Multimodal Sentiment Analysis: Exploiting Other Modalities for Improved Results

Vatsal Kansara[a], Madhuri Chopade[a]

[a]*Gandhinager Instiitute of Technology, Khatraj – Kalol Road, Moti Bhoyan, Gandhinagar-382721, India*

Abstract

An emerging field of NLP multimodal sentiment analysis (MSA) based on synthesized embeddings extracted from textual, visual, and acoustic sources predicts a speaker's sentiment tendencies based on both textual and acoustic information. These embeddings are generated by combining input unimodal raw data to produce a richer multimodal feature representation. By employing Transformer-based models, lexicon-based features perform better than other modalities due to their pre-training on large corpora of text. In spite of their powerful performance, learning a new self-supervised learning Transformer on new modality is often not possible due to an insufficient amount of data, which is the case in language learning using multiple modalities. Bidirectional Encoder Representations from Transformer (BERT) is efficient pre-trained language model which provides state-of-the-art results on question answering and natural language inference and many other. However, most of the work in sentiment analysis is worked based on textual data, how to learn better representations with multimodal information is still area on which need to be explored. There are multiple modalities which we can exploit to get better sentiment analysis of person's behavior. Here our approach is to leverage the audio-based features combined with lexical features to get efficient and powerful model for sentiment analysis trained on CMU-MOSI dataset.

Keywords: Multimodal Sentiment Analysis, CMU-MOSI Dataset, Transformers

1. Introduction

The human communication does not limit itself to words. It includes acoustic annotations, body movements, facial expressions, and even body language. With proliferation of social media platforms such as Facebook, WhatsApp, YouTube, people produce large volumes of multimodal data with rich sentiment information every day. In addition to its role in human-computer interaction, sentiment analysis plays a vital role in artificial intelligence development and has been widely used in a variety of applications, including automatic driving, human-machine communication, and much more [1]. Text-based features which basically expresses sentiment through words, phrases, and relations generally outperforms other modalities [2]. In recent times text sentiment analysis has achieved lot of attention. Using pre-trained word vectors as inputs, TextCNN [3] is able to give state-of-art results on four out of seven sentiment classification tasks.

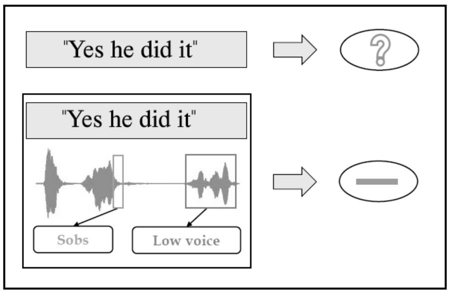

Fig. 1. Example of MSA [16]

However, in some cases it becomes difficult to identify the correct emotion of person when there is only lexicon-based features are available. Many times, text modality also coexist with audio modality. Audio modality includes loudness, pitch, energy, vocal effort and other frequency-related characteristics that conveys sentiment information [4]. A combination of text and audio channel can give more comprehensive picture of information and capture more emotions. In Figure - 1, we can see an example of intermodally of text and audio. The sentence "Yes he did it" is ambiguous as it can address variety of emotions in different situations. Its challenging to determine the sentiment of that kind of sentence just by taking those series of words as inputs. After introducing speaker's audio related information, we can clearly identify that the intent of speaker is negative in this scenario. Multimodal Sentiment Analysis (MSA) has been gaining increased popularity in recent years as a way to compensate for disadvantage of single modality of affective computing.

Recent large language models use Self-Supervised Learning and Transformer based methods utilizing large number of datapoints for pre-training in textual, vision and multimodal contexts. Collection of large volumes of data in multimodal context is not always easy. Recently, Bidirectional Encoder Representations from Transformer has presented state-of-the-art results on Eleven NLP tasks. The contextual representations generated by BERT are conditioned on both left and right context in each layer. So, the inference basically describes context content [5]. We explore approach to generalize pre-trained language model which exposes audio modality for better and efficient sentiment analysis.

