

Assignment Report: Deep Learning Techniques for Sentiment Analysis of Users on Social Media.

Student name: Wiam Alharbi.

Lecturer: Dr. Mohammed Anbar

Date: 28 Feb 2025

Table of contents:

- Summary research papers
- Critical Analysis
- Connections
- Conflicts between the findings
- Synthesis Matrix for Literature Review
- References

Summary on research papers:

Paper 1: Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model

In this paper, the author introduces an improved ensemble deep learning model designed to handle sentiment analysis more effectively. They enhanced the sentiment classification accuracy by combining advanced word embedding techniques with Long Short-Term Memory (LSTM) networks. The model was tested on several benchmark datasets, including a Twitter dataset focused on the COVID-19 pandemic, where it outperformed existing state-of-the-art sentiment analysis models.

The paper makes two key contributions:

1. It presents a strong framework that leverages word embeddings and an LSTM network to understand contextual relationships between words. This approach is particularly useful for recognizing rare words in emerging situations—like the COVID-19 pandemic—by identifying suffixes and prefixes from training data.
2. It improves sentiment analysis by introducing a hybrid ensemble model that effectively captures differences among state-of-the-art methods.

To validate their approach, the authors conducted multiple experiments using their own Twitter coronavirus hashtag dataset, along with publicly available review datasets from Amazon and Yelp.

Paper 2: Hybrid Deep Learning Framework for Sentiment Analysis of Social Media Data

This paper explores the use of hybrid deep learning models for sentiment analysis of social network data. The authors tested the effectiveness of combining Support Vector Machine (SVM), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) models while incorporating two word embedding techniques—Word2Vec and BERT. To evaluate adaptability, they conducted experiments on eight different datasets, including tweets and reviews, aiming to determine whether hybrid models could perform consistently across varying dataset types and sizes. The study also examined how different datasets, feature extraction methods, and deep learning architectures influenced the reliability of sentiment polarity analysis. By comparing the performance of hybrid models with standalone models, the research highlights the potential advantages of integrating multiple approaches for more robust sentiment classification.

Paper 3: Comparative Analysis of Deep Learning Models for Sentiment Classification on Social Media

This paper presents a comparative analysis of deep learning models for sentiment classification, evaluating eight different models—three Convolutional Neural Networks (CNNs) and five Recurrent Neural Networks (RNNs). The study explores two types of input structures: word-level and character-level embeddings, testing their impact on sentiment classification across 13 review datasets. The results reveal that model architecture and input representation play a crucial role in performance, providing valuable insights into selecting the most effective approach for sentiment analysis in different domains. However, the study also

acknowledges certain limitations: 1) model performance varies depending on dataset type and size, and 2) the findings are specific to sentiment polarity classification, making them less applicable to more complex tasks such as sarcasm detection or aspect-based sentiment analysis.

Paper 4: Emotion Recognition from Social Media Text Using Deep Learning

This paper explores a deep learning approach to identifying emotions in social media comments, specifically targeting joy, anger, and sadness. The method combines Latent Dirichlet Allocation (LDA) with an emotion dictionary to extract key emotional features, which are then converted into word vectors using Word2Vec. These vectors are fed into a CNN-BiLSTM model, where CNN identifies essential text patterns, while BiLSTM captures long-range dependencies to improve sentiment classification. When tested on Weibo comments, the model demonstrated strong performance, surpassing traditional machine learning methods by 3.66% in F1 score and outperforming deep learning models by 1.84%. These findings highlight its effectiveness in accurately detecting emotions in social media conversations.

