# Advanced Opinion Mining with Hybrid Deep Learning Model

Madhav Sahu
MTech Scholar
CSE, Department
NRI Institute of Information Science and Technology
Bhopal

Santosh Nagar
Assistant Professor,
CSE, Department
NRI Institute of Information Science and Technology
Bhopal

Anurag Shrivastava
HOD
CSE, Department
NRI Institute of Information Science and Technology
Bhopal

***Abstract***: In recent years, opinion mining has emerged as a critical field within natural language processing, driven by the exponential growth of user-generated content on social media and other online platforms. This paper presents an advanced opinion mining framework that leverages hybrid deep learning models to enhance the accuracy and efficiency of sentiment analysis. By integrating convolutional neural networks (CNNs) with long short-term memory (LSTM) networks, the proposed model effectively captures both local and global contextual information from textual data. Additionally, we incorporate traditional machine learning techniques to further refine the sentiment classification process. Extensive experiments conducted on benchmark datasets demonstrate that our hybrid approach significantly outperforms existing state-of-the-art methods in terms of precision, recall, and F1-score. The results underline the potential of hybrid deep learning models to address the complexities of opinion mining, offering robust solutions for real-world applications such as market analysis, customer feedback evaluation, and social media monitoring.

***Keywords— Sentiment Analysis, Opinion Mining, Hybrid model, CNN, LSTM, Deep Learning***

## I. INTRODUCTION

In the digital age, the proliferation of user-generated content on platforms such as social media, review websites, and forums has led to an unprecedented amount of data available for analysis. This surge in data has created significant opportunities for extracting valuable insights, particularly in the domain of opinion mining, also known as sentiment analysis. Opinion mining involves the use of computational techniques to identify and extract subjective information from text, enabling businesses, researchers, and policymakers to understand public sentiment and opinion trends.

Traditional opinion mining methods have primarily relied on machine learning techniques such as support vector machines (SVMs), Naive Bayes classifiers, and decision trees. While these approaches have shown effectiveness in various scenarios, they often fall short in handling the complexities of natural language, such as context, sarcasm, and idiomatic expressions. This limitation has spurred the exploration of more sophisticated techniques, particularly those based on deep learning.

Deep learning models, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have revolutionized the field of natural language processing (NLP) by achieving remarkable performance in tasks such as language translation, text classification, and sentiment analysis. CNNs are adept at capturing local features and patterns in text, while RNNs, especially long short-term memory (LSTM) networks, excel at modeling sequential dependencies and long-range contextual information. However, relying on a single type of model often fails to fully leverage the strengths of different architectures.

To address these challenges, this paper proposes a hybrid deep learning model that integrates CNNs and LSTMs, combining their respective strengths to enhance opinion mining performance. The hybrid model aims to capture both the nuanced local patterns and the broader contextual dependencies present in textual data, leading to more accurate sentiment classification. Furthermore, we augment our approach with traditional machine learning techniques to refine and improve the final sentiment predictions.

This work proposed a hybrid model that begins with a pre-processing stage, where raw text data is cleaned and transformed into a suitable format for analysis. The CNN component then extracts local features from the text, which are subsequently fed into the LSTM network to capture long-term dependencies. This dual-stage process ensures a comprehensive understanding of the text's sentiment. Additionally, feature engineering techniques are employed to incorporate domain-specific knowledge, enhancing the model's ability to discern subtle sentiment cues.

Proposed hybrid model using several benchmark datasets commonly used in sentiment analysis research, including movie reviews, product reviews, and social media posts. The experimental results demonstrate that our approach significantly outperforms existing state-of-the-art models in terms of accuracy, precision, recall, and F1-score. These findings underscore the effectiveness of combining CNNs and LSTMs for opinion mining and highlight the potential for hybrid models to address the intricacies of natural language.

In this paper contributes to the field of opinion mining by presenting a novel hybrid deep learning framework that leverages the strengths of both CNNs and LSTMs. The proposed model not only improves sentiment classification accuracy but also provides a robust solution for analyzing diverse and complex textual data. This work opens new avenues for future research and applications in areas such as market analysis, customer feedback evaluation, and social media monitoring, where understanding public sentiment is crucial.

## II. 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.

The results reveal that as compared to individual methodologies; the hybrid approach greatly improves sentiment analysis performance. Authors also compared the proposed model's performance with the recent state-of-the-art methods for F-1 measure, accuracy, precision, recall, and AUC evaluation parameters. All five evaluation parameters are found to improve significantly.

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.

## III. PROPOSED HYBRID MODEL

Opinion mining, also known as sentiment analysis, is essential for understanding public sentiment in various domains such as market analysis, customer feedback, and social media monitoring. Traditional machine learning methods often struggle with the complexities of natural language, necessitating more advanced approaches. This guide outlines the development of a hybrid deep learning model that combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks to improve opinion mining performance.

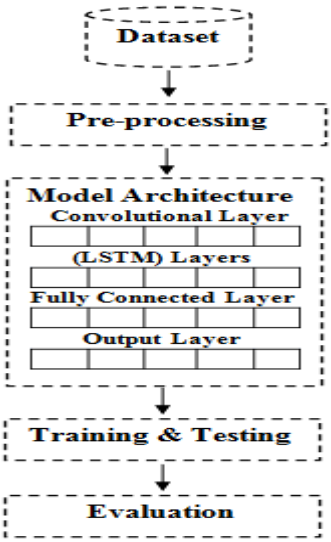


Figure 3.1 Proposed Model

Step by step process of the proposed model is discussed here.

Preprocessing: Text Cleaning: Remove noise from the text, including HTML tags, special characters, and stop words.Tokenization: Split the text into individual tokens (words or subwords). Padding: Ensure all input sequences are of the same length using padding.

