WWW.IJRPR.COM

International Journal of Research Publication and Reviews

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

"A Review on Sentiment Analysis of Social media text using Deep learning Technique"

Prince Kumar[1], Prof. Nitesh Gupta[2], Prof. Anurag Shrivastava[3]

1M.Tech Scholar, CSE, Department Of Computer Science & Engineering NIIST, Bhopal
2Assistant Professor , CSE, Department Of Computer Science & Engineering NIIST, Bhopal
3Assistant Professor CSE, Department Of Computer Science & Engineering NIIST, Bhopal
[1]princewithkumar@gmail.com, [2]9.nitesh@gmail.com, [3]anurag.shri08@gmail.com

Abstract:

Sentiment analysis of social media text has gained prominence as a crucial application of deep learning techniques due to the exponential growth of user-generated content. This study explores the effectiveness of advanced deep learning models in assessing sentiment in social media posts, which are often characterized by informal language, slang, and varying contextual nuances. We investigate several state-of-the-art architectures, including Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), and Transformer-based models like BERT (Bidirectional Encoder Representations from Transformers), to enhance sentiment classification accuracy. By leveraging large-scale datasets from platforms such as Twitter and Facebook, we address challenges such as irony, sarcasm, and context-dependent sentiment expression. The study demonstrates that Transformer models, in particular, significantly outperform traditional methods in capturing complex sentiment patterns and contextual meanings. Through comprehensive experimentation and analysis, we highlight the impact of model choice and hyperparameter tuning on performance. Our findings contribute to the development of more accurate sentiment analysis systems, providing valuable insights for businesses, policymakers, and researchers interested in understanding public opinion and social trends.

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

1. Introduction

Sentiment analysis, a subfield of Natural Language Processing (NLP), focuses on identifying and extracting subjective information from textual data. As the digital era progresses, understanding sentiment—whether positive, negative, or neutral—has become pivotal for various applications such as customer feedback, social media monitoring, and market research. Traditional sentiment analysis methods often relied on rule-based systems and simple machine learning models, which were limited in their ability to capture the nuances and complexities of human language [2]. Recent advancements in deep learning have revolutionized sentiment analysis, offering new ways to interpret and classify text with greater accuracy and contextual understanding. Deep learning models, particularly those based on neural networks, have demonstrated remarkable capabilities in processing and analyzing large volumes of textual data. These models leverage hierarchical structures and large-scale pre-trained embeddings to enhance their performance in sentiment classification tasks [4]. One of the significant breakthroughs in this field is the development of contextual word embeddings, such as those provided by BERT (Bidirectional Encoder Representations from Transformers). Unlike traditional word embeddings that represent each word with a static vector, BERT and similar models generate dynamic embeddings based on the surrounding context of each word. This ability to understand context allows for more nuanced sentiment analysis, capturing subtle variations in meaning that were previously challenging for conventional models [3]. In this paper, we explore the application of advanced deep learning techniques for sentiment analysis, focusing on the benefits and limitations of various models. We begin with a review of foundational data pre-processing methods that are essential for preparing textual data for deep learning models. Next, we delve into different deep learning architectures, including Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), Convolutional Neural Networks (CNNs), and Transformers, assessing their effectiveness in sentiment classification tasks [5]. Particularly, we highlight the use of BERT and its variants for sentiment analysis, demonstrating how these models outperform traditional methods in terms of accuracy and robustness. Through empirical evaluation on popular sentiment datasets, we showcase the superior performance metrics achieved by deep learning models, including accuracy, precision, recall, and F1-score. This introduction sets the stage for a detailed examination of how advanced deep learning techniques are transforming sentiment analysis, providing insights into their practical applications and future directions in the field of NLP [6].

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 emergence of transformer-based models like BERT, emphasizing their revolutionary role in achieving state-of-the-art performance in sentiment classification. Table 2 shows relevant research in sentiment analysis field.

Table 1 Summary of Related Work

S. No.	Title	Description	Future work
1	Enhancing Sentiment Analysis on Social Media Data with Advanced Deep Learning Techniques	Authors 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.	These strategic directions are intended to not only advance the technical aspects of sentiment analysis but also to increase its practical relevance and effectiveness in dynamic environments.
2	Hybrid Deep learning models for sentiment Analysis	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. 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.	In future various other combinations of models are tested on reliability and computation time. Reliability of the latter was significantly higher.
3	A Review of Hybrid and Ensemble in Deep Learning for Natural Language Processing	The work 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 future the synergistic alliance between ensemble methods and deep learning models in the realm of NLP epitomizes the scientific community's unwavering endeavor to continually redefine the boundaries of linguistic understanding and computational capabilities.
4	Opinion Mining using Hybrid Methods	In this work 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.	The work done in this research is only related to classification opinions into two classes, positive and negative class. The future work, a multiclass of sentiment classification such as positive, negative and neutral can be considered.
5	Opinion Mining of Movie Review using Hybrid Method of Support Vector Machine and Particle Swarm Optimization	This research 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%.	In the future development, a multiclass of sentiment classification such as positive, negative, neutral and so on might be taken into consideration.

