# FUSION OF MULTI MODAL DATA FOR EFFECTIVE OPINION MINING

Yoshitha Puvvala
*Department of Computer Science and Engineering*
*Institute of Aeronautical Engineering*
Hyderabad, India

Dr. M Lakshmi Prasad
*Department of Computer Science and Engineering*
*Institute of Aeronautical Engineering*
Hyderabad, India

*Abstract*—***Opinion mining is crucial for recognizing and analysing opinions of users in a variety of sectors. The goal of this research is to fuse multimodal data in order to get more accurate and complete opinion mining. The code created for this project provides the user with six different choices from which to select the required functionality.***

***Keywords—CNNs, Image-Processing, Image Detection, Speech Recognition, Natural Language Processing***

## I. INTRODUCTION

Sentiment analysis, also known as opinion mining, is crucial in natural language processing for assessing public sentiment towards various subjects. It involves analyzing opinions from sources like blogs, comments, reviews, and tweets. These insights are valuable in marketing for evaluating advertising campaigns, new product launches, and understanding demographic preferences.

However, sentiment analysis faces challenges. Interpretation of opinion words can vary based on context, and human expression of opinions is nuanced. Traditional text processing struggles to capture subtle differences in sentiment, especially when individuals express contradictory sentiments within the same text.

## II. STUDY EXISTING METHODOLOGIES

Our work builds upon previous studies on opinion mining, particularly in the area of sentiment analysis and emotion detection. The existing system comprises individual code modules that are designed to address specific functionalities within sentiment analysis. These modules are developed separately to cater to different aspects of the project.

### *A. PRIOR WORK*

In recent years, opinion mining and sentiment analysis have become pivotal areas of research. Li et al. [1] conducted a comprehensive survey on opinion mining, exploring its evolution from words to multiple modalities. Cambria and Hussain [2] provided foundational insights into opinion mining and sentiment analysis, laying the groundwork for subsequent research in the field [3]. Liu [4] further delved into sentiment analysis and opinion mining, offering in-depth analysis and methodologies.

Chui et al. [5] investigated the potential of machines to replace humans in various domains, shedding light on the evolving landscape of human-machine interaction. Kim and Hovy [6] contributed to determining the sentiment of opinions, presenting methodologies for sentiment analysis. Additionally, Cambria [7] explored affective computing and sentiment analysis, emphasizing the role of emotions in understanding textual data.

Turchi [8] introduced a cross-lingual approach to multilingual sentiment analysis, addressing challenges in sentiment analysis across different languages. Poria et al. [9] proposed aspect extraction for opinion mining using deep convolutional neural networks, advancing the field with innovative methodologies. Wang et al. [10] focused on exploiting sentiment information from structured reviews for online recommendation systems, leveraging sentiment analysis techniques for personalized recommendations.

Moreover, Balahur [11] contributed methods and resources for sentiment analysis in multilingual documents, enhancing the applicability of sentiment analysis across diverse linguistic contexts. Niles and Pease [12] linked lexicons and ontologies, providing valuable insights into semantic relationships for opinion mining tasks. Strapparava and Valitutti [13] extended WordNet to include affective dimensions, enriching lexical resources for sentiment analysis tasks.

Esuli and Sebastiani [14] developed SENTIWORDNET as a publicly available lexical resource for opinion mining, facilitating sentiment analysis research. Building upon this work, Baccianella et al. [15] introduced SENTIWORDNET 3.0, an enhanced lexical resource for sentiment analysis and opinion mining, further advancing the state-of-the-art in the field.

### *B. DISADVANTAGES OF EXISTING SYSTEM*

- **Lack of Integration:** One of the main disadvantages of the existing system is the lack of integration between the different code modules. Each module operates independently, requiring the user to run them separately to access specific functionalities. This lack of integration can result in a disjointed user experience and make it challenging to seamlessly analyse opinions using multiple modalities.

- **Limited Modality Coverage:** The existing system primarily focuses on text, image, facial, speech, and audio modalities for opinion mining. While these modalities cover a wide range of data sources, there may be other modalities, such as video or sensor data, that are not adequately addressed. The limited modality coverage restricts the system's ability to capture opinions expressed through alternative data sources.

- **Sensitivity to Data Quality:** Opinion mining heavily relies on the quality and diversity of the training data used for sentiment analysis and emotion detection.

If the training data is biased, limited, or not representative of the target user group, the system's performance may suffer. Ensuring high-quality training data can be challenging, as it requires extensive data collection and curation efforts.

