International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Enhanced Sentiment Analysis Accuracy Using Multi-Modal Deep Learning Techniques

Ankit Kumar[1], Nitesh Gupta[2], Anurag Shrivastava[3]

[1]M. Tech Scholar, CSE, Department NIIST, Bhopal ank14itk@gmail.com
[2]Associate Professor, CSE, Department NIIST, Bhopal 9.nitesh@gmail.com
[3]Associate Professor, CSE, Department NIIST, Bhopal anurag.shri08@gmail.com

ABSTRACT

Sentiment analysis, an essential task in natural language processing (NLP), traditionally relies on textual data to gauge user opinions and emotions. However, the growing availability of diverse data types, such as images, audio, and video, has made multi-modal deep learning a promising approach for improving sentiment analysis accuracy. By incorporating multiple data modalities, multi-modal deep learning models offer a more comprehensive understanding of sentiment, capturing subtle emotional cues that text alone may overlook. This work explores the application of multi-modal deep learning techniques to enhance sentiment analysis accuracy. It reviews state-of-the-art models, including Convolutional Neural Networks (CNNs) for image and video analysis, Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks for text and audio, and Transformer-based architectures for multi-modal fusion. Techniques like feature fusion, cross-modal attention, and co-attention mechanisms are discussed for their ability to integrate and align different modalities effectively. The result indicate that multi-modal approaches outperform traditional single-modal methods by capturing deeper emotional contexts and improving sentiment prediction. Challenges such as data alignment, missing modality handling, and the high computational costs of training multi-modal models are also addressed. Future research directions are suggested, focusing on real-time systems, model interpretability, and the inclusion of new data types for even more accurate sentiment detection.

***Keywords*— Sentiment Analysis, Hybrid model, CNN, RNN, Deep Learning**

Introduction

Sentiment analysis, also known as opinion mining, plays a pivotal role in understanding and analyzing subjective information, such as opinions, emotions, and attitudes, from user-generated content. Traditionally, sentiment analysis focuses on textual data, utilizing natural language processing (NLP) techniques to extract sentiments from reviews, social media posts, blogs, and other text-based sources. While these approaches have achieved reasonable success, they often fail to capture the full spectrum of emotional nuances, particularly in cases where text alone is insufficient to convey sentiment accurately. With the rapid growth of multimedia content, there is an increasing need to incorporate multiple data modalities—such as images, audio, and video—into sentiment analysis. For example, a person's tone of voice, facial expressions, or the context of a visual scene can provide valuable emotional cues that enhance the understanding of sentiment beyond what text alone can offer. This is where multi-modal deep learning comes into play. By integrating information from various modalities, multi-modal approaches allow for a more holistic view of sentiment, thereby improving the accuracy and depth of sentiment analysis models.

Recent advancements in deep learning have enabled the effective fusion of multi-modal data through techniques like Convolutional Neural Networks (CNNs) for image analysis, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks for text and audio, and Transformers for capturing long-range dependencies across different modalities. Moreover, attention mechanisms such as cross-modal attention and feature fusion further enhance the ability to align and integrate diverse data sources. This research work explores how multi-modal deep learning techniques enhance sentiment analysis accuracy by leveraging various data types. It delves into the architectures, methods, and challenges of implementing multi-modal models, along with the potential they hold for real-world applications in fields such as marketing, customer feedback analysis, and social media monitoring. Through this exploration, we aim to shed light on how multi-modal learning can significantly improve sentiment detection outcomes.

Figure 1.1. Sentiment Analysis

BACKGROUND AND RELATED WORK

This work [1] the rating of movie in twitter is taken to review a movie by using opinion mining This paper proposed a hybrid methods using SVM and PSO to classify the user opinions as positive, negative for the movie review dataset which could be used for better decisions.

Authors [2] found that PSO affect the accuracy of SVM after the hybridization of SVM-PSO. The best accuracy level that gives in this study is 77% and has been achieved by SVM-PSO after data cleansing. On the other hand, the accuracy level of SVM-PSO still can be improved using enhancements of SVM that might be using another combination or variation of SVM with other optimization method.

Authors [3] perform sentiment analysis from the point of view of the consumer review summarization model for capitalists. Author's outlined several research concerns and possible solutions for the challenges that occur when performing sentiment analysis for raw online reviews. Using the hybrid feature extraction method proposed in this work, the input pre-processed reviews can be transformed into meaningful feature vectors, allowing efficient, reliable, and robust sentiment analysis to take place.