3.FINDINGS OF THE SURVEY

The review on sentiment analysis of social media text using deep learning techniques reveals several key insights and trends in the field. Firstly, deep learning models, particularly those based on Transformer architectures like BERT and GPT, have demonstrated superior performance over traditional methods in understanding and classifying sentiment in social media content. These models excel due to their ability to capture contextual information and handle the informal and diverse language used on social media platforms. The study highlights that while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks were effective in earlier applications, Transformer models have set new benchmarks in accuracy and efficiency. Transformers benefit from their attention mechanisms, which allow them to focus on relevant parts of the text, thereby improving sentiment detection, especially in complex or nuanced expressions. Another significant finding is the impact of pre-training on large corpora of text. Pre-trained models, such as those fine-tuned on domain-specific social media datasets, achieve higher accuracy in sentiment classification by leveraging learned contextual relationships and linguistic patterns. The review also underscores the importance of handling the challenges specific to social media texts, such as slang, emojis, and sarcasm. Advanced models with fine-tuning strategies are better equipped to interpret these nuances. The findings further reveal that ensemble approaches, which combine multiple models or techniques, often result in improved performance. These approaches can integrate the strengths of various models to address the limitations of individual ones. Overall, the review concludes that while significant progress has been made, ongoing research is needed to address remaining challenges and enhance model robustness in the ever-evolving landscape of social media language.

CONCLUSION

The review on sentiment analysis of social media text using deep learning techniques underscores the transformative impact of advanced models in this domain. The application of deep learning, particularly Transformer-based architectures such as BERT and GPT, has markedly improved the accuracy and effectiveness of sentiment classification compared to traditional methods. These models excel in interpreting the nuanced, informal, and often ambiguous language prevalent on social media platforms, offering substantial improvements in handling context-dependent sentiment.The findings highlight that while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have provided valuable contributions, Transformer models have emerged as the state-of-the-art due to their sophisticated attention mechanisms and ability to capture intricate contextual relationships. This advancement has led to more precise sentiment analysis, addressing challenges such as sarcasm, irony, and slang, which are common in social media texts. Pre-training on extensive text corpora and fine-tuning on domain-specific datasets have proven critical in enhancing model performance. These techniques enable models to better understand and adapt to the unique characteristics of social media language. Additionally, ensemble methods, which combine multiple models or strategies, have shown promise in further improving classification outcomes by leveraging the strengths of various approaches. Despite these advancements, the review identifies several areas requiring further research, including improving model robustness against evolving social media language and addressing the limitations of current models. The dynamic nature of social media necessitates continuous adaptation and innovation in sentiment analysis techniques. In conclusion, while deep learning has significantly advanced sentiment analysis, ongoing efforts are needed to refine models and address emerging challenges. The continuous evolution in social media language and usage patterns will drive future research, aiming to enhance the accuracy and applicability of sentiment analysis tools in a rapidly changing digital landscape.

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] 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

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

[8] Cach N. Dang et al “Hybrid Deep Learning Models for Sentiment Analysis” Hindawi 2021

[9] 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.

[10] 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

[11] 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.

[12] 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.

[13] 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.
[14] 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.
[15] 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.
[16] Xiuzhen Zhang., Yun Zhou.: Holistic Approaches to Identifying the Sentiment of Blogs Using Opinion Words. In: Springer-Verlag Berlin Heidelberg, 5–28, 2011.
[17] 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).
[18] 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.
[19] Anuj Sharma., Shubhamoy Dey: An Artificial Neural Network Based approach for Sentiment Analysis of Opinionated Text. In: ACM, 2012.
[20] Yulan He. : A Bayesian Modeling Approach to Multi-Dimensional Sentiment Distributions Prediction. In: ACM, Aug. 2012.
[21] 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.
[22] Andrius Mudinas., Dell Zhang., Mark Levene. : Combining Lexicon and Learning based Approaches for Concept-Level Sentiment Analysis. In: ACM, Aug. 2012.
[23] 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
[24] 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
[25] 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)
[26] Pansy Nandwani1 · Rupali Verma1 "A review on sentiment analysis and emotion detection from text" https://doi.org/10.1007/s13278-021-00776-6
[27] 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
[28] 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