- **Difficulty in Handling Subjectivity:** Opinion mining involves analysing subjective information, including sentiments and emotions. Subjectivity can vary significantly across different individuals and cultures, making it challenging to develop universal models that accurately capture the diversity of opinions. The existing system may struggle to handle subjectivity effectively and may not provide nuanced analysis in certain contexts.
- **Computational Complexity:** Analysing opinions from multiple modalities, such as text, image, facial expressions, speech, and audio, can be computationally intensive. Processing and analysing data from various modalities in real-time or large-scale scenarios can pose challenges in terms of computational resources and time efficiency. The existing system may face limitations in scalability and real-time processing. Ethical Considerations: Opinion mining systems must address ethical considerations, such as privacy, data protection, and potential biases. Processing and analysing user opinions require handling sensitive user data, and ensuring its privacy and security is crucial. Additionally, the existing system should be designed to mitigate biases that may arise from the training data or the sentiment analysis algorithms.

- **User Dependence:** The existing system relies on user input and active participation to provide data for sentiment analysis. Users need to upload datasets, images, or audio recordings for analysis. This user dependence may limit the system's usability and effectiveness, as it relies on user engagement and cooperation

## III. Problem Description

The existing opinion mining techniques primarily focus on text analysis, neglecting other important modalities such as images, speech, and facial expressions. This approach may lead to incomplete and less accurate results. Additionally, users may have different preferences and requirements for analyzing opinions, requiring a flexible and customizable solution.

The main problem addressed by this research is the limited scope of traditional opinion mining techniques, which often rely solely on textual data and fail to consider other modalities. This leads to a less comprehensive and accurate analysis of user opinions.

## IV. Design and Development of Proposed System

### *A. Development of Proposed System*

The significant upgrades to the existing system by consolidating all the individual code modules into a single code base. This integration allows for a more streamlined and efficient approach to sentiment analysis and opinion mining. The upgraded system includes several new features and enhancements to improve user experience and provide more comprehensive analysis:

**Multi-language Support:** The system now supports multiple languages, enabling the analysis of emotions expressed in different languages in spoken content. This enhances the system's capability to understand and analyse opinions across diverse linguistic contexts.

**Custom Dataset Upload:** Users now have the ability to upload their own feedback datasets for analysis. The code analyses the data and provides insights into the associated emotions expressed in the feedback.

This customization feature allows users to gain valuable insights into the emotional tone of the feedback they receive.

**Feedback and Rating System:** A feedback and rating system has been implemented to allow users to provide comments and ratings on the system's performance. This feedback is invaluable for system enhancement and improvement, as users can contribute their suggestions and help shape the system's future development.

Users can select the specific approaches or options that best suit their requirements, allowing for a personalized and tailored experience.

The upgraded system provides a powerful platform for analysing opinions across different modalities and languages, empowering users to gain deeper insights into sentiments and emotions expressed in various forms of data.

**Face Detection:**

**Haarcascade** refers to a pre-trained classifier used for detecting frontal faces in images or video frames. It is part of the OpenCV library, a popular open-source computer vision and machine learning software library. This classifier is based on the Haar feature-based cascade classifiers proposed by Paul Viola and Michael Jones in their 2001 paper, "Rapid Object Detection using a Boosted Cascade of Simple Features".

**Facial Emotion Detection:**

An emotion classification CNN using Keras involves building a Convolutional Neural Network (CNN) using the Keras library in Python. This CNN is trained on a dataset of labeled images to classify emotions such as happiness, sadness, and anger. After training, the model can predict the emotion depicted in unseen images. Deployment involves integrating the trained model into applications for real-time emotion classification tasks.

### B. Description of the Techiques used in Proposed System

In multi-modal opinion mining, fusion techniques play a sentiment analysis. Early fusion concatenates feature vectors from different modalities into a single representation, allowing the classifier to learn from the combined information. Late fusion combines decision outputs from individual modality-specific classifiers, leveraging the strengths of each modality independently before making a final decision.

Decision-level fusion aggregates the decisions of individual classifiers using techniques such as voting or weighting schemes, aiming to enhance the robustness and reliability of the overall decision-making process. These fusion techniques offer various approaches to integrating information from multiple modalities, enabling more comprehensive and accurate analysis of opinions expressed through different data sources. crucial role in integrating information from diverse modalities to improve the overall accuracy and reliability of

## V. System Design

The architecture of the fusion of multimodal data for effective opinion mining involves multiple components and stages. It typically includes data acquisition from various modalities such as text, images, facial expressions, speech, and audio.

These modalities are processed and feature extraction techniques are applied to extract relevant information. The extracted features are then fused using appropriate fusion methods to capture the combined representation of user opinions.