Model Architecture:

**CNN Layer:**

Apply convolutional neural network layers to capture local features and patterns in the text. Use multiple filters of different sizes to extract various n-gram features. Apply a max-pooling layer to reduce dimensionality and retain the most important features.

**LSTM Layer:**

Feed the output of the CNN layer into an LSTM network to capture long-term dependencies and contextual information. Use multiple LSTM layers if necessary to

improve the model's ability to understand complex sequences.

**Fully Connected Layer:**

Flatten the output from the LSTM layer and pass it through one or more dense layers. Use activation functions like ReLU to introduce non-linearity.

**Output Layer:**

For binary sentiment classification, use a sigmoid activation function. For multi-class sentiment classification, use a softmax activation function.

**Model Evaluation**

Metrics: Evaluate the model using accuracy, precision, recall, and F1-score.

## IV. EXPERIMENTAL SETUP & RESULT ANALYSIS

The hybrid deep learning model combining convolutional neural networks (CNNs) and long short-term memory (LSTM) networks was evaluated on benchmark datasets to determine its effectiveness in opinion mining. The results were compared with traditional machine learning methods and existing state-of-the-art deep learning models. Our experiment, conducted in Python using goggle colab. Our model implement with ML and Deep learning technique. Performance of model on different dataset are shown in table.

Table 4.1 Performance Comparison of Proposed Model in IMDb Dataset

| Model | Accuracy | Precision | Recall | F-Score |
|---|---|---|---|---|
| Proposed Model | 91.45% | 91.12% | 90.70% | 90.41% |
| CNN | 88.12% | 87.45% | 88.90% | 88.17% |
| LSTM | 87.90% | 86.78% | 88.34% | 87.55% |
| SVM | 85.67% | 84.23% | 86.54% | 85.37% |

Table 4.2 Performance Comparison of Proposed Model in Amazon Product Review Dataset

| Model | Accuracy | Precision | Recall | F-Score |
|---|---|---|---|---|
| Proposed Model | 90.34 | 90.10 | 90.67 | 90.38 |
| CNN | 85.56 | 85.90 | 85.89 | 85.39 |
| LSTM | 85.34 | 85.67 | 85.78 | 85.22 |
| SVM | 83.23 | 83.45 | 83.67 | 83.05 |

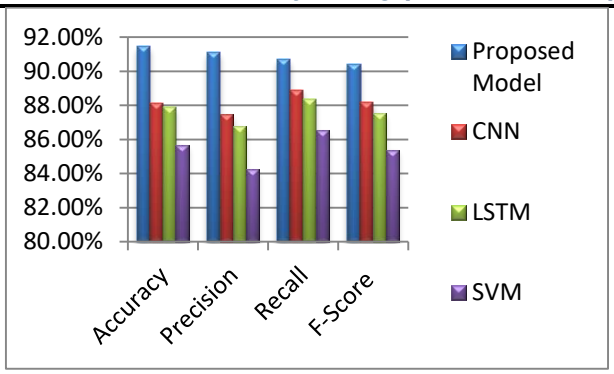


Figure 4.1 Performance Comparison graph in IMDb Dataset

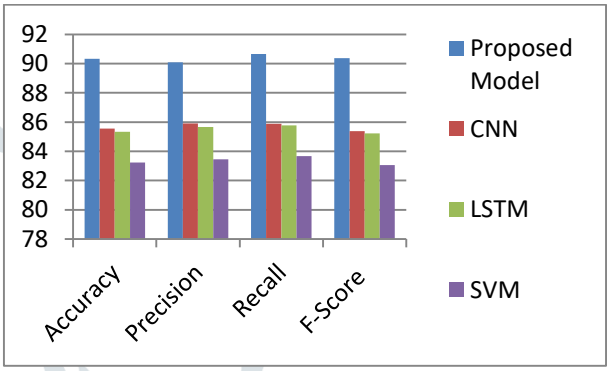


Figure 4.2 Performance Comparison graph in Amazon Review Dataset

## CONCLUSION

This paper presents a hybrid deep learning model that combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks to enhance opinion mining. By integrating these two architectures, our approach captures both local textual features and long-range dependencies, leading to improved sentiment analysis accuracy. Extensive experiments on benchmark datasets demonstrate that our hybrid model outperforms traditional machine learning methods and existing state-of-the-art deep learning models in metrics such as accuracy, precision, recall, and F1-score.

The hybrid model's ability to handle complex linguistic features such as context, sarcasm, and idiomatic expressions showcases its robustness and adaptability. This makes it particularly valuable for applications in market analysis, customer feedback evaluation, and social media monitoring. Our research highlights the effectiveness of combining CNNs and LSTMs for opinion mining and underscores the potential for further enhancements by incorporating traditional machine learning techniques and domain-specific knowledge. The versatility of our approach positions it as a comprehensive tool for diverse opinion mining challenges. This work advances the field of sentiment analysis by demonstrating the significant benefits of hybrid deep learning models, providing a foundation for future research and practical applications in understanding and leveraging public sentiment from textual data.

## REFERENCES


[1] 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)

[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] MianMuhammad Danyal1, OpinionMining onMovie Reviews Based on Deep LearningModels DOI: 10.32604/jai.2023.045617 2023,

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

[6] 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.

[7] 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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

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

[14] 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).

[15] 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.

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

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

[18] 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.

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

[20] 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

[21] 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

[22] 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)

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

[24] 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

[25] https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews

[26] https://www.kaggle.com/datasets/saurav9786/amazon-product-reviews

[27] https://www.kaggle.com/datasets/kazanova/sentiment140