2. Theory Background

An important subset of Natural Language Processing (NLP) is sentiment analysis, which focuses on identifying qualitative characteristics in data such as sentiment, opinion, thoughts, and behavioral intent. Many approaches [34] have been proposed to achieve state-of-the-art results on sentiment analysis. With recent advancements in supervised learning approaches have led to growth of field of Multimodal Sentiment Analysis.

2.1. Approaches

Based on the given feature, categorizing the polarity into Positive, Negative, or Neutral is important task. The approaches towards sentiment analysis are divided in to following categories.

2.1.1. Lexicon Based

A lexicon-based approach is traditional approach for sentiment analysis which uses a set of manually framed rules to classify the input into positive or negative opinions [26]. Words that elicit positive or negative feelings in humans are scanned through documents using lexicon-based methods. The method basically involves different Natural Language Processing techniques like Lexicons, Stemming, Part of Speech, Tokenization, etc. This approach counts the number of words representing various emotions in the input, then classifies them accordingly. Sentiment analysis shows to be extremely dependent on the domain of interest [27]. For example, analyzing specific product reviews can yield very different results compared to analyzing social media data due to different forms of language used. Which is one of the issues and adjusting model based on that can be time consuming process. However, lexicon-based methods do not require training data which can be beneficiary for small datasets.

There are two lexicon-based approaches:

- Dictionary-based
- Corpus-based

In Dictionary-based approach small set of sentiment words are prepared and then process iteratively expands the lexicon of sentiment words with synonyms and antonyms from existing dictionaries. Domain-specific lexicons can be created using corpus-based lexicons. The approach starts with a list of general-purpose sentiment words and discovers other sentiment words from a domain corpus based on co-occurring word patterns [29].

2.1.2. Machine Learning Based

Machine learning based approaches for sentiment analysis can be divided as below:

- Unsupervised Learning
- Semi-supervised Learning
- Supervised Learning.

Unsupervised Learning algorithms mostly work with unlabeled type data and with the help of algorithms different patterns/structures are discovered in data. One of the advantages of this kind of methods is that we can use large datasets without investing more effort on human supervision and labeling.

Supervised Learning is one of the most widely known machine learning method. In this approach the datasets are pre labeled with classes or ground truth. In this method, a model is trained with labeled source data, which can be used in making predictions using new unlabeled input data. In most cases, supervised learning often outperforms unsupervised and semi-supervised learning approaches, but the dependency on labeled training data can require lots of human effort and is therefore sometimes inefficient [26]. Some of the examples of this approach is Support Vector Machine (SVM), Neural Networks, Bayesian Networks and Naïve Bayes.

Semi-Supervised Learning can work with both labeled and unlabeled dataset where unlabeled data are complemented with the labeled examples for training the model. Compared to supervised learning this method requires less effort in labeling and yields decent accuracy results.

2.1.3. Hybrid approach

Hybrid approaches mainly aim to extend machine learning models with lexicon-based knowledge [11]. Using both lexicon and machine learning based features, the goal is to yield optimal results by combining both methods [12]. For example, Minaee et al. [13] developed an ensemble model using LSTM and CNN algorithm and demonstrated that this ensemble model provides better performance than the individual models.

Study of Pooja Mehta and Dr. Sharnil [14] shows that machine learning methods, such as SVM, Naive Bayes, and neural networks have the highest accuracy and can be considered as the baseline learning methods as well as in some cases lexicon-based methods are very effective. Hence our study on sentiment analysis will be mostly covering machine learning aspect.

2.2. Levels of Sentiment Analysis

Three different levels are considered while doing sentiment analysis namely the document level, the sentence level and the entity or aspect level.

2.2.1. Document Level

Document-level analysis considers the whole text document as a unit of analysis [21]. This task assumes that only one opinion holder is involved in creation of document. It should be noted that document analysis has its own issues, such as the fact that multiple mixed opinions are sometimes expressed in same document through implicit language [22].