Paper 5: Aspect-Based Sentiment Analysis Using Deep Learning

This paper introduces a novel approach to aspect-level sentiment analysis by integrating user-specific data with advanced deep learning techniques. The authors propose the U-ASGCN model, which combines User Comment Feature Extraction (UCFE) with Aspect-Sentence Graph Convolutional Networks (ASGCN) to enhance sentiment prediction. The core idea is straightforward yet impactful: by analyzing both user behavior and sentence context, the model makes more informed sentiment predictions. UCFE extracts patterns from a user's past comments to understand their sentiment tendencies, embedding this information into ASGCN, which identifies connections between aspect words (e.g., “price” or “service”) and overall sentiment. This fusion enables the model to handle complex sentiment scenarios where word-level analysis alone falls short. When tested on benchmark datasets like Yelp and SemEval, the model outperformed BERT and BiLSTM, achieving higher accuracy and F1 scores. These results highlight the importance of incorporating user context into sentiment analysis, marking a significant advancement in the field.

Paper 6: Dynamic Light Weight Recommendation System for Social Networking Analysis Using a Hybrid LSTM-SVM Classifier Algorithm

This paper presents an improved deep learning model for sentiment analysis on social media. By combining LSTM networks with advanced word embeddings, it enhances classification accuracy. Tested on datasets like Twitter (COVID-19 related), Amazon reviews, and Yelp, the model outperforms existing methods.

The paper makes two key contributions:

- 1) Context-Aware Sentiment Detection – Uses LSTM and word embeddings to better capture word relationships, especially in fast-changing topics like COVID-19.
- 2) Hybrid Deep Learning Approach – Integrates multiple models to improve accuracy and consistency across different datasets.

Extensive experiments confirm that this hybrid model is more effective than traditional sentiment analysis methods, particularly for evolving language patterns.

Critical Analysis:

Research on sentiment analysis in social media using deep learning has shown promising results but also presents challenges. Studies consistently find that hybrid deep learning models perform better than single-method approaches, offering higher accuracy and a better grasp of context. These models have been tested on real-world data from platforms like Twitter, Yelp, and Amazon, reinforcing their reliability. However, a major challenge is choosing the right word representation technique. While advanced methods like BERT improve accuracy, they require significant computational resources, making simpler approaches like Word2Vec more practical in some cases. The scope of sentiment analysis also varies widely, from basic positive/negative classification to deeper tasks like emotion detection and aspect-based sentiment analysis. Another ongoing debate is whether including user-specific data improves accuracy—some studies show clear benefits, while others achieve strong results without it.

Despite these advancements, the field still faces issues like reproducibility, dataset accessibility, and the clarity of model parameters. While research does a good job of testing models across multiple datasets, there's a need for more transparency and better documentation to make studies easier to replicate. Future efforts should focus on refining comparative analyses, balancing model complexity with efficiency, and ensuring sentiment analysis techniques are both powerful and practical for real-world applications

Connections:

1. **Focus on Hybrid Deep Learning Models:** All reviewed papers emphasize the superiority of hybrid or ensemble deep learning models over traditional standalone methods for sentiment analysis tasks, particularly on social media datasets.
2. **Word Embedding Importance:** Multiple studies highlight the critical role of advanced word embedding techniques such as Word2Vec and BERT in capturing semantic relationships, context, and evolving language, particularly in rapidly changing scenarios such as COVID-19-related social media data.
3. **Extensive Experimental Validation:** Each paper reinforces the importance of rigorous experimentation across multiple benchmark datasets, including platforms like Twitter, Yelp, and Amazon, demonstrating the necessity for adaptability and consistency in sentiment analysis.
4. **Context and User-specific Features:** Several articles underline the importance of context-aware sentiment detection, either through user behavior analysis or emotional context, improving the depth and accuracy of sentiment classification.

Conflicts between the findings:

- **Model Selection and Architecture:** While some studies advocate strongly for hybrid approaches integrating CNN, LSTM, and SVM, others find specific standalone architectures (e.g., certain RNN or CNN variants) sufficient or superior in certain contexts.
- **Embedding Methods:** There's variance in conclusions regarding the most effective word embedding techniques; while some results favor the advanced contextual embeddings provided by BERT, others highlight traditional embeddings (Word2Vec) as more effective or computationally efficient.

