. challenges of extracting sentiment from these diverse sources, allowing for a more comprehensive understanding of public sentiment. Multimedia sentiment analysis goes beyond textual data and captures emotions and opinions.

II. LITERATURE SURVEY

[1] Anu J Nair , Aadithya Vinayak, Veena G , **"Comparative study of Twitter Sentiment On COVID - 19 Tweets"**, IEEE, 6, (2021). In this tweets related to COVID hashtag will be extracted and then sentiment analysis algorithms like Logistic Regression BERT and VADER will applied It is aspect based sentiment analysis. In this BERT is used which is more accurate than other algorithms because they look for the aspect of the sentence. The limitation of this existing system is it is domain based and we have not looked towards the mood of the user.

[2] Amalia Anjani Arifiyanti, Eka Dyar Wahyuni **"Emoji and Emoticon in Tweet Sentiment Classification"**, IEEE, 6, (Oct. 2020). In this study, they use three different scenarios regarding emoji and emoticon conversion. First converting emoji into its Unico de naming convention. Second converting emoji and emoticon into its sentiment category. Last step is text pre-processing is stemming. Different from emoticons, emojis not only show a representation of the human face but also many other things in the form of two-dimaensional pictographs. It is only limited for tweets only.

[3] Payal K. Punde, Rasika S. Wagh, **"A Survey Paper on Different Approaches for Sentiment Analysis"**, IEEE, 5, (March 2018). The abstract emphasizes the significance of sentiment analysis, its relevance in various domains, and the categorization of sentiment analysis, its relevance in various domains, and the categorization of sentiment analysis into document, sentence, and aspect levels, while also acknowledging the challenges and potential for future research in this area. The article highlights the importance of sentiment analysis in text data mining, provides a comprehensive overview of sentiment analysis approaches, and discusses the challenges and future research directions in the field. The paper lacks specific details on the methodologies and techniques used in sentiment analysis, making it less useful for researchers seeking in depth insights into the field.

[4] Gen Li, QiuSheng Zheng, Long Zhang, SuZhou Guo, LiYue Niu **"Sentiment Infomation based Model For Chinese text"**, IEEE, 6, (2020). The paper combining sentiment dictionary knowledge and deep learning. It employs a hybrid task learning approach for improved sentiment analysis, achieving better performance compared to other models on Chinese text datasets. The paper lacks specific details about model architecture and hyperparameters. Limited explanation of the emotional dictionary's construction.

[5] Namita Mittal, Divya Sharma, Manju Lata Joshi **"Image Sentiment Analysis using Deep Learning"**, IEEE, 4, (2018). The paper highlights the importance of image sentiment analysis in the context of social media and discusses the suitability and limitations of different deep learning models. CNN can achieve high accuracy with fewer features and reduced training time, making it an efficient choice. Fast R-CNN offers improved efficiency compared to R-CNN by reducing processing time and memory usage. The choice of the most suitable deep learning model depends on the specific dataset and problem, which can make model selection challenging.

[6] Shen Ao **"Sentiment Analysis Based on Financial Tweets and Market Information"**, IEEE, 6, (2018). It outlines the procedure for financial news sentiment analysis, which involves data mining, text processing, and sentiment analysis. The analysis helps in understanding how sentiments expressed in financial news correlate with stock market movements. This insight can be valuable for investors, traders, and regulators. Determining the sentiment of words and phrases in the financial domain can be challenging, as the sentiment of financial terms can vary widely and depend on context.

[7] Irene Irawaty, Rachmadita Andreswari, Dita Pramesti **"Vectorizer Comparison for Sentiment Analysis on Social Media Youtube: A Case Study"**, IEEE, 6, (September2020). The research employs machine learning algorithms and vectorization techniques to classify sentiments, highlighting the effectiveness of support vector machine with TFIDF vectorizer for sentiment analysis in the context of Nokia's products on YouTube. This research utilizes sentiment analysis on YouTube comments to gauge public sentiment towards Nokia's products, providing valuable insights for market research and product improvement. The study's focus on YouTube comments may not captur e sentiments expressed on other social media platforms, potentially limiting the comprehensiveness of the analysis.

[8] Yun Liang, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama **"Cross-domain Semi-supervised Deep Metric Learning For Image Sentiment Analysis"**, IEEE. It combines teacher-student models, multi-level feature extraction, and cross-domain training to improved sentiment classification. However, it requires careful hyperparameter tuning and is more complex compared to some existing methods. The use of multi-level deep metric learning allows for more effective sentiment analysis by capturing both shallow and deep-level image features. The proposed method involves multiple components, including teacher-student models, joint losses, and center loss, making it relatively complex to implement.