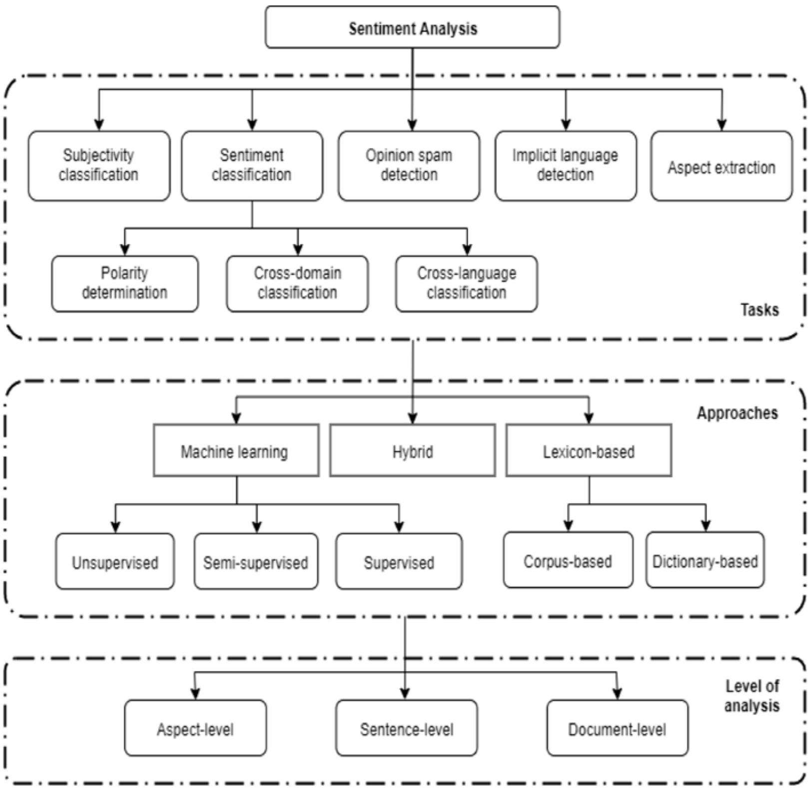

Fig. 2. Sentiment Analysis Concept Overview [25]

2.2.2. Sentence Level

In sentence-level analysis, specific sentences in text are considered and categorized into different categories based on their sentiments. Basically, method analyses individual sentences in a document to detect whether the sentence contains facts or emotions and opinions. Sentence is neutral when it does not imply any opinion. In case a sentence is neutral, it's termed as objective sentence, which is sentence which reveals facts, as opposed to subjective sentences which present opinions and subjective views. Subjectivity and objectivity classification are the main advantage of sentence analysis [22].

2.2.3. Aspect Level

Aspect level sentiment analysis also known as entity-level or feature-level sentiment analysis refers to analyzing sentiments about specific entities and their aspects in a text document, not merely the overall sentiment of the document. The output of this Aspect level more detailed analysis as compared to both document and sentence level analysis. It's possible for opinion holder to have divergent opinions about specific aspects of entity, despite the general sentiment of document which can be either positive or negative [22]. To measure aspect-level opinion, aspects of the entity need to be identified [24]. stated that aspect-based sentiment analysis is beneficial to the business manager because customer opinions are extracted in a transparent way.

2.3. Deep Learning Based Methods

Figure – 3 Demonstrates recent advancements in sentiment analysis space and its evident that Language models are one of the latest and biggest advancement in NLP and sentiment analysis space. One of sub-branch of machine learning is Deep Learning which consists of approaches like CNN, RNN, LSTM[30], etc.

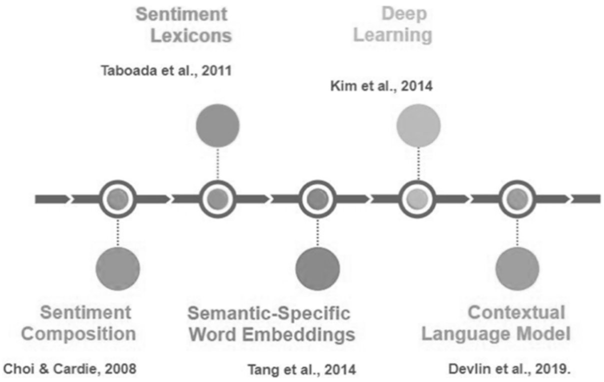

Fig. 3. Milestones of Sentiment Analysis [25]

The latest advancement Language Models are built on top of many deep learning-based concepts like RNN, LSTM, Transformers, Word Embeddings. Many Language models like BERT[33], GPT-2[32] and GPT-3[31] are provide state of the art performance on sentiment analysis data. To train them we just have to leverage pre-trained model with fine-tunning layer added on top of that for general NLP tasks.

