INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 11, Issue 1 - V11I1-1559)
Available online at: https://www.ijariit.com

Advancing Sentiment Analysis: A Comprehensive Review of Data Augmentation Strategies and Deep Learning Techniques

Rashidkhan R Pathan
the.rashid.khans@gmail.com
L.D. College of Engineering, Ahmedabad, Gujarat

Dr. Pradip Patel
pradippatel@ldce.ac.in
L.D. College of Engineering, Ahmedabad, Gujarat

ABSTRACT

Sentiment analysis has gained significant attention in natural language processing (NLP) due to its applications in social media monitoring, customer feedback analysis, and opinion mining. However, a major chall enge in sentiment classification is class imbalance, where certain sentiment categories are underrepresented, leading to biased models and reduced accuracy. This research addresses the issue by leveraging data augmentation techniques, specifically synonym replacement and back-translation, to balance the dataset and enhance model performance. Unlike conventional deep learning approaches that rely on complex architectures, this study proposes a simplified yet more effective model by utilizing a Decision-based Recurrent Neural Network (D-RNN) trained on an augmented dataset. Additionally, Aspect-based and Priority-based augmentation techniques are introduced to ensure semantic consistency and emphasize critical contextual information during augmentation. Experimental results demonstrate that the proposed approach effectively reduces class skewness and improves sentiment classification accuracy, surpassing traditional models in performance while maintaining lower computational complexity. This research highlights the significance of data augmentation as a powerful strategy to enhance sentiment analysis, offering a cost-effective and scalable solution without the need for more complex deep learning models.

Keywords—*Sentiment Analysis, Opinion Mining, Data Augmentation, Neural Network, Deep Learning, State of Art Models, Synonym Replacement, Back-translation.*

INTRODUCTION

Sentiment analysis, also known as opinion mining, is a critical area of natural language processing (NLP) that focuses on determining the emotional tone behind a piece of text. With the rapid expansion of social media, e-commerce, and online reviews, sentiment analysis has become an essential tool for understanding public opinion and user feedback. Traditional deep learning models, such as BERT (Bidirectional Encoder Representations from Transformers) and Decision-based Recurrent Neural Networks (D-RNN), have demonstrated significant success in sentiment classification tasks[1]. However, one of the key challenges in sentiment analysis is class imbalance, where certain sentiment classes, such as neutral or negative sentiments, are underrepresented in the dataset. This imbalance leads to biased models that favor the majority class, ultimately reducing the effectiveness of sentiment predictions.

To address this issue, this research explores the role of data augmentation techniques in improving sentiment classification. Specifically, synonym replacement and back-translation are employed to artificially increase the number of samples in underrepresented classes, thereby reducing class skewness and enhancing model performance. Unlike conventional approaches that rely on highly complex architectures, this study proposes a less complex yet more effective solution by leveraging a Decision-based RNN (D-RNN) model trained on an augmented dataset.

Furthermore, a combined augmentation approach is introduced, integrating Aspect-based and Priority-based techniques to refine data augmentation strategies[2].

Aspect-based augmentation ensures that key contextual elements of sentiment remain intact, preventing distortion in meaning, while Priority-based augmentation selectively enhances the most critical portions of text to maximize the impact on classification accuracy.

By incorporating these augmentation strategies, the proposed methodology aims to enhance the robustness of sentiment analysis models, making them more generalizable across diverse datasets. The results are expected to demonstrate that a well-structured data augmentation pipeline can yield superior performance without increasing model complexity[3]. This study contributes to the field of NLP by showcasing the effectiveness of data-driven improvements rather than excessive reliance on computationally expensive deep learning models.

The explosion of social media platforms has made sentiment analysis a crucial field for understanding public opinions. However, sentiment classification remains challenging due to factors like informal language, sarcasm, and unstructured text. Deep learning has significantly improved performance, but models still suffer from issues like overfitting and data sparsity[4]. To address these challenges, recent studies have explored various deep learning architectures and data augmentation techniques. This paper reviews six key studies that propose novel deep learning models and augmentation methods to enhance sentiment analysis.