2.1 Summary Of Literature Survey

From survey of above these eight research papers cover various aspects of sentiment analysis in different domains and using diverse techniques. Some highlight the importance of sentiment analysis in areas like social media, finance, and multimedia, while others delve into the technical aspects of sentiment analysis methods. Notably, the use of deep learning models, such as CNN and Fast R-CNN for image sentiment analysis, is discussed. Challenges such as domain-specific sentiment analysis and model selection are acknowledged. Additionally, some papers introduce novel approaches like combining sentiment dictionaries with deep learning, which enhance sentiment analysis accuracy. While these papers provide valuable insights and contributions to the field, they also have limitations, such as lacking specific details on methodologies, model architecture, or hyperparameters. They are all based on specific single media sentiment analysis. Overall, these papers collectively contribute to the growing field of sentiment analysis, emphasizing its relevance in various applications and its potential for improving decision-making processes.

2.2 Limitations Of Existing System

- Most of the existing system are based on single media sentiment analysis. Most are based on text only.
- A single media source may have inherent bias or a specific agenda.
- Sentiment analysis on a single media item may not capture the full range of emotions and sentiments expressed by people. It may miss important nuances and details.
- Single media analysis may struggle to capture the emotional tone conveyed through non-verbal cues like facial expressions or tone of voice.
- Single media sentiment analysis may not be suitable for tasks that require a comprehensive understanding of emotions and sentiments, such as customer sentiment analysis for a brand with multimedia interactions.

III. PROBLEM STATEMENTS & OBJECTIVES

3.1 Problem Statements

The problem statement for sentiment analysis on single media focuses on analyzing and understanding the sentiment expressed within a single type of media, such as text, to determine whether it is positive, negative, or neutral. In contrast, sentiment analysis on multimedia aims to extend this analysis to diverse media formats, including text, images, audio, and video, in order to discern and interpret sentiments, emotions, and opinions expressed across various media channels, thus providing a more comprehensive and holistic view of public sentiment in the digital age in a single web app.

3.2 Objectives

The objective of conducting sentiment analysis on multimedia content is to leverage advanced artificial intelligence and natural language processing techniques to extract and analyze sentiments, emotions and opinions expressed in various forms of multimedia, including text, images, audio and video. Our objective aims to provide valuable insights into public perception, user feedback, and emotional responses related to a wide range of multimedia sources.

One key aspect of this objective is to develop and implement robust machine learning and deep learning models capable of accurately identifying and categorizing sentiments within multimedia data. This involves training models to recognize not only the sentiment polarity (positive, negative, neutral) but also more nuanced emotions such as happiness, anger, sadness, or excitement. By achieving this, the goal is to gain a deeper understanding of the emotional context within multimedia content.

IV. METHODOLOGY

4.1 Introduction

Our proposed system Multimedia sentiment analysis is the process of using various forms of media, such as text, images and videos to determine and analyze the emotional tone, sentiment or feeling expressed within the content. Our proposed system works by reviewing the existing literature on sentiment analysis, with a primary focus on deep learning-based multimodal sentiment analysis methods. It examines the major applications of sentiment analysis across various domains, such as finance, politics, health, tourism and more. The system reviews sentiment analysis methods for different modalities, including text, image, video and audio and discusses various levels of sentiment analysis. Additionally, it explores fusion strategies for combining information from different media to achieve better sentiment analysis result.

The system evaluates the challenges faced in the field and provides insights into possible solutions and future research directions. It also evaluates the limitations and complexities of existing models and methods and discusses potential future directions in the field of sentiment analysis.

- **Multimedia Data:** This is the raw multimedia data that needs sentiment analysis, which can include text, audio, visual content, and more.
- **Data Preprocessing:** Data preprocessing is essential to clean and prepare the data for analysis. It involves tasks like text cleaning, audio normalization, and image resizing.
- **Feature Extraction & Data Transformation:** This step extracts relevant features from the multimedia data. For text, it might involve tokenization and feature engineering. For audio, it may involve extracting acoustic features or converting speech to text. For visual data, it could involve extracting features using deep learning models.
- **Sentiment Analysis:** The data, now in a suitable format, goes through sentiment analysis modules tailored to the data type. This involves using Natural Language Processing (NLP) models for text, audio analysis models for audio, and computer vision modelsfor visual data.
- **Integrated Sentiment Analysis:** After individual sentiment analysis, results from different data types are integrated to provide a holistic sentiment analysis across all multimedia data. The integration step can include combining scores or considering the weight of each type of data.
- **Sentiment Classification:** The integrated sentiment analysis results are categorized into sentiment classes, typically "Positive," "Negative," and "Neutral." This classification simplifies the sentiment assessment for reporting and further analysis.
- **Output:** The final sentiment analysis results can be presented in a user-friendly format, such as sentiment scores (positive, negative, neutral), emotion categorization, or visualizations.