2.4. Multimodal Sentiment Analysis

As an extension to traditional text-based sentiment analysis, multimodal sentiment analysis includes features such as text, speech as well as visual. For multimodal sentiment analysis variety of two modality combinations image+text, speech+text, speech+image combinations can be used which are called bimodal systems or some proposed models uses combination of all three of them. The use of visual features can describe something more effectively than a long list of written words. Audio data can also provide important indicators such as pauses, larger number of pauses indicates natural sentiment. The image below illustrates different phases of sentiment analysis based on multimodal fusion combining audio-visual features.

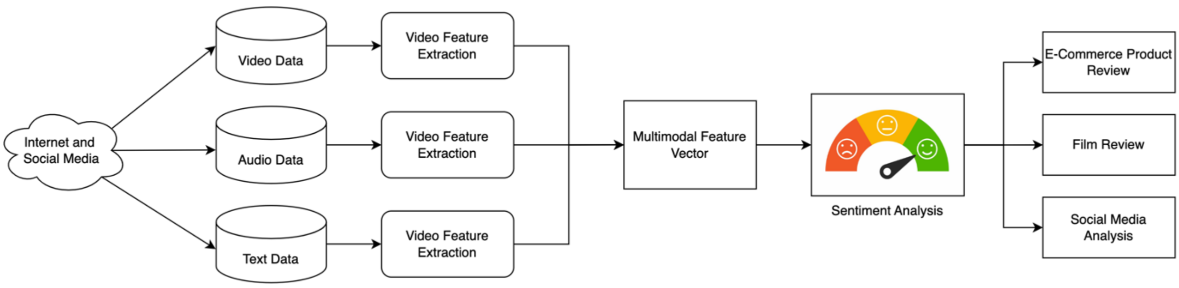

Fig. 4. Multimodal Sentiment Analysis Process

3. Literature Survey

3.1. Contextual Inter-modal Attention for Multi-modal Sentiment Analysis

An important challenge associated with multimodal sentiment analysis is combining text, visual and acoustic inputs effectively. RNN based approach was taken by Deepanway Ghosaly, Md Shad Akhtary, Dushyant Chauhany et al. [20] for multi-modal attention framework that leverages the contextual information for utterance-level sentiment prediction. The approach taken applies attention on multi-modal multi-utterance representations and tries to learn the features. This approach was able to mark F1 score of 82.31.

3.2. Multimodal Transformer for Unaligned Multimodal Language Sequences

Yao-Hung Hubert Tsai, Shaojie Bai, Paul Pu Liang et al.[19] explores the challenges like inherent data non-alignment due to variable sampling rates for multimodal sequences and long range dependencies between elements across modalities. Multimodal Transformer (MulT) introduced here generically address these issues without explicitly aligning the data. Directional pairwise cross-modal attention attends to interaction between multimodal sequences across distinct time steps and latency adapt streams across modalities. The model is evaluated with IEMOCAP, CMU-MOSI and CMU-MOSEI dataset with F1 score of 81.6 on CMU-MOSEI dataset.

3.3. Gated Mechanism For Attention Based Multimodal Sentiment Analysis

In recent times key to achieving state-of-art performance on downstream NLP tasks is to fine-tune the trained contextual language model on task-specific datasets. For tasks related to lexical data its straight-forward while, it's not trivial for multimodal language. Challenges like cross modal sentiment analysis, learning long-term dependencies and Fusion of unimodal and cross modal cues multimodal sentiment analysis where addressed Ayush Kumar and Jithendra Vepa [18]. The aim in their approach is to learn interaction between different modalities controlled by learnable gates. By this approach F1 score of 81.17 was achieved also thy discuss issues related to audio quality leading to lower speech recognition accuracy eventually affecting sentiment score.