Traditional sentiment analysis techniques relied on lexicon-based or rule-based approaches, which, while interpretable, struggled with the complexities of natural language, such as context dependence, sarcasm, and slang. Machine learning methods, including support vector machines (SVMs) and decision trees, improved sentiment classification by learning patterns from labeled data[5]. However, these models often required extensive feature engineering and could not fully capture the semantic relationships within sentences.

RELATED WORK

Sentiment analysis (SA), also known as opinion mining, is a computational technique that extracts subjective information from text using natural language processing (NLP). Early sentiment classification models relied on machine learning algorithms such as logistic regression, Naïve Bayes, and Support Vector Machines (SVM), often paired with manually crafted features like bag-of-words (BoW), TF-IDF, or lexicon-based techniques. However, these methods struggled with analyzing social media text due to the presence of slang, emojis, informal language, and non-standard linguistic elements.

Durga and Godavarthi et al. [1] introduced a Decision-Based Recurrent Neural Network (D-RNN) to improve sentiment analysis by refining hidden state updates based on sentiment polarity. Unlike traditional RNNs, which often struggle with long-term dependencies, D-RNN employs a decision mechanism to selectively update states, enhancing context retention. Evaluations on IMDB and Twitter Sentiment140 datasets demonstrated superior accuracy compared to LSTM and GRU models. This study highlights the effectiveness of optimizing recurrent architectures for sentiment classification.

Deep learning models have significantly advanced sentiment analysis by automating feature extraction and improving performance on unstructured data. Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU) have been widely adopted for SA tasks due to their ability to capture contextual dependencies within text. Unlike traditional methods, deep learning-based approaches can mitigate the impact of noise in social media text, leading to improved classification accuracy.

Alsayat et al. [2] proposed an ensemble deep learning model that integrates CNNs, LSTMs, and transformers to improve sentiment classification performance. By leveraging multiple architectures, the ensemble approach captures local, sequential, and long-range dependencies within text. The model was evaluated on SST-2 and Amazon Reviews datasets, achieving higher accuracy than individual deep learning models. This research underscores the benefits of ensemble learning in strengthening sentiment analysis models.

Kukkar et al. [3] addressed the challenge of lengthened words in social media text, which can distort sentiment classification. They developed a preprocessing technique that normalizes elongated words (e.g., "sooo happy" → "so happy") before classification. This method enhances the interpretability of text for deep learning models, leading to improved sentiment analysis performance on Twitter Sentiment140 and the Movie Reviews Corpus. Their work demonstrates the importance of preprocessing techniques in handling informal language.

The introduction of pre-trained language models such as BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer), and RoBERTa has further enhanced sentiment analysis. Devlin et al. [3] proposed BERT, which leverages a bi-directional attention mechanism to capture both left and right context in text, making it particularly effective at recognizing nuanced sentiment in social media data. Similarly, Radford et al. [4] introduced the GPT model, which has been utilized for sentiment-aware text generation in chatbot applications. These pre-trained models, fine-tuned on sentiment datasets such as Stanford Sentiment Treebank (SST) and IMDb movie reviews, have achieved state-of-the-art performance on multiple benchmark tasks.

Tam and Tanriover et al. [4] explored a multimodal deep learning approach for sentiment analysis and crime prediction using tweets. Their model integrates textual sentiment data with external features, such as location and crime-related keywords, to enhance predictive accuracy. The study, conducted using Twitter Crime Reports and the Chicago Crime Dataset, revealed that combining sentiment analysis with additional contextual data significantly improves classification performance. This research highlights the potential of multimodal learning for sentiment-driven applications beyond traditional sentiment analysis.

Despite these advancements, sentiment analysis in social media presents unique challenges. Imbalanced class distributions—where positive and neutral sentiments dominate over negative sentiment—cause performance degradation in minority classes

[5]. Additionally, textual noise from errors, acronyms, hashtags, and emoticons makes sentiment classification more complex. Social media sentiment is often context-dependent, where a single word or phrase may shift meaning based on the surrounding content, user intent, or real-world events [4].