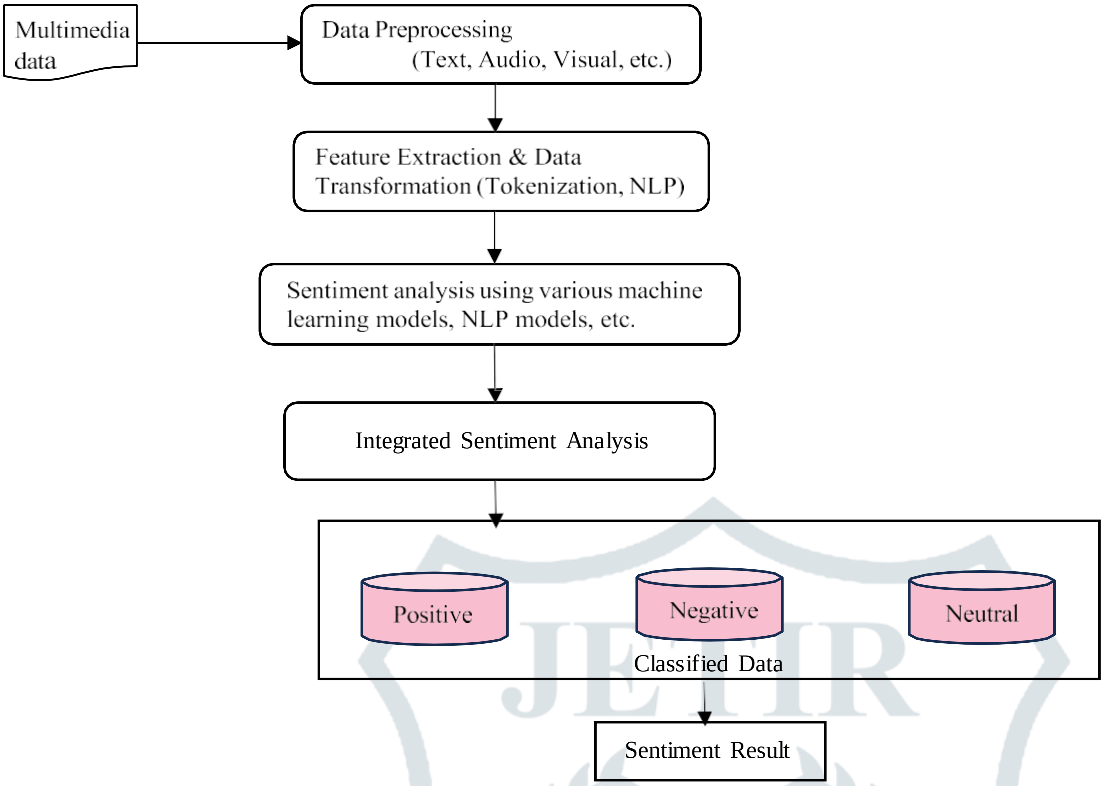

Fig 4.1.1 : Methodology of System

A. Text Sentiment Analysis

Initially, user provides input text in the provided text area. The input text is converted to lowercase to ensure consistency. Punctuation marks are removed from the text to simplify tokenization. Tokenization is the process of breaking down the text into individual words or tokens. Stopwords, which are common words that do not contribute much to the meaning of the text (e.g., "the", "is", "and"), are removed. A predefined dictionary containing words associated with emotions is used to detect emotions in the text. Each word in the tokenized text is checked against the dictionary to determine if it corresponds to an emotion.

NLTK's Vader sentiment analyzer is used to analyze the overall sentiment of the text. The sentiment analyzer provides scores for positive, negative, and neutral sentiments based on the text's content. The detected sentiment and emotions, if any, are displayed to the user.