Finally, classification algorithms are applied to classify the fused features and determine the sentiment or emotions expressed by users. The architecture ensures the integration of different modalities to provide a comprehensive and accurate analysis of user opinions.

System models play a crucial role in the design phase of the fusion of multimodal data for effective opinion mining. These models provide visual representations of the system's structure, behaviour, and data flow, helping to understand and communicate the design effectively. The following system models are utilized in this phase:

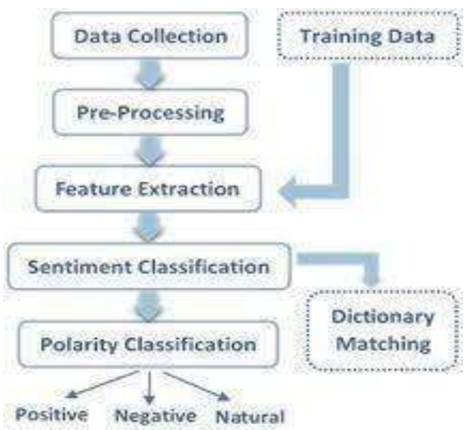



Fig. 1. Architecture.

**Architectural model:** The architectural model provides an overview of the system's structure and the interaction between its various components. It defines the high-level design of the system, including the modules, their dependencies, and the flow of data between them.

The architectural model ensures that the system is organized and scalable, allowing for seamless integration and communication between different modalities.

The architecture of the fusion of multimodal data for effective opinion mining involves multiple components and stages. It typically includes data acquisition from various modalities such as text, images, facial expressions, speech, and audio. These modalities are processed and feature extraction techniques are applied to extract relevant information.

The extracted features are then fused using appropriate fusion methods to capture the combined representation of user opinions. Finally, classification algorithms are applied to classify the fused features and determine the sentiment or emotions expressed by users.

The architecture ensures the integration of different modalities to provide a comprehensive and accurate analysis of user opinions

The proposed system offers the following key features:

Multimodal Data Fusion: The system combines multiple modalities, including text, image, facial expressions, speech, and audio, to perform effective opinion mining. By leveraging the strengths of each modality, the system can provide more accurate and comprehensive analysis of user opinions.

**Text Emotion Analysis:** The system employs sentiment analysis techniques to determine the emotional tone of textual content. It can classify text as positive or negative and identify the emotions associated with the text.

**Custom Dataset Analysis:** Users have the flexibility to upload their own datasets that contain feedback. The system analyses the feedback and displays the emotions associated with each feedback entry. This feature enables users to gain insights into the emotional tone of the feedback they receive.

**Image Emotion Detection:** Using computer vision techniques, the system can evaluate emotions expressed in images. It can recognize emotions such as happiness, sadness, fear, and more. Users can upload images to obtain an analysis of the emotional content within the image.

**Real-time Facial Emotion Detection:** The system activates the webcam, allowing users to assess their current emotional state in real-time. Facial recognition algorithms analyse facial expressions and detect emotions. This feature provides users with immediate feedback on their emotional expressions.

**Speech-to-Text Emotion Analysis:** The system supports speech-to-text conversion in multiple languages, including Telugu, Hindi, and English. It can transcribe spoken words into text and analyse the emotional content of the speech.

**Audio-to-Text Emotion Recognition:** The system takes audio input and converts it to text, enabling emotional analysis of spoken information. This feature enhances the system's ability to record and evaluate user opinions

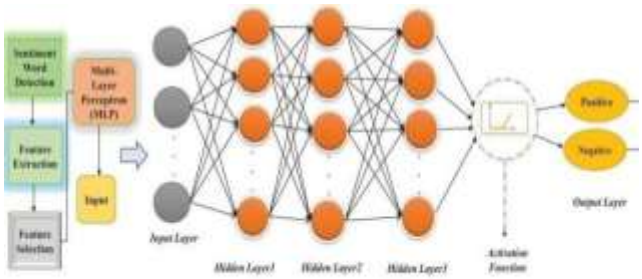


Fig. 2. Feature Extraction

Feature extraction and classification are crucial steps in multimodal data fusion for effective opinion mining. These processes involve extracting informative features from different modalities, combining them, and applying classification algorithms to classify user sentiments. Features can include word frequencies, visual descriptors, facial expressions, acoustic characteristics, and more.

Fusion techniques can be applied to combine the features at different levels. Classification algorithms such as SVM, decision trees, or neural networks are used to classify opinions based on the fused features. These steps enable a comprehensive analysis of user sentiments expressed through multiple modalities.