3.4. VAE-Based Adversarial Multimodal Domain Transfer for Video-Level Sentiment Analysis

Yanan Wang et al. [17] introduced VAE-based adversarial multimodal domain transfer (VAE-AMDT) and jointly trained it with multi-attention module to reduce the distance difference between unimodal representations. For visual, linguistic and acoustic representations to follow common distributions they perform variational autoencoder (VAR) and then transfer their all-unimodal representations to joint embedding space using adversarial space. They jointly train VAE-AMDT and multi-attention module which consist of self-attention, cross-attention and triple-attention components reaching to F1 score of 84.3 on MOSI dataset. Also Yanan Wang et al. [17] discussed about future steps to improve their approach by using more powerful encoders to extract unimodal representations such as face identification methods for emotion.

3.5. CM-BERT: Cross-Modal BERT for Text-Audio Sentiment Analysis

Similar to Ayush Kumar and Jithendra Vepa [18] In Kaicheng Yang et al.[16] proposed Cross Modal BERT (CM-BERT), which relies on the interaction of text and audio modality to fine-tune pre-trained BERT model. The input of model consist of text features which are basically output encoding of BERT last layer and Audio features got by some pre pre-processing such as word-level alignment. Using text and audio modality together, the introduced multimodal attention dynamically adjusts word weight. They evaluated the modal on CMU-MOSI and CMU-MOSEI dataset achieving F1 score of 84.5.

3.6. Integrating Multimodal Information in Large Pretrained Transformers

Wasifur Rahman et al. [15] designed framework that allows BERT and XL-Net core structures to remain intact, and only attaches a carefully designed Multimodal Adaption Gate to the models. They basically use conditional attention on nonverbal behaviors, the gate essentially maps the informative visual and acoustic factors to vector with trajectory magnitude and during fine-tuning this adaption vector modifies the internal state of BERT and XL-Net, allowing the models to seamlessly adapt to nonverbal inputs. By using this model achieved powerful performance (MAG-BERT - F1 Score 86.00) on CMU-MOSI (Multimodal Opinion Sentiment Index) dataset.

4. Challenges and Issues

As per the literature review few of the challenges and issues which are there in the multimodal analysis space can be found below.

4.1. Extend and test approaches to real-world datasets and scenarios

The effectiveness of multimodal sentiment analysis approaches needs to be tested on a variety of real-world datasets and scenarios to ensure that they are robust and reliable in practical applications.

4.2. Handling implicit sentiments like sarcasm

Identifying implicit sentiments, such as sarcasm or irony, is a significant challenge in multimodal sentiment analysis since these sentiments are often conveyed through indirect or ambiguous language.

4.3. Improving model accuracy and efficiency

Multimodal sentiment analysis models often require large amounts of data and computational resources, which can make them slow and computationally expensive. There is a need to improve the accuracy and efficiency of these models, while also reducing their computational requirements.

4.4. Handling different representations of different modalities

Different modalities, such as text, image, and audio, have different representations, and integrating them in a meaningful way requires sophisticated algorithms that can extract relevant features from each modality.

4.5. Limited availability of data for non-text modalities:

There is often a scarcity of labeled data for modalities other than text, such as images and audio, which makes it challenging to develop robust multimodal sentiment analysis models.

4.6. Generalizability of the models

Multimodal sentiment analysis models may perform well on specific datasets or in certain scenarios, but their generalizability to new datasets or scenarios is often limited. There is a need to develop models that can generalize well to a wide range of datasets and scenarios.

5. Problem Statement, Aims and Objectives

With great advancement in digital world and communication technology, increasing popularity of the various portable devices large amount of data is being uploaded as audio, video, and text. Considering an example of consumers of the products who record their reviews on product on social media platforms like YouTube, Facebook to inform larger population about their views and opinions. This huge amount of data being generated everyday and requirement of organizations to make quick decision about product and services from feedbacks requires accurate and automated solution which can give insights about opinions and emotions about the individuals.