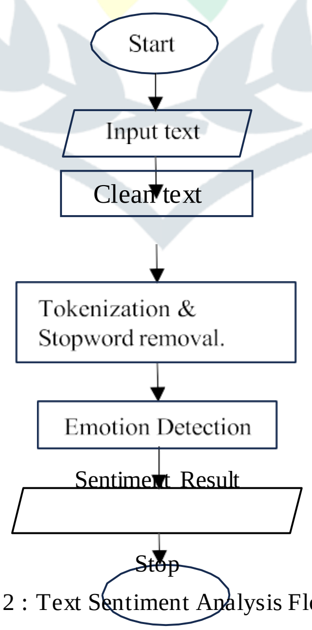

Fig 4.1.2 : Text Sentiment Analysis Flowchart

B. Image Sentiment Analysis

User uploads an image file (e.g., JPEG, PNG). The uploaded image is converted to grayscale to simplify processing. The image is resized to match the input size expected by the pretrained model. The pre-trained deep learning model analyzes the image to predict the dominant emotion depicted in the image. The model has been trained on a dataset of facial expressions to recognize emotions. The predicted emotion is displayed to the user. A pre-trained deep learning model is used for emotion recognition in images. The model is loaded from a JSON file (emotion_model.json) and corresponding weights are loaded from an HDF5 file (emotion_model.h5).

The emotion recognition model is likely based on Convolutional Neural Networks (CNNs), a type of deep learning architecture commonly used for image classification tasks.

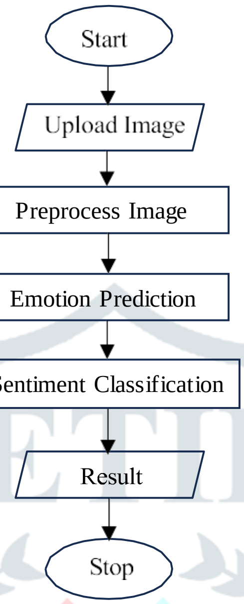

Fig 4.1.3 : Image Sentiment Analysis Flowchart

C. Video Sentiment Analysis

User initiates the webcam feed to capture real-time video. The webcam captures frames, and OpenCV's face cascade classifier is used to identify faces in each frame. For each detected face, the pre-trained model predicts the dominant emotion depicted by analyzing facial expressions. Detected emotions are overlaid onto the video feed, allowing the user to see the emotions associated with each identified face in real-time.

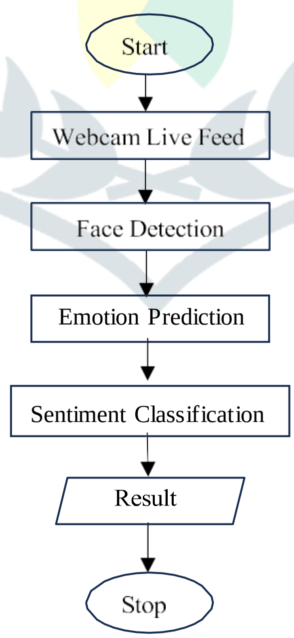

Fig 4.1.4 : Video Sentiment Analysis Flowchart

D. Audio Sentiment Analysis

User selects either "Upload Audio" or "Record Audio". User uploads an audio file in WAV format. Or User records audio using the microphone. Speech recognition software converts the audio input into text. The recognized text is converted to lowercase and punctuation marks are removed for consistency. Tokenization breaks down the text into individual words or tokens. Stopwords, common words that do not carry significant meaning, are removed from the text. Emotions are detected in the text using a predefined dictionary. NLTK's Vader sentiment analyzer evaluates the sentiment expressed in the text. The detected sentiment and emotions, if any, are displayed to the user.