In the context of speech to text and audio to text, initially it transcribes the audio or speech to text using the SpeechRecognition library based on the specified language. Then, sentiment analysis is conducted using TextBlob to determine the emotional tone of the transcribed text.

Users can select the specific approaches or options that best suit their requirements, allowing for a personalized and tailored experience. The upgraded system provides a powerful platform for analyzing opinions across different modalities and languages, empowering users to gain deeper insights into sentiments and emotions expressed in various forms of data.

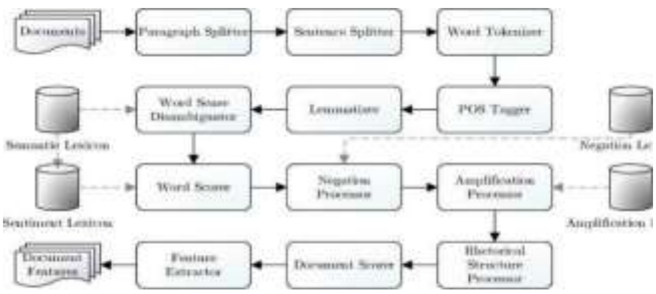


Fig. 3. Classification

## APPENDIX A:

| **Data Modality** | **Preprocessing Steps** |
|---|---|
| Text | Tokenization, stop word removal, Stremming / Lemmatization |
| Image | Landmarkimage resizing, feature extraction (eg: CNN, SIFT, Surf). |
| Facial | Face detection, Facial detection, Emotion recognition |
| Speech | Speech-to-text conversion, Audio feature extraction (eg: mfcc). |
| Audio | Audio feature extraction (eg: mfcc, chroma feature). |

Table A.1: Preprocessing Steps for Multi-Modal Data Fusion

## APPENDIX B:

| **Data Modality** | **Feature Extraction Method** |
|---|---|
| Text | Bag-of-words, tf-idf, word Embeddings (eg: word2vec, glove). |
| Image | Convolutional neural Network(CNN), local binary Pattern |
| Facial | Facial Landmark-based Features, Histrogram of Oriented gradients |
| Speech | Mel-frequency cepstral Coefficients(MFCC), prosody Features |
| Audio | MFCC chroma Features, spectral contrast. |

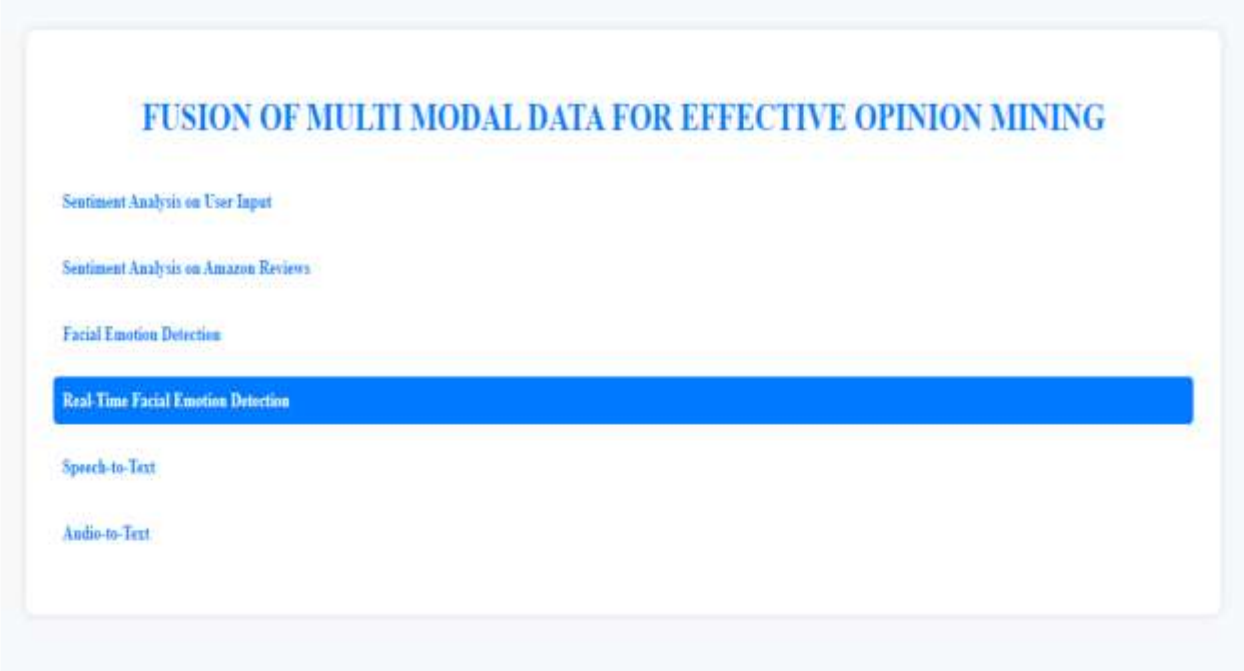
Table B.1: Fusion Feature Extraction Methods for Multi-Modal Data Fusion