The primary advantage of analyzing audio and video over simple text analysis, for detecting emotions and sentiment, is the surplus of tonal and behavioral cues and as the research on similar subjects shows that even just including audio modalities with text modalities can give us significant boost to the sentiment analysis accuracy.

Recently many approaches have been proposed for multimodal sentiment analysis giving state-of-the-art results. However as discussed major issues and challenges remains there like impact of each modality across datasets, generalization ability of multimodal sentiment classifier, improving accuracy for larger databases, handling different representations from different modalities.

The main goal here would be to make a powerful and efficient NLP modal which can do Multimodal Sentiment Analysis on real-world streaming data which then can be used in many applications of NLP such as Conversational AI, human- computer interactions etc.

Specifically primary focus is to evolve current solution with below Aims and Objective,

1. To Improve accuracy
2. To lookout for solution with real time sentiment detections

6. Proposed Work

Language model built based on transformer provide powerful way of learning for NLP tasks. Since lexical features are trained on large text of corpora, they usually outperform other modalities in this domain. However as discussed introducing multimodal data can provide us more cue about person's opinion which can help us to get better at task [2].

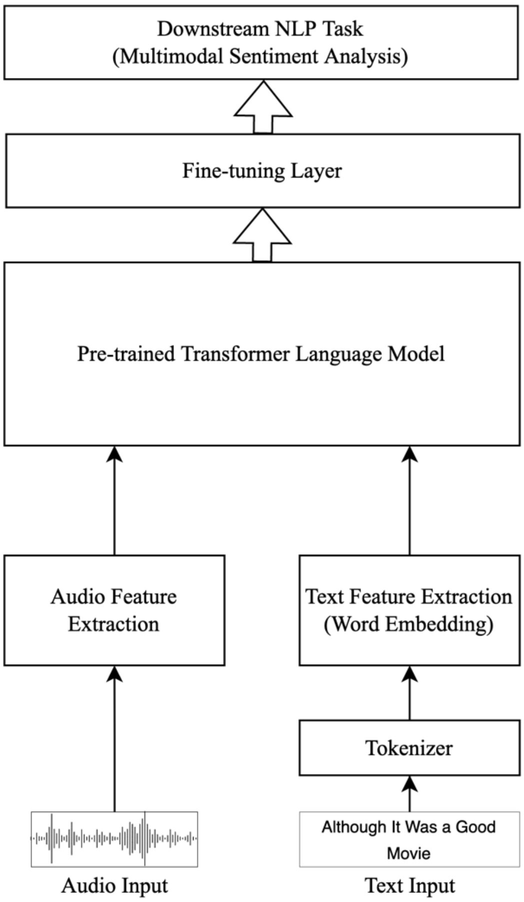

Fig. 5. Proposed Model

Our focus would be more on to work with audio modality in addition to lexicon features as working on raw speech data is less computationally intensive also it provides more informative advantage in multimodal language settings [2].

As with major language models approach, we plan to train model in two phases. 1.Train Pre-Trained Language Model with audio features on training tasks (Ex. Masked Language Model) with for smaller number of epochs. 2. Train the model on actual dataset for sentiment analysis with finetuning layer added on top of trained model.

Main challenge here would be to make audio feature extraction module which we can add as prefix to already trained model on large lexical data. Providing small number of epochs in first step with prefixed input will help model to adept prefixed audio features. Then in second step we can train the model on actual database with NLP task at hand i.e. multimodal sentiment analysis. The proposed approach will be more efficient in terms of training as compared to current systems which uses different branch for preprocessing audio and then generating features and combining output before fine-tuning layer.

OpenAI Whisper APIs will be used to get real time and accurate transcription of the user speech which will serve as inputs to multimodal sentiment analysis model. Overall flow for prediction can be found in below image.

We purpose to train such model on multimodal dataset available like CMU-MOSI, CMU-MOSEI, IEMO-CAP, etc. and to test on real world scenario to make efficient and powerful model.