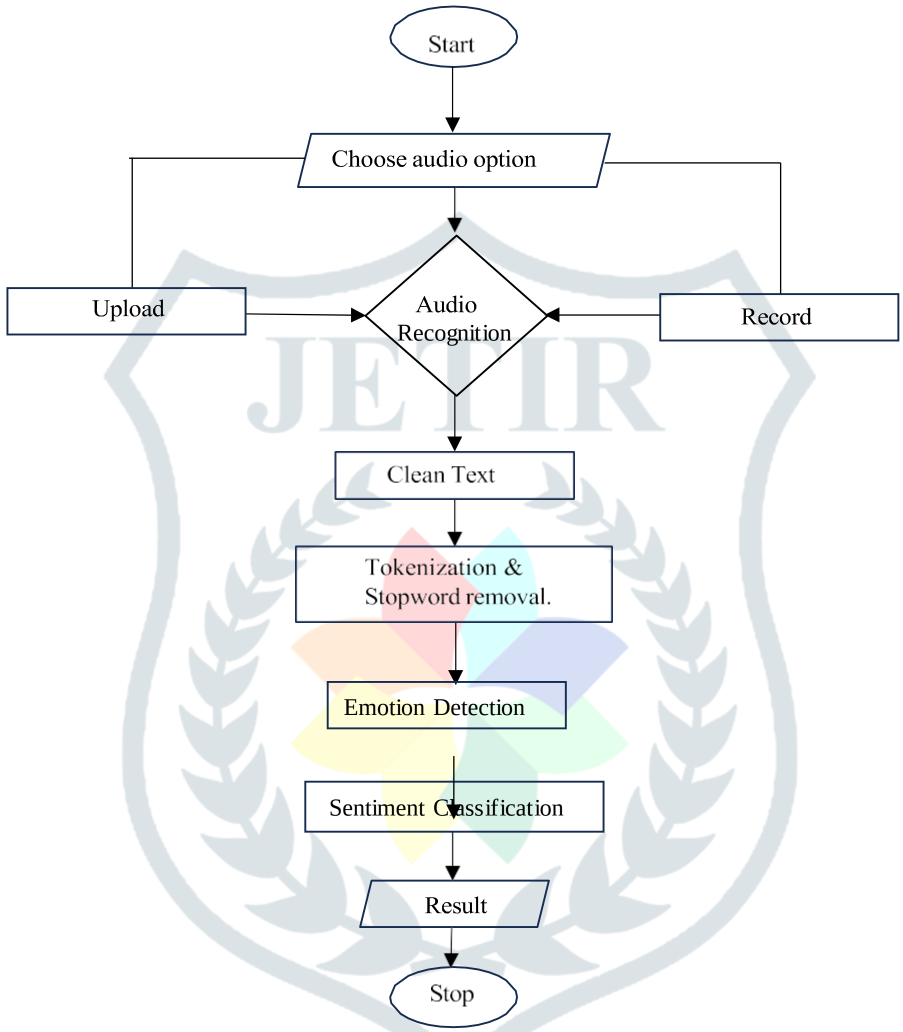

Fig 4.1.5 : Audio Sentiment Analysis Flowchart

V. EXPERIMENTAL SET UP

5.1 Details Of Selected DataSet

To find a dataset for Sentiment analysis on Kaggle, we visited the Kaggle website and search for relevant keywords like "Sentiment analysis" or "NLP". Browsed through the search results to find a dataset with features and labels that suit our needs. We made sure to check dataset's size, format, description, and licensing terms. We took dataset of tweet to perform sentiment analysis on that.

1. **Input Data:** The system employs various multimedia inputs, including text, images, audio, and video, for sentiment analysis.
2. **Data Preprocessing:** Text inputs undergo preprocessing, including lowercasing, punctuation removal, tokenization, and stopword removal.
3. **Data Split:** The dataset was divided into three parts: the training set to teach our model, the validation set to fine-tune its performance, and the testing set to evaluate its accuracy.
4. **Data Augmentation:** To enhance the model's robustness, we employed data augmentation techniques, creating variations of images through rotations, scaling, and other transformations.
5. **Database Management:** The system does not involve a database for this implementation.
6. **User Interaction:** Users can choose between text input, image upload, audio upload, or audio recording for sentiment analysis.
7. **External Libraries:** External libraries such as NLTK, OpenCV, TensorFlow, Streamlit, SpeechRecognition, and PyDub are utilizedfor text processing, image manipulation, audio processing, and user interface design

5.2 Performance Evaluation Parameters For Validation

In this project, performance evaluation is crucial for assessing the effectiveness of sentiment analysis across different multimedia inputs. The following parameters are considered:

1. **Accuracy:** Measures the correctness of sentiment predictions across text, image, audio, and video inputs.
2. **Precision:** Indicates the percentage of correctly identified sentiments compared to all positive predictions.
3. **Recall:** Determines the system's ability to capture all instances of sentiments present in the input data.

5.3 Accuracy & Loss Graph of Our Project

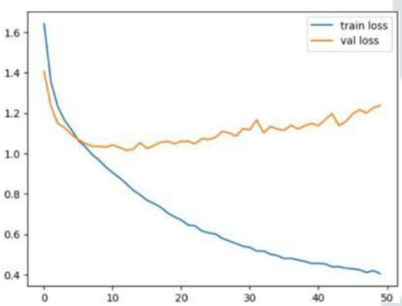

Fig. 5.3.1 Loss Graph

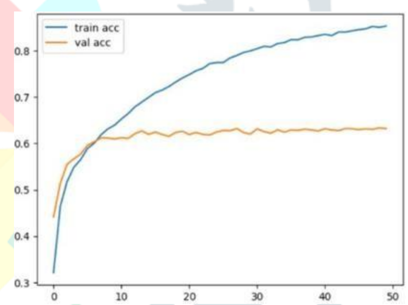

Fig. 5.3.2 Accuracy Graph

VI. SYSTEM REQUIREMENT

❖ **Software Requirements :-**

- Python
- Python Libraries like streamlit, io, tempfile.
- Natural Language Processing (NLP) libraries like nltk, nltk.corpus, nltk.sentiment.vader, nltk.tokenize and Audio processing libraries like speech_recognition(SpeechRecognition), pydub, AudioSegment(from pydub)
- Image processing libraries like tensorflow, Keras, cv2(OpenCV), numpy, PIL(Python Imaging Library)
- Video processing libraries like streamlit_webrtc, VideoTransformerBase(from streamlit_webrtc), tensorflow.keras.models.
- Dataset and data storage solutions (Kaggle)
- Various sentiment analysis algorithms and machine learning models, such as support vector machines (SVM)