Authors [4] results show that sentiment analysis is an effective technique for classifying movie reviews. This analysis focused primarily on English-language movie reviews, and the models may not perform as effectively when applied to other languages due to linguistic variations and cultural differences. This study introduces a sentiment analysis approach using advanced deep learning models: Extra-Long Neural Network (XLNet), Long Short-Term Memory (LSTM), and Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM).

Authors [5] Hybrid deep sentiment analysis learning models that combine long short-term memory (LSTM) networks, convolutional neural networks (CNN), and support vector machines (SVM) are built and tested on eight textual tweets and review datasets of different domains. %e hybrid models are compared against three single models, SVM, LSTM, and CNN. Both reliability and computation time were considered in the evaluation of each technique. %e hybrid models increased the accuracy for sentiment analysis compared with single models on all types of datasets, especially the combination of deep learning models with SVM. %e reliability of the latter was significantly higher.

Authors [6] examine primary taxonomy and newly released multimodal fusion architectures. Recent developments in MSA architectures are divided into ten categories, namely early fusion, late fusion, hybrid fusion, model-level fusion, tensor fusion, hierarchical fusion, bi-modal fusion, attention-based fusion, quantum-based fusion and word-level fusion. A comparison of several architectural evolutions in terms of MSA fusion categories and their relative strengths and limitations are presented. Finally, a number of interdisciplinary applications and future research directions are proposed.

Authors [7] review the multimodal sentiment analysis by combining several deep learning text and image processing models. These fusion techniques are RoBERTa with EfficientNet b3, RoBERTa with ResNet50, and BERT with MobileNetV2. This work focuses on improving sentiment analysis through the combination of text and image data. The performance of each fusion model is carefully analyzed using accuracy, confusion matrices, and ROC curves. The fusion techniques implemented in this study outperformed the previous benchmark models. Notably, the EfficientNet-b3 and RoBERTa combination achieves the highest accuracy (75%) and F1 score (74.9%).

PROPOSED METHODOLOGY

This work focuses on sentiment classification analysis using deep learning techniques. We employ diverse deep learning algorithms, including CNN and RNN, to categorize image and text data into sentiment classes like positive, negative, and neutral. Through rigorous experimentation and evaluation, incorporating performance metrics such as accuracy, precision, recall, and F1 score, we assess the efficacy of these models in sentiment analysis tasks. Our research contributes to advancing sentiment analysis methodologies, shedding light on the effectiveness of machine learning approaches in deciphering and analyzing sentiments expressed in textual data.

Steps to Build the Model:

1. **Load Twitter Dataset**: Load Twitter dataset, ideally with text and any available metadata (like images, emojis, or audio features) for a multi-modal model.
2. **Preprocess the Data**: The data needs to be cleaned by removing noise such as stopwords, URLs, mentions, hashtags, and emojis. Text tokenization, stemming, or lemmatization will be applied for textual data. If images are present, resize and normalize them.
3. **Select Best Attribute**: Select key attributes (such as tweet text, user interactions, or image embeddings) for modeling. Feature selection can be based on importance scores or domain knowledge.
4. **Split Data into Training and Testing Sets**: Divide the dataset into 80% training data and 20% testing data
5. **Apply Deep Learning Multi-Modal Model**: Apply multi- model with two branches—one for text processing (e.g., using an LSTM) and another for other data modalities like images (e.g., using CNNs).
6. **Classify the Result**: The output layer classifies the sentiment into positive or negative categories.