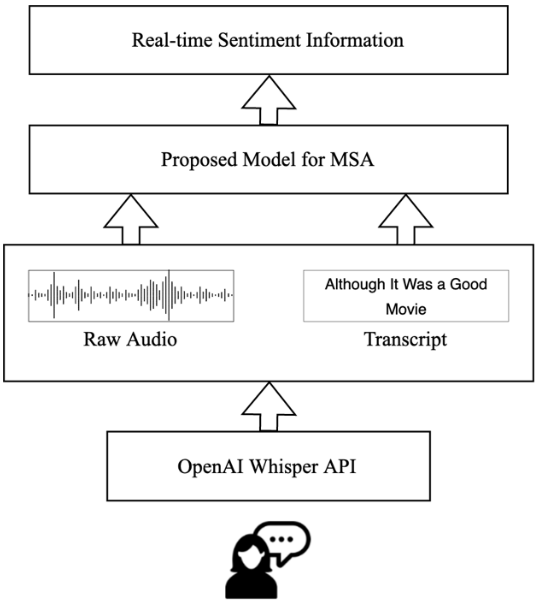

Fig. 6. Prediction Flow

7. Conclusion

Combining other modalities with lexical data and harnessing the power of already trained model with transformers on large text corpora can give us significant boost in performance for NLP tasks at hand specifically sentiment analysis. The proposed model is efficient in terms of leveraging already known features from lexical data model compared to other approaches that requires training new cross model transformer model which we can extend to larger data sets for better generalizability.

8. Future work

In future research, there is a need to develop strategies to handle implicit sentiments, including sarcasm, as they can significantly impact the sentiment analysis results. Furthermore, to enhance the practical applicability of the models, it is essential to focus on improving their generalizability across different domains. One potential approach could be to incorporate visual data, such as images and videos, to better capture contextual information and improve the overall accuracy of the sentiment analysis models.

References

1. Tadas Baltrušaitis, Chaitanya Ahuja, and Louis-Philippe Morency. 2018. Multimodal machine learning: A survey and taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence.
2. Erik Cambria, Devamanyu Hazarika, Soujanya Poria, Amir Hussain, and R. B. V. Subramanyam. 2017. Benchmarking Multimodal Sentiment Analysis
3. Yoon Kim. 2014. Convolutional Neural Networks for Sentence Classification.
4. Runnan Li, Zhiyong Wu, Jia Jia, Yaohua Bu, Sheng Zhao, and Helen Meng. 2019.Towards discriminative representation learning for speech emotion recognition. In Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI). 5060–5066.
5. Xinlong Li, Xingyu Fu, Guangluan Xu, Yang Yang, Jiuniu Wang, Li Jin, Qing Liu, and Tianyuan Xiang. 2020. Enhancing BERT Representation With Context-Aware Embedding for Aspect-Based Sentiment Analysis.
6. Jiahong Yuan and Mark Liberman. 2008. Speaker identification on the SCOTUS corpus.
7. Gilles Degottex, John Kane, Thomas Drugman, Tuomo Raitio, and Stefan Scherer. 2014. COVAREP A collaborative voice analysis repository for speech technologies.
8. Choi Y, Cardie C. 2008. Learning with compositional semantics as structural inference for subsentential sentiment analysis.
9. Duyu Tang, Furu Wei, Nan Yang, Ming Zhou, Ting Liu, and Qin Bing. 2014. Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification.
10. Devlin J, Chang MW, Lee K, Toutanova K. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding.