❖ **Hardware Requirements :-**

- Monitor
- CPU
- RAM

VII. RESULTS

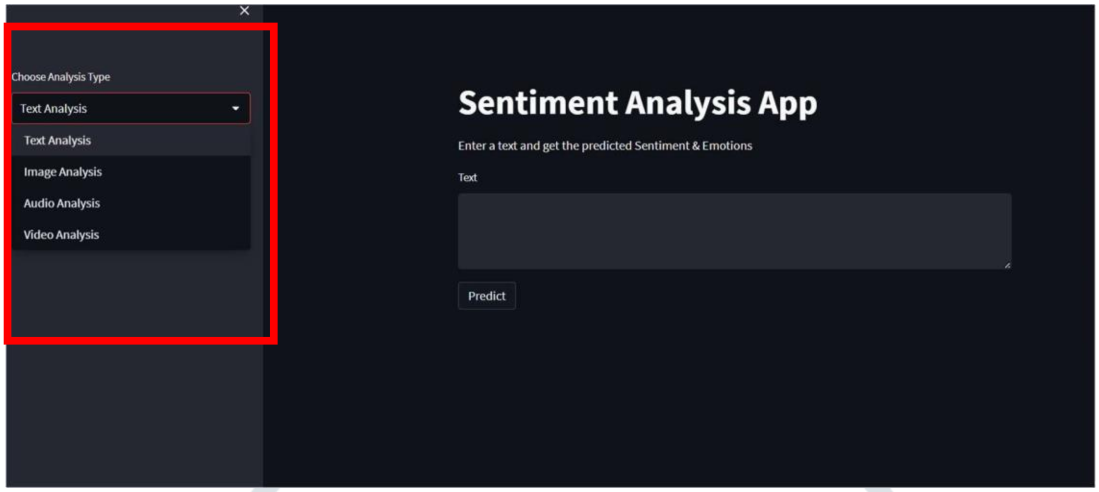

Fig. 6.1 Side bar of Sentiment Analysis

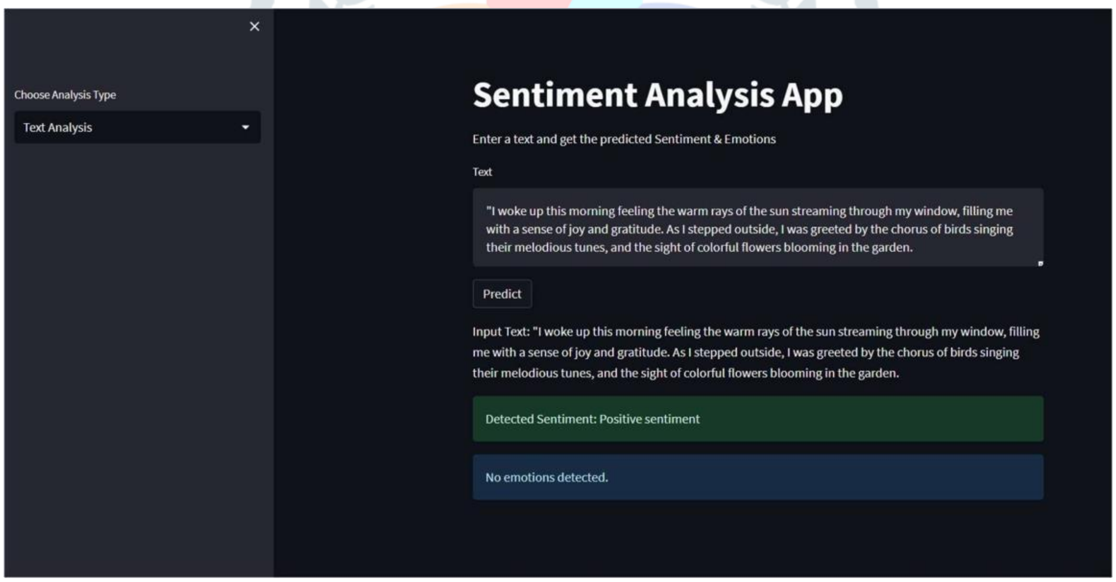

Fig. 6.2 Text Sentiment Analysis

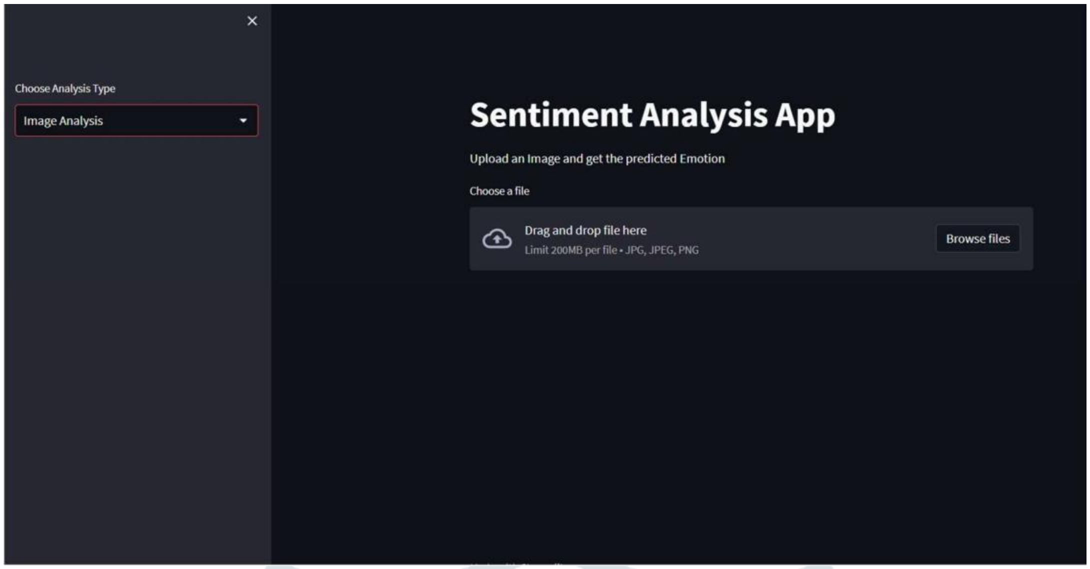

Fig. 6.3 (a) Image Sentiment Analysis

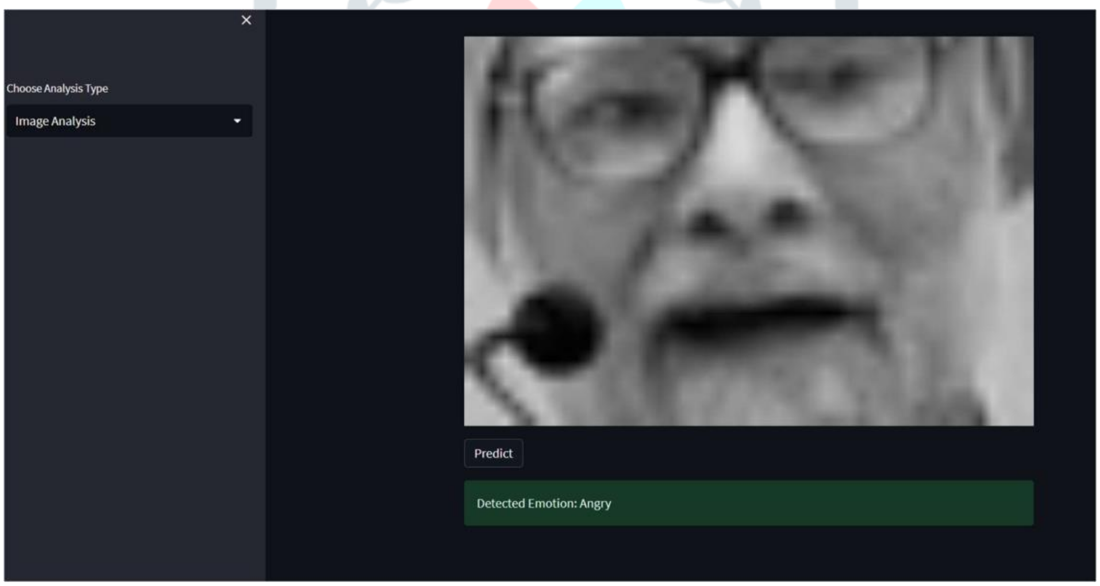

Fig. 6.3 (b) Image Sentiment Analysis

Fig. 6.4 Audio Sentiment Analysis

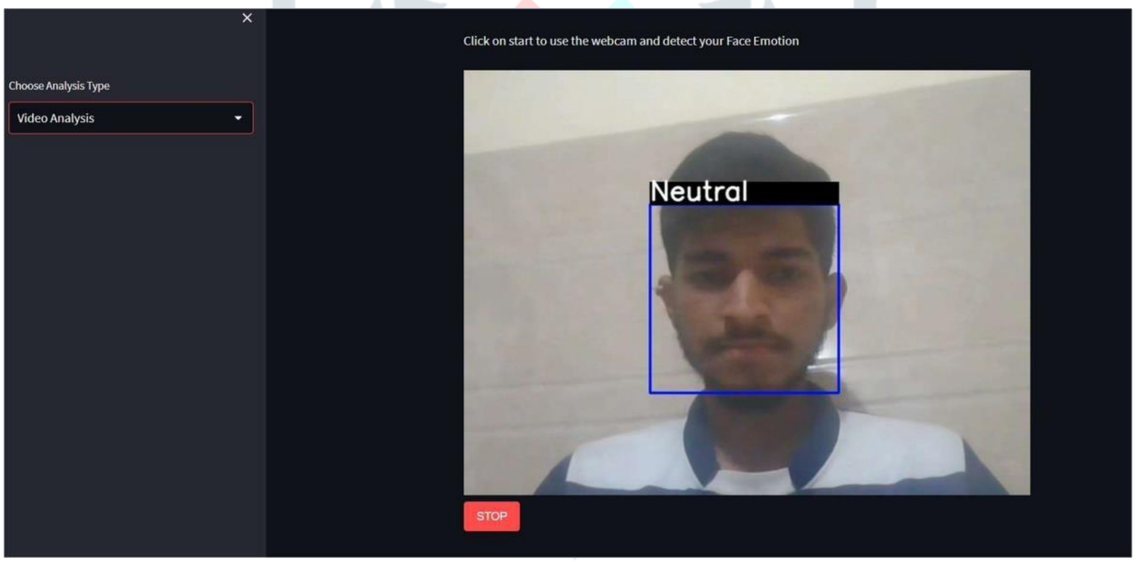

Fig. 6.5 Video Sentiment Analysis

VIII. CONCLUSION

In conclusion, multimedia sentiment analysis, which involves the analysis of sentiments expressed in various forms of media, including text, images, audio, and video, represents a powerful and versatile approach to understanding human emotions, opinions, and attitudes. The proposed system offers a comprehensive understanding of sentiment analysis, from its various levels to the challenges and advancements in the field. It underscores the growing importance of analyzing