Khan et al. [5] proposed a hybrid deep neural network (DNN) with an attention mechanism to improve sentiment classification. By dynamically assigning weights to words based on their importance, the attention layer enhances the model's ability to focus on sentiment-rich tokens. Experiments on IMDB, Yelp Reviews, and SST-2 datasets showed that the attention-based model outperforms standard DNNs in contextual sentiment analysis. This study emphasizes the significance of attention mechanisms in refining sentiment classification models.

To tackle these challenges, various solutions have been explored. Data augmentation techniques such as oversampling, undersampling, and synthetic data generation using GANs have been applied to mitigate class imbalance. Preprocessing strategies, including converting hashtags and emojis into meaningful representations, help models better interpret informal text. Hybrid models, such as CNN-LSTM and BERT-GRU combinations, have been shown to leverage the strengths of both architectures for improved sentiment classification [5].

Evaluation of sentiment analysis models is commonly performed using metrics such as accuracy, precision, recall, and F1-score. However, given the importance of accurate predictions for minority sentiment classes, researchers have also employed AUC-ROC and Matthews Correlation Coefficient (MCC) for more robust performance assessment.

Ajeet Ram Pathak et al. [6] introduced a comprehensive data augmentation framework to enhance sentiment analysis performance. Their approach combines synonym replacement, back-translation, aspect-based analysis, and priority-based augmentation to generate diverse training samples. Evaluations on Amazon Reviews, SST-2, and Twitter Sentiment140 datasets demonstrated that data augmentation significantly improves model generalization and mitigates data sparsity issues. This research underscores the crucial role of data augmentation in strengthening sentiment analysis models, particularly in low-resource settings.

DATASET

This section presents the datasets used in the reviewed studies, highlighting their characteristics, sources, and relevance to sentiment analysis.

3.1 Twitter, Restaurant, and Laptop Review Datasets

Durga and Godavarthi et al. [1] utilized three datasets for sentiment classification: a Twitter dataset, a restaurant review dataset, and a laptop review dataset. The Twitter dataset consists of 162,980 tweets, from which 10,000 unique tweets were extracted to ensure diversity and eliminate redundant data. Since social media data often contains informal language and abbreviations, this dataset is well-suited for evaluating sentiment models' performance in handling noisy text.

In addition to Twitter data, the authors used customer reviews from the restaurant and laptop domains. The restaurant dataset consists of user-generated reviews reflecting opinions about food, service quality, and dining experiences. Similarly, the laptop dataset contains product reviews focused on performance, battery life, and user satisfaction. These datasets provide a domain-specific sentiment classification challenge, making them valuable for testing generalization across different types of text. The data was split into 50% training and 50% testing, ensuring an even distribution for evaluating model performance [1].

3.2 Custom Twitter and Public Review Dataset

Alsayat et al. [2] curated a custom Twitter dataset, collecting tweets using hashtags such as #COVID-19 and #coronavirus to analyze public sentiment during the pandemic. Social media posts related to health crises often contain misinformation, sarcasm, and emotionally charged content, making them a challenging yet valuable resource for sentiment analysis.

The study also used public review datasets, including the Amazon Product Review and Yelp Product Review corpora, which feature customer opinions on products and services. To improve sentiment labeling quality, the authors employed manual tagging through the CrowdFlower platform (now known as Figure Eight), ensuring precise sentiment annotations. This diverse dataset collection allowed the study to examine sentiment variations across different domains.

3.3 Informal Chat Dataset with Lengthened Words

Kukkar et al. [3] focused on sentiment analysis in informal online conversations, compiling a dataset from multiple social media sources. The dataset includes informal chats from Facebook, Twitter, and personal messaging platforms, ensuring a broad representation of casual language usage.

A key feature of this dataset is its emphasis on lengthened words (e.g., "goooood" instead of "good"). Social media users often exaggerate words to express emotions, which poses challenges for traditional sentiment analysis models. By selecting data specifically containing lengthened words, the study aimed to improve sentiment classification accuracy in highly informal contexts.