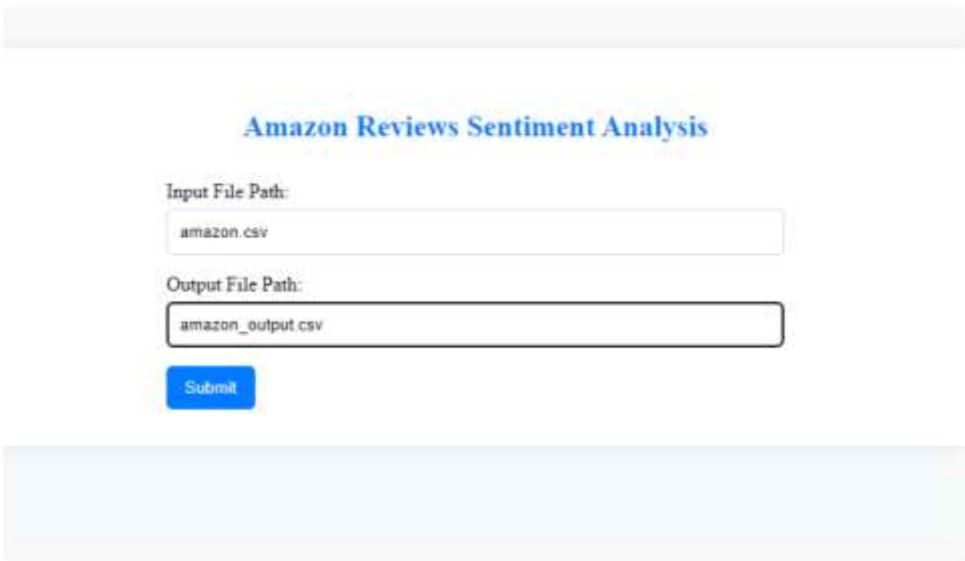
## APPENDIX C:

| Fusion Technique | Description |
|---|---|
| Early fusion | Concatenating feature vector from different Modalities into a single Representation |
| Late fusion | Combining the decision Outputs from individual Modality-specific classifiers |
| Decision-level fusion | Aggregation the decision of individual classifiers using Voting or weighting schemes |