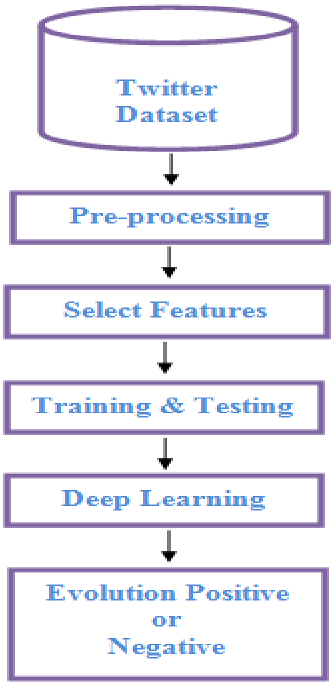

Figure: 3.1 Proposed Architecture for Sentiment Analysis

Twitter Dataset: This comprehensive dataset offers a valuable resource for researchers, data scientists, and enthusiasts interested in exploring the dynamics of social media communication. Spanning diverse topics, industries, and regions, our dataset provides an in-depth understanding of trends, sentiments, and user behavior within the Twitterverse. Analyze engagement metrics, track the evolution of hashtags, uncover influential users, and gain insights into the ever-changing landscape of online conversations. With our Twitter dataset, you can unlock the power of social media data to fuel research, drive innovation, and gain a deeper understanding of the world through the lens of Twitter [25]. Twitter is the preferred platform for individuals to express their thoughts and emotions through tweets. To gather data from Twitter, users subscribe to the Twitter API and authenticate it using access_token, access_secret, consumer_key, and consumer_secret. Initially, 2000 tweets were utilized to train the algorithms. Following training, the model processes pre-processed data. Out of 3000 tweets, 260 were removed, leaving 2680 true negatives and 160 true positives. The training dataset comprises 80% of the original data, while the testing dataset comprises 20%.

EXPERIMENTAL SETUP AND RESULT ANALYSIS

The proposed multi-modal for sentiment analysis using a Twitter dataset, follow the steps outlined below. The process includes loading the dataset, preprocessing it, selecting the best attributes, splitting the data into training and testing sets, applying a multi model of deep learning, and classifying the results as positive or negative. For experimental purpose Python tools and libraries are used.

Table 1: PERFORMANCE OF MULTI MODEL DEEP LEARNING MODEL

Parameter	**Proposed Multi Model**
Accuracy	**88.3%**
Precision	**87.2%**
Recall	**86.1%**
F-1 Score	**85.2%**

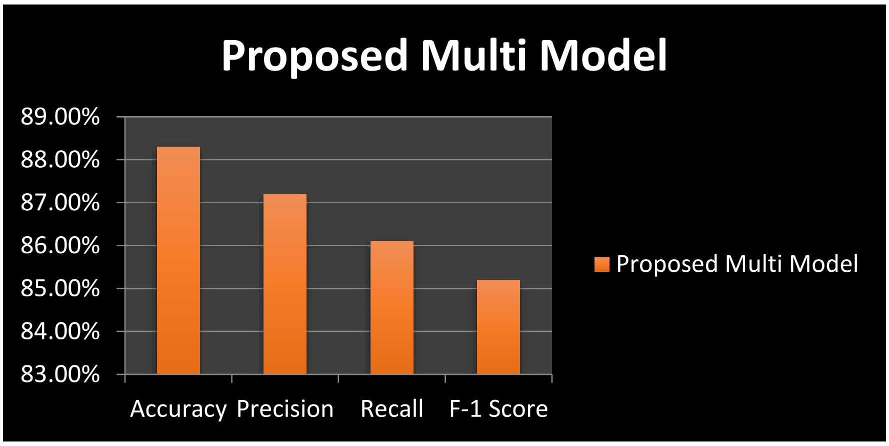

Figure 2.2 Performance comparison GRAPH.

CONCLUSION

The application of multi-modal deep learning techniques in sentiment analysis has shown significant potential for improving accuracy by incorporating a broader range of data sources, such as text, images, audio, and video. Traditional text-based sentiment analysis methods, while effective, often miss the subtle emotional cues found in other modalities. By integrating multiple data types, multi-modal approaches provide a more holistic and nuanced understanding of sentiment, leading to more accurate predictions. Deep learning models such as CNNs, RNNs, LSTMs, and Transformer-based architectures have been instrumental in effectively combining information from various modalities. Techniques like feature fusion and cross-modal attention have enhanced the ability to align and integrate data from different sources, leading to better sentiment detection. However, challenges remain, including the computational complexity of training multi-modal models, the alignment of heterogeneous data, and handling missing or incomplete modalities. Looking forward, the field presents exciting opportunities. Future research should focus on improving the efficiency of multi-modal models for real-time sentiment analysis, increasing their interpretability, and exploring new data types, such as physiological signals, to further enrich sentiment predictions. As multi-modal sentiment analysis becomes more refined, it will play a crucial role in applications ranging from social media monitoring and customer feedback analysis to marketing and user experience enhancement.

References

[1] K.Uma maheswari, Ph.D et al "Opinion Mining using Hybrid Methods" International Journal of Computer Applications (0975 – 8887) International Conference on Innovations in Computing Techniques (ICICT 2015)

[2] Abd. Samad Hasan Basaria et al "Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization" 1877-7058 © 2013 The Authors. Published by Elsevier Ltd.