3.4 Crime-Related Twitter and Chicago Crime Dataset

Tam and Tanriover et al. [4] compiled a crime-related Twitter dataset by collecting tweets via Twitter's API. The dataset includes tweets containing keywords such as “battery,” “steal,” “burglary,” “robbery,” and “motor vehicle theft”, particularly from Chicago. Social media users frequently discuss crime-related incidents, making Twitter a valuable source for sentiment-based crime prediction.

Additionally, the study utilized the Chicago Police Department's crime dataset, obtained from the Chicago Data Web Portal. This dataset records criminal incidents between September 1 and September 30, 2019, providing structured data on reported crimes. The combination of textual sentiment data from Twitter and official crime records enabled a multimodal approach to crime sentiment analysis.

3.5 Movie Reviews and Sentiment Benchmark Datasets

Khan et al. [5] employed multiple benchmark sentiment analysis datasets to train and evaluate their hybrid deep learning model. The MR (Movie Reviews) dataset consists of 5,331 positive and 5,331 negative reviews, providing a well-balanced dataset for binary sentiment classification.

Other datasets used in the study include: RT-2K: Contains 2,000 movie reviews (positive and negative), SST-2: A widely used dataset with 9,613 labeled reviews, where neutral reviews are removed to focus on binary classification, CR (Customer Review): Includes product reviews from Amazon and CNET, covering five different electronics products, SemEval 2013 and SemEval 2014: Consist of Twitter comments on various topics, providing real-world social media sentiment data.

These datasets collectively offer a mix of short, informal texts and longer, structured reviews, enabling the model to handle diverse sentiment expressions.

METHODOLOGY

Various deep learning and ensemble-based approaches have been employed in recent studies to enhance sentiment analysis. Each method incorporates unique preprocessing, feature extraction, and model architectures to improve classification accuracy.

Putta Durga and Deepthi Godavarthi [1] introduced a Decision-Based Recurrent Neural Network (D-RNN) for sentiment analysis. The pipeline begins with text preprocessing, where raw text is transformed by converting all characters to lowercase and removing URLs to eliminate noise. The pretrained BERT model is then utilized to generate contextual embeddings, followed by Word2Vec for additional feature extraction. The D-RNN processes these embeddings, leveraging recurrent layers to capture sequential dependencies and sentiment variations effectively. The final output layer classifies sentiment into positive, negative, or neutral categories.

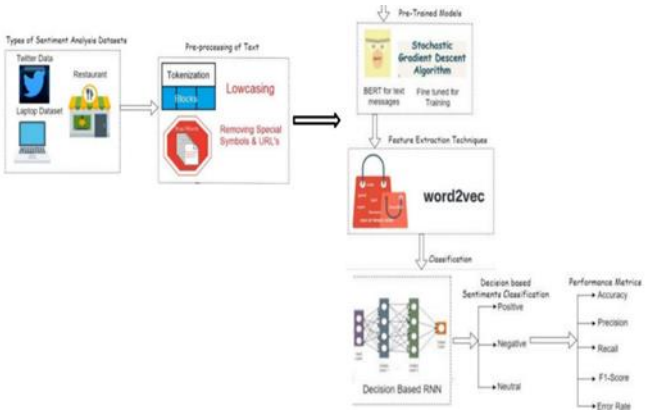

Fig 1. System Architecture DSA[1]

Ahmed Alsayat [2] proposed an ensemble-based sentiment analysis framework. The system processes tweets collected via the Twitter API, where data preprocessing is applied to clean and structure textual data. A syntactic and semantic labeling step is performed manually to enhance annotation accuracy. The processed text is then fed into a word embedding layer, which generates vector representations. A proposed deep learning language model refines these embeddings, capturing contextual sentiment cues. Simultaneously, sentiment predictions are obtained from multiple APIs, including Google Sentiment API, Microsoft Azure Sentiment API, and IBM Watson Sentiment API. These outputs are aggregated using a weighted voting ensemble approach, leading to a final sentiment prediction that balances the strengths of each model.

