IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Survey On Deep Learning Multi-Modal For Sentiment Analysis

Payal Patel
M.Tech Scholar
CSE, Department
NIIST, Bhopal
upasanapatel11211@gmail.com

Prof. Anurag Shrivastava
Assistant Professor,
CSE, Department Of Computer Science & Engineering
NIIST, Bhopal

Prof. Nitesh Gupta
Assistant Professor,
CSE, Department Of Computer Science & Engineering
NIIST, Bhopal

Abstract: Sentiment analysis, a critical task in natural language processing, has advanced significantly with the integration of deep learning and multi-modal approaches. This survey explores the evolving landscape of multi-modal sentiment analysis, which combines text, audio, visual, and contextual data for richer emotional understanding. We analyze key deep learning architectures, including transformers, convolutional neural networks, and recurrent models, highlighting their roles in handling diverse modalities. Challenges such as modality fusion, data alignment, and scalability are discussed, along with potential solutions. Furthermore, the paper reviews prominent datasets and applications across domains like social media, healthcare, and entertainment. Future research directions are proposed to enhance multi-modal sentiment analysis performance.

Keywords— Sentiment Analysis, Deep Learning, Natural Language Processing (NLP), Multi Modal, Text Classification

I. Introduction

In recent years, the proliferation of social media platforms like Twitter has transformed how people express opinions, emotions, and sentiments online. With over 500 million tweets posted daily, Twitter offers a rich and dynamic source of real-time data for sentiment analysis. Sentiment analysis, or opinion mining, involves extracting and interpreting the emotions, attitudes, or opinions expressed in text. While traditional sentiment analysis primarily relies on text data, recent advancements in multi-modal deep learning have opened new possibilities for integrating diverse data modalities such as text, images, audio, and metadata, enabling more comprehensive and accurate sentiment analysis. Twitter data, in particular, poses unique challenges and opportunities due to its informal language, use of emojis, abbreviations, and the integration of multimedia content such as images, GIFs, and videos. A single tweet may carry nuanced sentiments that are difficult to detect from text alone. For instance, an image or emoji accompanying a sarcastic tweet may completely alter its intended sentiment. This has spurred the development of multi-modal approaches that aim to analyze and combine information from multiple sources to achieve a holistic understanding of the sentiment conveyed. Deep learning has emerged as the dominant technique for multi-modal sentiment analysis, owing to its ability to model complex relationships across different modalities. Architectures like convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformers such as BERT and GPT have demonstrated remarkable success in capturing intricate patterns in text. When extended to multi-modal tasks, these architectures incorporate advanced techniques like attention mechanisms and modality-specific encoders to process and integrate data from text, images, and other formats. This integration allows for more robust sentiment predictions, particularly for challenging scenarios like sarcasm detection, sentiment shifts, and ambiguous content.

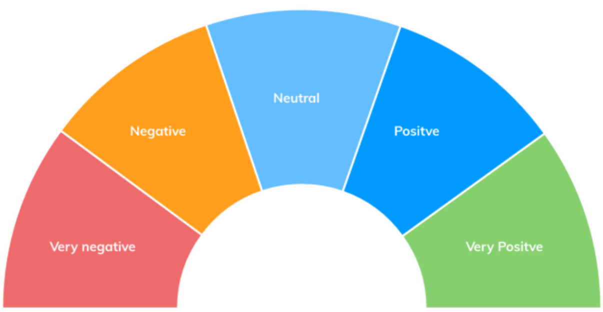

Figure 1.1: Sentiment Analysis

This survey aims to provide a comprehensive overview of multi-modal sentiment analysis using Twitter data with a focus on deep learning techniques. We explore state-of-the-art models and their applications in analyzing tweets, examining how different modalities are utilized and fused to enhance sentiment classification. Additionally, we discuss the unique challenges posed by multi-modal Twitter data, including noisy data, modality alignment, and the imbalance of sentiment classes. Prominent datasets, evaluation metrics, and domain-specific applications are also reviewed [4].

Finally, we highlight emerging trends and future directions in multi-modal sentiment analysis, such as the integration of contextual information, real-time processing, and improved scalability for large-scale datasets. By addressing the challenges and opportunities in this field, this survey aims to guide researchers and practitioners in leveraging multi-modal deep learning for more effective sentiment analysis on Twitter data.

II. LITRETURE REVIEW

The literature survey explores key developments in sentiment analysis, focusing on traditional machine learning approaches and their limitations. The survey also examines the ML model emphasizing their revolutionary role in achieving state-of-the-art performance in sentiment classification.

Authors [1] proposed approache outperform comparative techniques. These results provide valuable insights for implementing deep learning in sentiment analysis and contribute to setting benchmarks in the field, thus advancing both the theoretical and practical applications of sentiment analysis in real-world scenarios.

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. 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. Authors [2] find that 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. Reliability of the latter was significantly higher.

The work [3] systematically introduces each task, delineates key architectures from Recurrent Neural Networks (RNNs) to Transformer-based models like BERT, and evaluates their performance, challenges, and computational demands. The adaptability of ensemble techniques is emphasized, highlighting their capacity to enhance various NLP applications. Challenges in implementation, including computational overhead, over-fitting, and model interpretation complexities, are addressed, alongside the trade-off between interpretability and performance.

In this work [4] the rating of movie in twitter is taken to review a movie by using opinion mining. Author proposed 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.