Table C.1: Fusion Algorithms for Multi-Modal Data

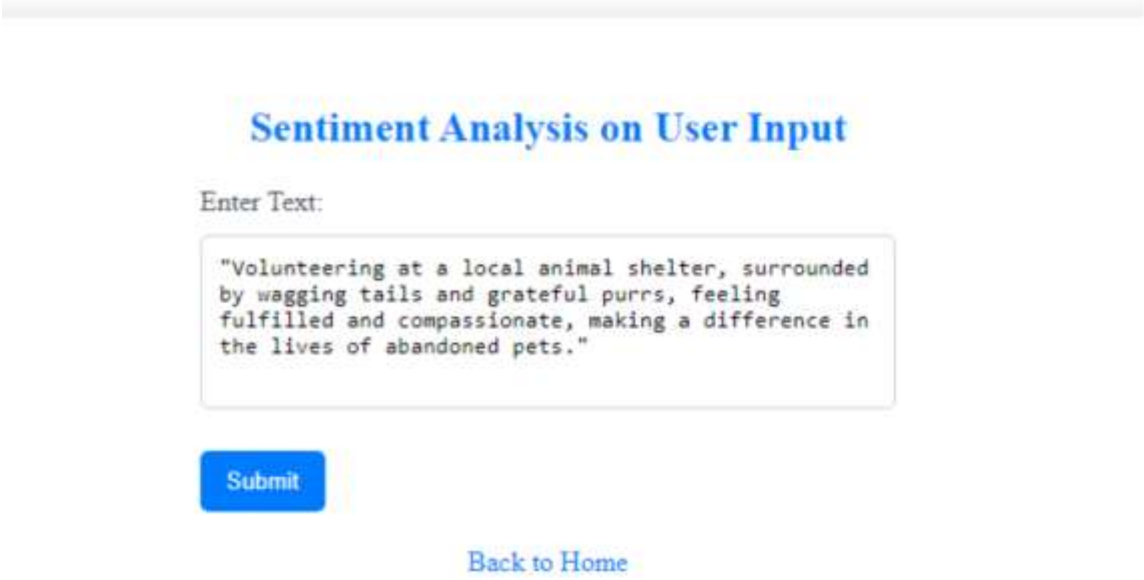
## VI. RESULT



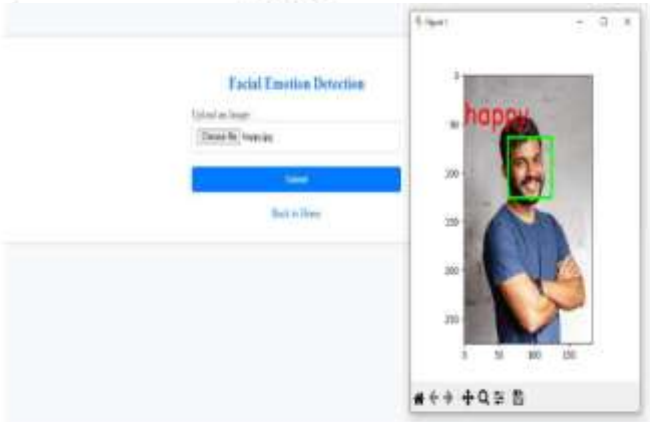

| | | |
|---|---|---|
| 1 | text | sentiment |
| 2 | reviews.text | neutral |
| 3 | I initially had trouble deciding between the paperwhite and the voyage because reviews more or less said the same thing: the paperwhite | positive |
| 4 | Allow me to preface this with a little history. I am (was) a casual reader who owned a Nook Simple Touch from 2011. I've read the Harry Pc | positive |
| 5 | I am enjoying it so far. Great for reading. Had the original Fire since 2012. The Fire used to make my eyes hurt if I read too long. Haven't e | positive |
| 6 | I bought one of the first Paperwhites and have been very pleased with it its been a constant companion and I suppose Ive read, on average | positive |
| 7 | I have to say upfront - I don't like coroporate, hermetically closed stuff like anything by Apple or in this case, Amazon. I like having devices | positive |
| 8 | My previous kindle was a DX, this is my second kindle in years. Love the form factor and all but I do miss the physical buttons for page tun | positive |
| 9 | Allow me to preface this with a little history. I am (was) a casual reader who owned a Nook Simple Touch from 2011. I've read the Harry Pc | positive |
| 10 | Just got mine right now. Looks the same as the previous generation except for the Kindle logo (it's black this time), feels a little heavier, th | positive |
| 11 | I initially had trouble deciding between the paperwhite and the voyage because reviews more or less said the same thing: the paperwhite | positive |
| 12 | I am enjoying it so far. Great for reading. Had the original Fire since 2012. The Fire used to make my eyes hurt if I read too long. Haven't e | positive |
| 13 | As reviewed by the wife This is the perfect thing for a new mommy who loves to read books! As soon as I had my baby girl, I had to stop re | positive |
| 14 | My new Kindle Paperwhite came from the USA to a small town in Sweden in just five days. Payment of tax and customs handled by Amazo | positive |
| 15 | I am enjoying it so far. Great for reading. Had the original Fire since 2012. The Fire used to make my eyes hurt if I read too long. Haven't e | positive |
| 16 | Had older model, that you could text to speech, this one hasn't. Liked the smaller size, but having to buy a different cover! Still getting use | positive |
| 17 | This is a review of the Kindle Paperwhite launched July 2015. Essentially. the same as the previous Kindle Paperwhite but with a fantastic | positive |
| 18 | I love my kindle! I got one for my fiance on his birthday, he loved it!!!!!!! so I decided to get one for myself as well and I am glad I did. It is | positive |
| 19 | Vraiment bon petit appareil , lger et facile d emploiJ ai hte de m en servir sur les plages cet hiverBelle bibliothque de livres disponiblesBor | neutral |
| 20 | Exactly what it is supposed to be. Works great and I love the built-in light. Perfect reader, and very quick delivery. | positive |
| 21 | Trs heureux que les livres soient sur Icloud. Aprs m'tre fait voler, c'est bien de pouvoir retrouver tous mes livres avec toutes les notes que | neutral |