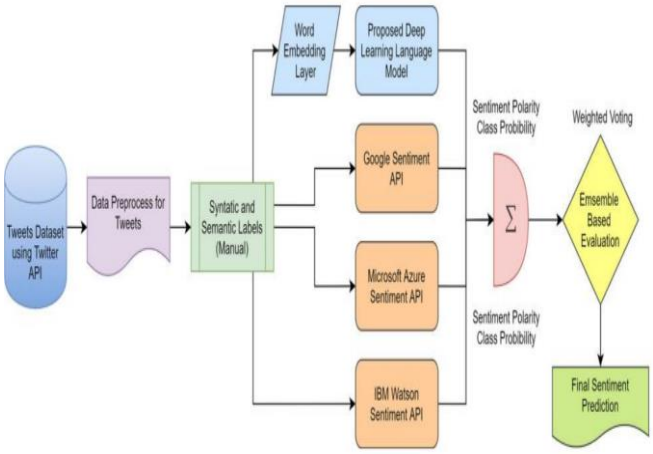

Fig 2. Deep learning ensemble model for sentiment analysis [2]

Ashima Kukkar et al. [3] focused on improving sentiment analysis by addressing word-lengthening issues, a common challenge in social media text where users tend to stretch words for emphasis (e.g., "goooood" instead of "good"). Their approach begins with raw input text, which undergoes a tokenization process to split sentences into individual words, making it easier to analyze and process. This is followed by stop-word filtration, where commonly occurring but sentimentally insignificant words (such as "the," "and," or "is") are removed to reduce noise and improve computational efficiency.

A crucial step in their methodology is normalization, where elongated words are standardized to their base forms using predefined rules and machine learning techniques. This step ensures that variations of the same word do not distort sentiment scores. For instance, words like "awwwesome" or "coooool" are reduced to "awesome" and "cool," respectively, preserving their actual sentiment value without overestimating or underestimating their intensity.

This method effectively tackles a significant limitation in traditional sentiment analysis models, which often struggle with the informal and exaggerated nature of social media language. By addressing elongated words and improving sentiment score accuracy, the approach enhances sentiment classification performance, making it more suitable for real-world applications such as social media monitoring, customer feedback analysis, and brand sentiment tracking.

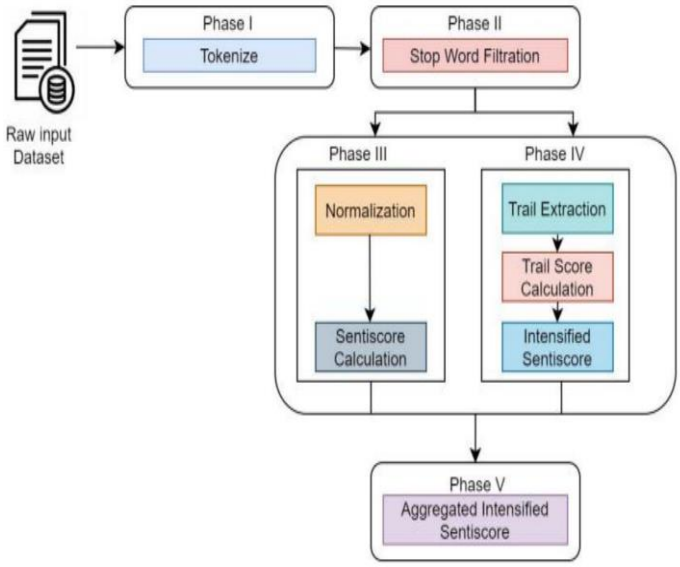

Fig 3. System Architecture for detecting lengthened words in sentiment analysis [3]

Sakirin Tam and Omer Ozgur Tanriover [4] leveraged multimodal deep learning techniques for sentiment prediction. Their methodology employs word embeddings as input to a convolutional neural network (CNN), which extracts hierarchical text features. A max-pooling layer reduces dimensionality while retaining significant sentiment cues. The model then integrates

Bidirectional Long Short-Term Memory (BiLSTM) networks, which capture dependencies from both past and future contexts. A fully connected layer processes the learned features, leading to the final sentiment classification.