multimodal data in social networks and the role of deep learning in enhancing sentiment analysis accuracy. Researchers and practitioners can benefit from this survey to gain insight s into the current state of sentiment analysis and explore opportunities for further improvements.

IX. ACKNOWLEDGEMENT

We would like to thank our project guide Prof. Reena Deshmukh along with our Head of Department Dr. Uttara Gogate and our Principal Dr. Pramod R. Rodge for providing for providing us an opportunity of the project work in SSJCOE Dombivli and provid ing us all support and guidance which made us to pursue the project duly. We are extremely thankful to them for providing such a nice support and guidance. We are thankful to and fortunate enough to get constant encouragement, support and guidance from all Teaching s taffs of Computer Engineering which helped us in pursuing our project work.

REFERENCES

[1] Anu J Nair , Aadithya Vinayak, Veena G , **"Comparative study of Twitter Sentiment On COVID - 19 Tweets"**, IEEE, 6, (2021).

[2] Amalia Anjani Arifiyanti, Eka Dyar Wahyuni **"Emoji and Emoticon in Tweet Sentiment Classification"**, IEEE, 6, (Oct. 2020).

[3] Payal K. Punde, Rasika S. Wagh, **"A Survey Paper on Different Approaches for Sentiment Analysis"**, IEEE, 5, (March 2018).

[4] Gen Li, QiuSheng Zheng, Long Zhang, SuZhou Guo, LiYue Niu **"Sentiment Infomation based Model For Chinese text"**, IEEE, 6, (2020).

[5] Namita Mittal, Divya Sharma, Manju Lata Joshi **"Image Sentiment Analysis using Deep Learning"**, IEEE, 4, (2018).

[6] Shen Ao **"Sentiment Analysis Based on Financial Tweets and Market Information"**, IEEE, 6, (2018).

[7] Irene Irawaty, Rachmadita Andreswari, Dita Pramesti **"Vectorizer Comparison for Sentiment Analysis on Social Media Youtube:A Case Study"**, IEEE, 6, (September2020).

[8] Yun Liang, Keisuke Maeda, Takahiro Ogawa, Miki Haseyama **"Cross-domain Semi-supervised Deep Metric Learning For Image Sentiment Analysis"**, IEEE