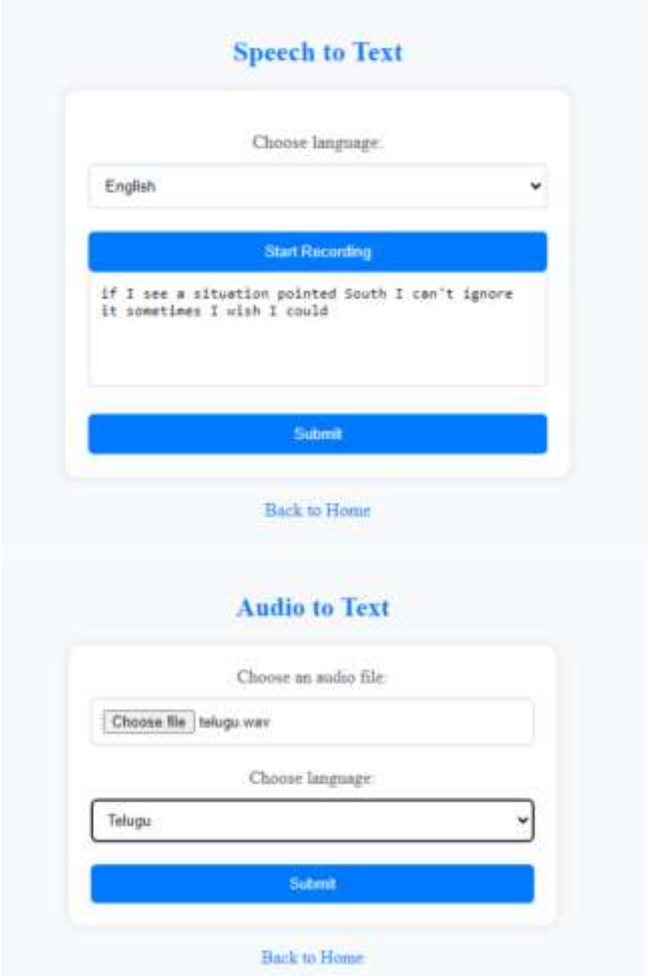
- Upload dataset with textual feedback




| Sentences | Expected Output | Actual Output | R |
|---|---|---|---|
| "Walking along the beach at sunset, feeling the warm sand beneath my feet, watching the waves crash gently, fills me with tranquility and contentment." | Positive | Positive | T |
| "Stuck in traffic on a rainy Monday morning, late for work again, frustration building with every passing minute, wishing for a teleportation device, feeling hopeless." | Negative | Negative | T |
| "Exploring a new city, getting lost in its winding streets, stumbling upon hidden cafes and boutiques, meeting friendly locals, fills me with excitement and wonder." | Positive | Positive | T |
| "Driving along a familiar road, radio softly playing in the background, mind wandering, neither focused nor distracted, just going through the motions." | Neutral | Positive | F |
| "Sitting alone in a quiet room, memories flooding back, of happier times long gone, feeling a deep sense of longing and sadness, wishing things were different." | Negative | Negative | T |
| "Surrounded by loved ones on a cozy winter evening, sharing stories and laughter, sipping hot cocoa by the fireplace, feeling grateful for these precious moments together." | Positive | Positive | T |
| "Attending a crowded concert, music pulsating through the air, energy coursing through my veins, dancing with strangers, feeling alive and exhilarated." | Positive | Positive | T |
| "Wandering aimlessly through the park, watching children play, feeling the breeze on my skin, thoughts drifting, neither happy nor sad." | Neutral | Positive | F |
| "Suffering from a terrible flu, body aching, head pounding, unable to get out of bed, feeling utterly miserable and defeated." | Negative | Negative | T |
| "Hiking through lush green mountains, breathing in the crisp, fresh air, surrounded by breathtaking views, feeling a sense of awe and appreciation for nature's beauty." | Positive | Positive | T |
| "Enduring a long and tedious meeting, listening to endless presentations, checking the clock every few minutes, feeling bored and impatient for it to end." | Negative | Negative | T |
| "Sitting in a crowded café, sipping on lukewarm coffee, surrounded by chatter, neither energized nor drained, simply existing in the moment." | Neutral | Neutral | T |
| "Cooking a delicious meal from scratch, the aroma filling the kitchen, the sound of sizzling ingredients, tasting the first bite, feeling proud and accomplished." | Positive | Positive | T |
| "Basking in the warm sunlight on a spring afternoon, feeling rejuvenated and alive, surrounded by blooming flowers and chirping birds." | Positive | Positive | T |
| "Struggling to meet deadlines at work, feeling overwhelmed and stressed, with tasks piling up and little time for rest." | Negative | Negative | T |
| "Taking a leisurely walk in the park, breathing in the crisp autumn air, feeling peaceful and content, admiring the vibrant colors of the leaves." | Positive | Positive | T |
| "Walking through the city streets, observing the hustle and bustle, neither overwhelmed nor underwhelmed, just taking in the sights." | Neutral | Neutral | T |
| "D[illegible] unexpected rain." | Negative | Negative | T |
| "Volunteering at a local animal shelter, surrounded by wagging tails and grateful purrs, feeling fulfilled and compassionate, making a difference in the lives of abandoned pets." | Positive | Neutral | F |