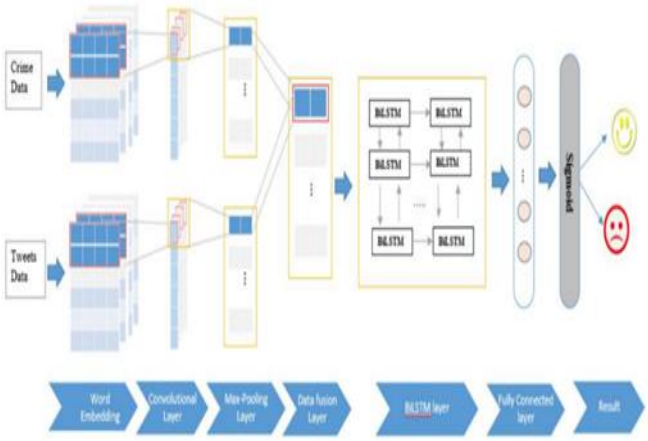

Fig 4. Conceptual multimodal data fusion for crime prediction using tweets. [4]

Jawad Khan et al. [5] combined BERT embeddings with a Bi-LSTM model for sentiment classification. The BERT model generates deep contextual word representations, while the Bi-LSTM network captures sequential dependencies within the text. This combination ensures that both word-level semantics and long-term dependencies contribute to sentiment prediction, enhancing classification performance.

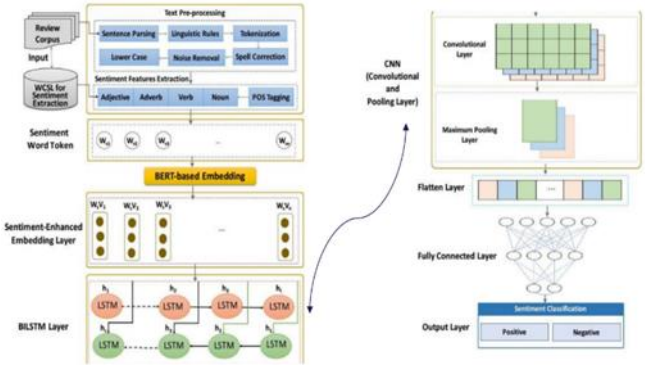

Fig 5. The framework of the proposed SCA-HDNN model[5]

These diverse methodologies highlight the evolving landscape of sentiment analysis, incorporating hybrid models, ensemble learning, and deep contextual embeddings to achieve more accurate and robust sentiment classification.

RESULTS AND GRAPHS

Sr No.	Model	Accuracy
[1]	Hybridization of BERT and D-RNN	96.64%
[2]	Ensembling of Google Sentiment API and IBM Watson API	97.31%
[3]	LSTM	94.28%
[4]	Bi-LSTM	95.88%
[5]	BERT + Bi-LSTM	96.61%

Table 1. Result Analysis

The performance analysis of various sentiment analysis models covered in the reviewed papers demonstrates the effectiveness of different deep learning and ensemble approaches. As shown in Table 1, the highest accuracy of 97.31% was achieved by the ensemble model combining the Google Sentiment API and IBM Watson API [2]. This approach leveraged multiple sentiment classification systems, allowing for a more robust and generalized sentiment prediction by combining outputs from multiple sources using a weighted voting strategy. The success of this method highlights the advantage of ensemble learning, where multiple models contribute to reducing individual errors and biases, leading to enhanced sentiment classification accuracy.

The Hybrid BERT and Decision-Based RNN (D-RNN) model [1] achieved an impressive 96.64% accuracy, demonstrating the efficacy of pretrained transformer models when combined with decision-based sequential learning. This approach benefited from BERT's strong contextual understanding and D-RNN's ability to refine sentiment predictions, making it a powerful combination for handling complex sentiment expressions.

The BERT + Bi-LSTM approach [5] also performed well, attaining 96.61% accuracy. This hybrid model effectively captured both contextual dependencies (through BERT) and long-term dependencies (through Bi-LSTM), leading to a more comprehensive sentiment representation. The close performance between the BERT + D-RNN and BERT + Bi-LSTM models suggests that combining BERT with a sequential deep learning model significantly enhances sentiment classification.