11. Behera RN, Manan R, Dash S. 2016. Ensemble based hybrid machine learning approach for sentiment classification-a review.
12. Ahmad M, Aftab S, Ali I, Hameed N. 2017. Hybrid tools and techniques for sentiment analysis: a review.
13. Minaee S, Azimi E, Abdolrashidi A. 2019. Deep-sentiment: sentiment analysis using ensemble of cnn and bi-lstm models.
14. Pooja Mehta, Dr. Sharnil Pandya. 2020. A Review On Sentiment Analysis Methodologies, Practices And Applications
15. Wasifur Rahman, Md. Kamrul Hasan, Sangwu Lee, Amir Zadeh, Chengfeng Mao, Louis-Philippe Morency, Ehsan Hoque. 2020. Integrating Multimodal Information in Large Pretrained Transformers.
16. Kaicheng Yang, Hua Xu, Kai Gao. 2020. CM-BERT: Cross-Modal BERT for Text-Audio Sentiment Analysis
17. Yanan Wang, Jianming Wu, Kazuaki Furumai, Shinya Wada, Satoshi Kurihara. 2022. VAE-Based Adversarial Multimodal Domain Transfer for Video-Level Sentiment Analysis
18. Ayush Kumar and Jithendra Vepa. 2020. Gated Mechanism for Attention Based Multimodal Sentiment Analysis
19. Yao-Hung Hubert Tsai, Shaojie Bai, Paul Pu Liang, J. Zico Kolter, Louis-Philippe Morency, Ruslan Salakhutdinov. 2019. Multimodal Transformer for Unaligned Multimodal Language Sequences
20. Deepanway Ghosaly, Md Shad Akhtary, Dushyant Chauhany, Soujanya Poria, Asif Ekbaly, Pushpak Bhattacharyya .2018. Contextual Inter-modal Attention for Multi-modal Sentiment Analysis
21. Wang G, Sun J, Ma J, Xu K, Gu J. 2014. Sentiment classification: the contribution of ensemble learning. Decis Support Syst 57:77–93
22. Kumar A, Sebastian TM. 2012. Sentiment analysis: a perspective on its past, present and future. Int J Intell Syst Appl 4(10):1–14.
23. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M. 2011. Lexicon-based methods for sentiment analysis.
24. Valdivia A, Luzón MV, Herrera F. 2017. Sentiment analysis in tripadvisor.
25. Ligthart, A., Catal, C. & Tekinerdogan, B. Systematic reviews in sentiment analysis: a tertiary study.
26. Hemmatian F, Sohrabi MK. 2017. A survey on classification techniques for opinion mining and sentiment analysis.
27. Vinodhini G, Chandrasekaran RM. 2012. Sentiment analysis and opinion mining: a survey
28. Shayaa S, Jaafar NI, Bahri S, Sulaiman A, Seuk Wai P, Wai Chung Y, Piprani AZ, Al-Garadi MA. 2018. Sentiment analysis of big data: Methods, applications, and open challenges.
29. Mite-Baidal K, Delgado-Vera C, Solís-Avilés E, Espinoza AH, Ortiz-Zambrano J, Varela-Tapia E. 2018. Sentiment analysis in education domain: a systematic literature review.
30. Ralf C. Staudemeyer, Eric Rothstein. 2019. Morris Understanding LSTM - a tutorial into Long Short-Term Memory Recurrent Neural Networks
31. Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child. Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei. 2020. Language Models are Few-Shot Learners.
32. Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever. 2019. Language Models are Unsupervised Multitask Learners.
33. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
34. Shah, A., Chopade, M., Patel, P., & Patel, P. 2022. Survey: Emotion Recognition from Text Using Different Approaches. In Futuristic Trends in Networks and Computing Technologies: Select Proceedings of Fourth International Conference on FTNCT 2021 (pp. 433-445). Singapore: Springer Nature Singapore.

Acknowledgement

I would like to thank my esteemed research guide Prof. Madhuri Chopade for her invaluable supervision, support and tutelage during the thesis work providing all necessary information and discussion about my thesis progress with devotion of her valuable time. Her immense knowledge and plentiful experience have encouraged me all the time of my academic research and daily life. Additionally, this endeavor would not have been possible without the generous support from our Head of Computer Department Dr. Nimisha Patel, who provided valuable feedbacks toward improving my work.

I would also like to extend my sincere gratitude to Dr. H N Shah Director of Gandhinagar Institute of Technology, Gandhinagar, for his very humble support, motivation, continuous encouragement for innovation with new ideas, and for providing wonderful positive environment for research work.

I am thankful to and fortunate enough to get constant encouragement, support, and guidance from all Faculty members of Computer Engineering Department which helped me in successfully completing thesis work. At last, I hereby take this opportunity to thank all my colleagues, family, and friends and those who directly or indirectly helped me in preparing this thesis work.