❖ Sample test cases for sentiment analysis on user input




## VII. Conclusion and Future Scope

### Conclusion

In conclusion, the fusion of multimodal data for effective opinion mining has shown promising results in capturing and analysing user opinions across various modalities, including text, image, facial expressions, speech, and audio. The proposed system combines different approaches and techniques to provide a comprehensive framework for sentiment analysis and emotion detection.

By integrating multiple modalities, the system can capture a more accurate and nuanced understanding of user sentiments, leading to improved decision-making and enhanced user experiences.

However, there are still several areas for future enhancements and research. Firstly, the system can be further improved by incorporating latest ML algorithms and NLP techniques. This can enhance the system's ability to extract and interpret emotions from textual content, improving the accuracy of sentiment analysis.

Additionally, the system can benefit from expanding language support to include more languages, enabling a wider range of users to analyze opinions in their native languages. This would require developing language-specific models and datasets for sentiment analysis and emotion detection.

### Future Scope

The fusion of multimodal data for effective opinion mining presents numerous opportunities for future enhancements and advancements. Here are some potential areas for further development:

- Integration of Advanced Machine Learning Techniques
- Contextual Analysis
- Sentiment Analysis in Social Media
- Multilingual Opinion Mining
- Enhanced Visualization and Reporting

## Acknowledgment

I am grateful for the infrastructure we have provided for the project work, as well as for Sri. M Rajashekar Reddy, Chair of the University. I thank Dr. Madhusudhana Rao, Professor and Head of the Department of Computer Science and Engineering, for extending his support for this project work and I would like to thank the beloved Principal, Dr. L. V Narasimha Prasad for having been a source of great information for our work. I would like to express my gratitude to my internal supervisor, Dr. M. Lakshmi Prasad, Professor, Computer Science and Engineering Department , for their internal support and their professionalism in helping me achieve a successful project.

## References

[1] Li, S., Wang, D., Zhang, J., & Wang, H. (2019). A Survey on Opinion Mining: From Words to Multiple Modalities. IEEE Transactions on Knowledge and Data Engineering, 31(10), 1923- 1940.

[2] Cambria, E., & Hussain, A. (2012). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 5(1-2), 1-135.

[3] Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.

[4] Liu, B. (2012). Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers.

[5] Chui, M., Manyika, J., & Miremadi, M. (2016). Where machines could replace humans— and where they can't (yet). McKinsey Quarterly.

[6] Kim, S. M., & Hovy, E. (2004). Determining the sentiment of opinions. In Proceedings of the 20th international conference on Computational Linguistics (COLING'04), 1367-1373.

[7] Cambria, E. (2016). Affective computing and sentiment analysis. IEEE Intelligent Systems, 31(2), 102-107.

[8] Turchi, M. (2015). Multilingual sentiment analysis using machine translation: A cross-lingual approach. Information Processing & Management, 51(5), 646-661.

[9] Poria, S., Cambria, E., & Gelbukh, A. (2016). Aspect extraction for opinion mining with a deep convolutional neural network. Knowledge-Based Systems, 108, 42-49.

[10] Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2010). Exploiting sentiment information from structured reviews for online recommendation. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'10), 1277-1286.

[11] A. Balahur, Methods and Resources for Sentiment Analysis in Multilingual Documents of Different Text Types, PhD Thesis, University of Alicante, Spain, 2011, 273 pages.

[12] I. Niles, A. Pease, Linking Lexicons and ontologies:mapping WordNet to the Suggested Upper Merged Ontology, Proceedings of the 2003 International Conference on Information and Knowledge Engineering (IKE 03), Las Vegas, 2003, pp. 23–26.

[13] C. Strapparava, A. Valitutti, WordNet-Affect: an affective extension of WordNet, Proceedings of LREC, vol. 4, 2004, pp. 1083–1086.

[14] A. Esuli, F. Sebastiani, SENTIWORDNET: a publicly available lexical resource for opinion mining, Proceedings of the 5th Conference on Language Resources and Evaluation LREC-06, Genoa, Italy, 2006, pp. 417–422, See also: http://sentiwordnet.isti.cnr.it/.

[15] S. Baccianella, A. Esuli, F. Sebastiani, SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining, Proceedings of LREC-10, Malta, 2010, pp. 2200–2204.