[3] Gagandeep Kaur1,2* and Amit Sharma3 "A deep learning‑based model using hybrid feature extraction approach for consumer sentiment analysis"

[4] Mian Muhammad Danyal1, Opinion Mining on Movie Reviews Based on Deep Learning Models DOI: 10.32604/jai.2023.045617

2023,

[5] Cach N. Dang et al "Hybrid Deep Learning Models for Sentiment Analysis" Hindawi 2021

[6] Ankita Gandhi et. al. “Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions” https://doi.org/10.1016/j.inffus.2022.09.025, www.elsevier.com/locate/inffus, 2023

[7] Muhaimin Bin Habib1 et al “Multimodal Sentiment Analysis using Deep Learning Fusion Techniques and Transformers” (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 15, No. 6, 2024

[8] Lei Zhang and Bing Liu : Aspect and Entity Extraction for Opinion Mining. Springer-Verlag Berlin Heidelberg 2014., Studies in Big Data book series, Vol 1, pp. 1-40, Jul. 2014.

[9] Zhen Hai, Kuiyu Chang, Gao Cong : One Seed to Find Them All: Mining Opinion Features via Association. ACM CIKM’12., LNCS

6608, pp. 255-264, Nov. 2012

[10] Zhen Hai, Kuiyu Chang, Jung-Jae Kim, and Christopher C. Yang : Identifying Features in Opinion Mining via Intrinsic and Extrinsic

Domain Relevance. ieee transactions on knowledge and data engineering, Volume 26, No. 3 pp. 623-634, 2014.

[11] Hui Song, Yan Yan, Xiaoqiang Liu : A Grammatical Dependency Improved CRF Learning Approach for Integrated Product Extraction.

IEEE International Conference on Computer Science and Network Technology, pp. 1787-139, 2012.

[12] Luole Qi and Li Chen : Comparison of Model-Based Learning Methods for Feature-Level Opinion Mining. IEEE International Conferences on Web Intelligence and Intelligent Agent Technology, pp. 265-273, 2011.

[13] Arjun Mukherjee and Bing Liu: Aspect Extraction through Semi- Supervised Modeling. In: Association for Computational Linguistics.,

vol. 26, no. 3, pp. 339-348, Jul. 2012.

[14] Liviu, P.Dinu and Iulia Iuga.: The Naive Bayes Classifier in Opinion Mining:In Search of the Best Feature Set. Springer-Verlag Berlin Heidelberg, 2012.

[15] Xiuzhen Zhang., Yun Zhou.: Holistic Approaches to Identifying the Sentiment of Blogs Using Opinion Words. In: Springer-Verlag Berlin Heidelberg, 5–28, 2011.

[16] M Taysir Hassan A. Soliman., Mostafa A. Elmasry., Abdel Rahman Hedar, M. M. Doss.: Utilizing Support Vector Machines in Mining Online Customer Reviews. ICCTA (2012).

[17] Ye Jin Kwon., Young Bom Park.: A Study on Automatic Analysis of Social NetworkServices Using Opinion Mining. In: Springer-Verlag Berlin Heidelberg, 240–248, 2011.

[18] Anuj Sharma., Shubhamoy Dey: An Artificial Neural Network Based approach for Sentiment Analysis of Opinionated Text. In: ACM, 2012.

[19] Yulan He. : A Bayesian Modeling Approach to Multi-Dimensional Sentiment Distributions Prediction. In: ACM, Aug. 2012.

[20] Danushka Bollegala, David Weir and John Carroll : Cross-Domain Sentiment Classification using a Sentiment Sensitive Thesaurus. ieee transactions on knowledge and data engineering, pp. 1-14, 2012.

[21] Andrius Mudinas., Dell Zhang., Mark Levene. : Combining Lexicon and Learning based Approaches for Concept-Level Sentiment Analysis. In: ACM, Aug. 2012.

[22] Vamshi Krishna. B, Dr. Ajeet Kumar Pandey, Dr. Siva Kumar A. P “Topic Model Based Opinion Mining and Sentiment Analysis” 2018 International Conference on Computer Communication and Informatics (ICCCI -2018), Jan. 04 – 06, 2018, Coimbatore, INDIA

[23] Rita Sleiman, Kim-Phuc Tran “Natural Language Processing for Fashion Trends Detection” Proc. of the International Conference on Electrical, Computer and Energy Technologies (ICECET 2022)

20-22 June 2022, Prague-Czech Republic

[24] 1d.sai tvaritha, 2nithya shree j, 3saakshi ns 4surya prakash s, 5siyona ratheesh, 6shimil shijo “a review on sentiment analysis applications and approaches” 2022 JETIR June 2022, Volume 9, Issue 6 www.jetir.org (ISSN-2349-5162)

[25] https://www.kaggle.com/datasets/goyaladi/twitter-dataset