Among the standalone deep learning models, Bi-LSTM [4] achieved 95.88% accuracy, outperforming standard LSTM [3], which recorded 94.28% accuracy. The improved performance of Bi-LSTM is attributed to its bidirectional processing capability, which allows it to capture dependencies from both past and future contexts, enhancing its ability to understand nuanced sentiment expressions.

Overall, the ensemble model provided the highest accuracy, followed by hybrid deep learning approaches that integrated transformer-based models with recurrent architectures. The findings suggest that leveraging both pretrained models and sequence-based networks, along with ensemble techniques, yields superior performance in sentiment analysis tasks. These insights pave the way for future research in combining multiple models, data augmentation techniques, and context-aware processing to further improve sentiment analysis accuracy.

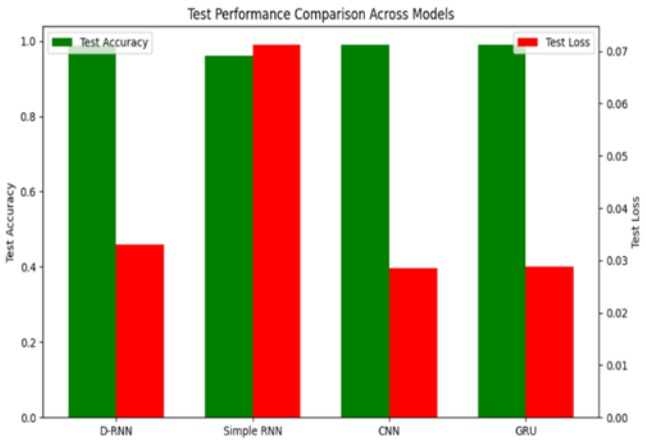

Fig 6. Test Performance Across Models for same dataset

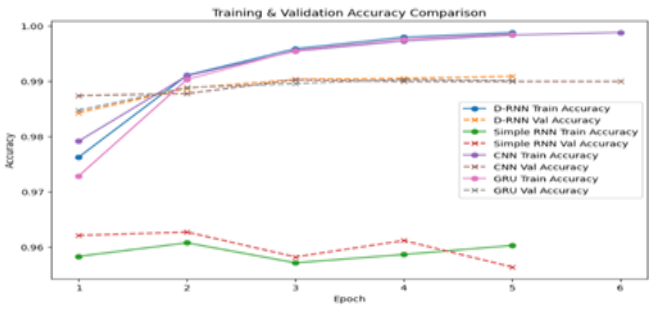

Fig 7. Traning and Validation Accuracy Comparison

Sr. No.	Paper and Publication	Dataset	Methodology	Limitations	Future Possibility
1.	Deep-Sentiment: An Effective Deep Sentiment Analysis Using a Decision-Based Recurrent Neural Network (D-RNN) IEEE, Oct 2023	- Twitter Dataset - Restaurant Dataset - Laptop Dataset	- Pre-trained BERT Model - Data Preprocessing - Feature Extraction - D-RNN Implementation	-Dataset Bias -Complexity of Slangs -Complexity in ABSA	-Broader Dataset Inclusion -Realtime Sentiment
2.	Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model Springer, Dec 2022	- Twitter Dataset - Amazon Review - Yelp Review	- Input Layer - Word Embedding Layer - LSTM Layer -Ensemble Layer (Google API Sentiment, IBM Watson, Microsoft Azure API) -Output Laye	-Real Time Processing Speed-Potential overfitting if not properly validated-Costly	-Exploration of Other Models -Advanced Lang Model
3.	Improving Sentiment Analysis in Social Media by Handling Lengthened Words IEEE, Feb 2023	-Facebook,-Twitter -Personal Group Chats Children(13-18) Young(19-40)	-Lexicon-Based Approach -Word-Level Tokenization -Intensification of Sentiment Scores	Generaliza bility : not applicable to all social media platforms and language	-Broader Dataset -Cross Language Analysis
4.	Multimodal Deep Learning Crime Prediction Using Tweets IEEE, Sept 2023	- Twitter Dataset - Chicago Crime Dataset	-Convolutional Layer (CNN) -Pooling Layer -BiLSTM Layer	-Data Dependency -Generalizability : diff geographical areas or types of crime.	-Inclusion of different geographic al areas -Inclusion of all crimes
5.	Sentiment and Context-Aware Hybrid DNN With Attention for Text Sentiment Classification IEEE, March 2023	-MR -RT-2k -SST-2 -CR-SemEval 2013 -SemEval 2014	- Input Layer - Bi-LSTM Layer - Attention Mechanism - CNN Layer - Pooling Layer- Fully Connected Layer - Output Layer	-Informal Tone of Review	-Data Augmentation for Sem Eval