- **Granularity of Analysis:** Discrepancies exist between studies focusing strictly on sentiment polarity classification versus those extending their analyses to finer-grained tasks like emotion detection and aspect-based sentiment analysis.
- **Incorporation of User Context:** While certain studies strongly advocate the integration of user-specific data (such as historical comment analysis), others achieve competitive results without explicitly considering user-level contextual data.
- **Emotion Detection vs. General Sentiment Analysis:** Differences emerge regarding the suitability of deep learning methods optimized specifically for emotion detection versus broader sentiment analysis approaches. Some studies suggest highly specialized models significantly outperform generalized ones, while others propose flexible hybrid models as universally effective.

Synthesis Matrix for Literature Review:

Study	Authors	Year	Research Design	Key findings/Result	Methodology	Limitations
Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model	Ahmed Alsayat	2022	Experimental	The research demonstrated that the proposed ensemble deep learning model, which combined FastText word embedding and LSTM networks, significantly improved sentiment analysis performance on social media platforms. The optimal model configuration, which included 200 neurons and two hidden layers, achieved the highest classification accuracy across multiple datasets, including Twitter (COVID-19), Yelp, and Amazon reviews. The ensemble approach, which integrated predictions from custom deep learning models	Data Collection & Preprocessing: The dataset was gathered using Twitter's API, focusing on COVID-19-related tweets, manually labeled for sentiment. Additional data from Amazon, Yelp, and Web 2.0 platforms (Twitter, MySpace, YouTube, BBC) were included for comparison. Cleaning involved removing hashtags, URLs, usernames, and irrelevant content. Emojis were assigned	Dependence on Large, Labeled Datasets: The model's rely on complex labeled datasets for training it can limit its effectiveness when applied to smaller datasets. High Computational Requirements: Training the model on large datasets demands computational resources, with training times go up to 60 hours, making it challenging for all users. Model Complexity and Scalability Issues: The ensemble approach, combining multiple classifiers, contributes to higher computational costs and longer training times,

along with Google, Microsoft, and IBM sentiment analysis APIs, provided a 2-5% improvement in classification accuracy compared to individual models. Statistical significance testing confirmed that the proposed model outperformed existing methods, with p-values showing strong evidence of its superiority. The model also demonstrated efficiency in terms of runtime, with the prediction time for a single record averaging 10 seconds. Overall, the study confirmed that combining various deep learning techniques, including word embedding and LSTM, alongside ensemble learning, significantly enhanced sentiment analysis accuracy, especially for emerging topics like the COVID-19 pandemic.

sentiment values to capture emotional expression.

Word Embedding (FastText):
FastText was used to handle social media text by breaking words into smaller character units, improving recognition of variations like "COVID19" and "COVID-19." This helped the model understand misspellings and rare words.

Model Architecture (LSTM):
The model used Long Short-Term Memory (LSTM) networks, which are effective for processing sequential text. LSTMs helped capture word dependencies and negations, improving sentiment classification.

Ensemble Learning & Hyperparameter Tuning:
To boost accuracy, multiple classifiers were combined using an ensemble approach. Hyperparameters were

potentially shackling the scalability for real time applications.

Limited Generalization Across Domains: While the model performs well on social media data, its generalization to other types of unstructured text or domains with different languages has not been thoroughly evaluated.

Focus on English-Language Datasets: The research primarily focuses on the English language datasets, limiting its applicability for multilingual sentiment analysis.