This research [5] concerns on binary classification which is classified into two classes. Those classes are positive and negative. The positive class shows good message opinion; otherwise the negative class shows the bad message opinion of certain movies. This justification is based on the accuracy level of SVM with the validation process uses 10-Fold cross validation and confusion matrix. The hybrid Partical Swarm Optimization (PSO) is used to improve the election of best parameter in order to solve the dual optimization problem. The result shows the improvement of accuracy level from 71.87% to 77%

The purpose of this survey [29] is to provide a concise, nearly comprehensive overview of TSA techniques and related fields. The primary contributions of the survey are the detailed classifications of numerous recent articles and the depiction of the current direction of research in the field of TSA.

III. FINDINGS OF THE SURVEY

The survey on deep learning-based multi-modal sentiment analysis using Twitter data reveals significant advancements, challenges, and opportunities in this rapidly evolving field. The integration of multiple modalities, including text, images, audio, and contextual metadata, has proven to enhance the performance of sentiment analysis systems, particularly in addressing the ambiguities and complexities inherent in Twitter data. Below are the key findings from the study:

Multi-modal models outperform traditional text-only sentiment analysis techniques, as they leverage additional data such as images, emojis, and videos to provide richer contextual information. Tweets often include non-textual elements, and ignoring these modalities can lead to misinterpretation of sentiment. Multi-modal approaches effectively bridge this gap, especially in scenarios involving sarcasm, humor, or indirect sentiment expressions [5].

Deep learning techniques like transformers (e.g., BERT, RoBERTa, and Vision Transformers), CNNs, and RNNs have been successfully adapted to multi-modal sentiment analysis. These architectures incorporate advanced mechanisms such as attention models, cross-modal encoders, and fusion strategies to handle the complexity of combining diverse data types. Transformer-based models, in particular, dominate the field due to their ability to capture long-term dependencies and cross-modal relationships effectively [7].

One of the major challenges in multi-modal sentiment analysis is the alignment and fusion of information from different modalities. Misalignment, where data from one modality (e.g., image) does not correspond to another (e.g., text), affects model accuracy. Fusion techniques like early fusion, late fusion, and hybrid approaches have been explored, but achieving seamless integration remains a critical area of research.

The availability of annotated multi-modal datasets for Twitter is limited. While datasets like MOSI, MOSEI, and Memotion provide valuable resources, most are not tailored specifically for Twitter's unique characteristics, such as hashtags, informal language, and emoji usage. This highlights the need for more comprehensive datasets for benchmarking multi-modal sentiment analysis.

Multi-modal sentiment analysis on Twitter has shown significant potential across domains, including politics, marketing, public health, and disaster response. For instance, analyzing user sentiment during political campaigns or health crises provides actionable insights for decision-makers.

CONCLUSION

This survey highlights the growing importance and potential of deep learning-based multi-modal sentiment analysis on Twitter data. By integrating diverse modalities such as text, images, emojis, and contextual metadata, multi-modal approaches offer a richer understanding of sentiments compared to traditional text-only methods. Advances in deep learning architectures, especially transformers, have significantly improved the accuracy and robustness of sentiment predictions across modalities. However, challenges such as modality alignment, fusion strategies, and the limited availability of annotated datasets specific to Twitter remain critical barriers.

Despite these challenges, the field has demonstrated its applicability across domains like politics, marketing, and public health, with promising opportunities for future research in real-time analysis and personalized sentiment detection. Multi-modal sentiment analysis thus stands as a crucial step toward understanding the complexity of human emotions in online environments.

REFERENCES

[1] Huu-Hoa Nguyen "Enhancing Sentiment Analysis on Social Media Data with Advanced Deep Learning Techniques" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 15, No. 5, 2024

[2] Cach N. Dang et al. "Hybrid Deep Learning Models for Sentiment Analysis" Hindawi Complexity Volume 2021, Article ID 9986920, 16 pages https://doi.org/10.1155/2021/9986920

[3] Jianguo Jia et al. "A Review of Hybrid and Ensemble in Deep Learning for Natural Language Processing" https://doi.org/10.48550/arXiv.2312.05589

[4] K.Umamaheswari, 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)

[5] 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.

[6] Yili Wang et. al. "Review Sentiment Analysis of Twitter Data" Appl. Sci. 2022, 12, 11775. https://doi.org/10.3390/app122211775

[7] Gagandeep Kaur1,2* and Amit Sharma3 "A Deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis" Kaur and Sharma Journal of Big Data (2023) https://doi.org/10.1186/s40537-022-00680-6

[8] MianMuhammad Danyal1, Opinion Mining on Movie Reviews Based on Deep Learning Models DOI: 10.32604/jai.2023.045617 2023,

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

[10] 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.

[11] 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

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

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

[18] 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).

[19] 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.

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

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

[22] 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.

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

[24] 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

[25] 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

[26] 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)

[27] Pansy Nandwani1 · Rupali Verma1 "A review on sentiment analysis and emotion detection from text" https://doi.org/10.1007/s13278-021-00776-6

[28] Hoong-Cheng Soong, Norazira Binti A Jalil, Ramesh Kumar Ayyasamy, Rehan Akbar "The Essential of Sentiment Analysis and Opinion Mining in Social Media" 978-1-5386-8546-4/19/$31.00 ©2019 IEEE

[29] Muhammet Sinan et al. "Sentiment Analysis with Machine Learning Methods on Social Media" Advances in Distributed Computing and Artificial Intelligence Journal Regular Issue, Vol. 9 N. 3 (2020), 5-15 eISSN: 2255-2863 DOI: https://doi.org/10.14201/ADCAIJ202093515