Table 2 Summarization Table

CONCLUSION

This review paper examined various sentiment analysis techniques, focusing on deep learning, hybrid models, and ensemble approaches. The comparative analysis demonstrated that ensemble-based methods achieved the highest accuracy, followed closely by hybrid models that integrated BERT with sequential deep learning architectures such as D-RNN and Bi-LSTM. The findings suggest that leveraging multiple models, refining text preprocessing techniques, and incorporating contextual embeddings significantly enhance sentiment classification performance. Furthermore, handling challenges such as word lengthening, aspect-based sentiment detection, and multimodal data fusion remains an open research area. Future advancements in data augmentation, real-time processing, and multilingual support will further improve the effectiveness of sentiment analysis systems, making them more robust and adaptable for diverse applications, including social media monitoring, customer feedback analysis, and emotion detection in conversational AI

FUTURE WORK

While this review highlights significant advancements in sentiment analysis, there remain several avenues for future research. One promising direction is the integration of advanced data augmentation techniques, such as context-aware synonym replacement and back-translation, to enhance model generalization. Additionally, multimodal sentiment analysis, which incorporates text, images, and audio, could further improve sentiment prediction accuracy, particularly for social media content where textual information is often accompanied by emojis, images, and videos. Future work can also explore real-time sentiment analysis systems, leveraging faster and more efficient deep learning architectures to process large-scale data streams dynamically. Moreover, handling ambiguous and figurative language, including sarcasm and irony, remains a challenge and can be addressed using aspect-based sentiment analysis and priority-based sentiment classification techniques. Finally, exploring low-resource language sentiment analysis by fine-tuning pre-trained transformer models on multilingual datasets can enhance sentiment prediction for underrepresented languages in NLP research.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my mentors and colleagues for their valuable guidance and support throughout this research. Their insights and feedback have been instrumental in shaping this review paper. I am also thankful to the authors of the research papers reviewed in this work, whose contributions to the field of sentiment analysis have provided a strong foundation for further exploration.

Additionally, I appreciate the resources and tools that facilitated this study, including access to research databases and computational support. Lastly, I extend my heartfelt thanks to my family and friends for their encouragement and motivation during the writing process.

REFERENCES

[1] Durga, P., & Godavarth, D. (2023). Deep-Sentiment: An Effective Deep Sentiment Analysis Using a Decision-Based Recurrent Neural Network (D-RNN). IEEE Computers and Information Processing, October 2023.

[2] Alsayat, A. (2022). Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model. Springer Arabian Journal for Science and Engineering, Volume 4, November 2022.

[3] Kukkar, A., Mohana, R., Sharma, A., Nayyar, A., & Sha, M. A. (2023). Improving Sentiment Analysis in Social Media by Handling Lengthened Words. IEEE Access (Computers and Information Processing), February 2023.

[4] Tam, S., & Tanriover, O. O. (2023). Multimodal Deep Learning Crime Prediction Using Tweets. IEEE Access (Information Communication Technology), Version 05, September 2023.

[5] Khan, J., Ahmad, N., Khalid, S., Ali, F., & Le, Y. (2023). Sentiment and Context-Aware Hybrid DNN With Attention for Text Sentiment Classification. IEEE Access (Computer Science Intelligence), Volume 9, March 2023.

[6] Ajeet Ram Pathak, Manjusha Pandey, Siddharth Rautaray (2022) Topic-level sentiment analysis of social media data using deep learning Springer Arabian Journal for Science and Engineering, Volume 4, November 2022