					fine-tuned through k-fold cross-validation, adjusting neuron counts (100-300) and hidden layers (1-3). Statistical Validation: Performance was tested using a binomial test, comparing results with models from Google, Microsoft, and IBM. Statistically significant p-values confirmed the ensemble model's superiority. Performance Evaluation: Training times varied—2.5 hours for Twitter COVID-19 data, up to 60 hours for Amazon reviews. Predictions averaged 10 seconds per record, making the model efficient for large-scale sentiment analysis.	
Hybrid Deep Learning Framework for Sentiment Analysis of Social Media Data	Cach N. Dang, María N Moreno-García, Fernando De la Prieta.	2021	Experimental	The key findings on the paper highlights the effectiveness of hybrid deep learning models for sentiment analysis. The experiments showed that combining SVM, CNN, and LSTM models resulted in better performance compared to individual models, which show the advantages of combining multiple	Data Collection: The main dataset was collected using Twitter's API with tweets on COVID-19, manually labeled for sentiment. They also added more datasets from Amazon and Yelp reviews, as well as Web 2.0 data from platforms like Twitter,	Computational Complexity: it combines CNN, LSTM, and SVM, which will requires significant computational resources, leading to longer training times. Overfitting Risk: With the increased complexity of the hybrid model, the system is more prone to

techniques to capture different aspects of the data. The two word embedding methods, Word2Vec and BERT, both contributed to improving the model performance, with BERT showing excellent results due to its deep framed word representations. Also, the hybrid models proved to be adaptable across various datasets of different sizes, maintaining a strong performance on both small and large datasets. The study also focused on the choice of dataset and feature extraction techniques that significantly influence the reliability of sentiment polarity analysis. Overall, the hybrid approach enhanced the accuracy and reliability of sentiment analysis, effectively combining the strengths of each model to handle the complexities of social media data.

MySpace, YouTube, BBC, were used for comparison for more accuracy.

Data Preprocessing: Data was cleaned by removing hashtags, URLs, usernames, and irrelevant content. Emojis were assigned sentiment values to capture emotional expression, which is crucial for social media analysis.

Word Embedding (Word2Vec and BERT): The study used Word2Vec and BERT embeddings to represent words contextually, which will improve the model's understanding of language nuances and ambiguities in social media data.

Model Architecture (SVM, CNN, and LSTM): A hybrid model combined SVM, CNN, and LSTM to handle high-dimensional classification (SVM), local text features (CNN), and sequential dependencies (LSTM),

overfitting, particularly with small or imbalanced datasets.

Domain Dependency: The performance of the model may across many different domains, especially when working with social media data like slangs.

Dependence on Data Quality: The accuracy of the hybrid model heavily relies on the quality and preprocessing of the data.

					optimizing sentiment analysis accuracy. This approach was chosen because each model brings unique strengths: SVM efficiently handles complex feature spaces, CNN captures local patterns in text, and LSTM models sequential data and context, making it ideal for analyzing social media content where word order and context are crucial. Performance Evaluation: The model was evaluated based on accuracy and runtime. Training times ranged from 2.5 hours for Twitter data to 60 hours for larger datasets. Prediction times averaged 10 seconds per record.	
Comparative Analysis of Deep Learning Models for Sentiment Classification on Social Media	Seungwan Seo, Czangyeob Kim, Haedong Kim, Kyoung Hyun	2020	Experimental	The research demonstrated that RNN-based models generally outperformed CNN-based models when using word-level inputs, with the bidirectional LSTM model showing the best results across most datasets. In	Data Collection: The study used 13 datasets, including product and service reviews from platforms like Amazon and Stanford, with sentiment labels	A key limitation is the dependence on large, labeled datasets, which may not be available in all cases, affecting model accuracy when data is limited or inadequately labeled.

	Mo, Pilsung Kang			contrast, CNN models performed better when using character-level inputs, particularly the 29-layer CNN. Additionally, the study revealed that larger training datasets led to improved sentiment classification performance. While increasing the number of layers in CNN models didn't always enhance their performance, simpler, shallow CNN models sometimes performed better. For RNN models, adding more layers and incorporating LSTM or GRU units consistently improved the classification results. The study found that word-level inputs were more effective for RNN models, while deeper CNN architectures showed better performance with character-level inputs. The combination of bidirectional LSTM with word-level input was the most effective model, yielding the best performance for five out of the 13 datasets used in the study.	indicating positive or negative feedback. Data Preprocessing: The data was split into training, validation, and test sets in a 50%-20%-30% ratio. To handle class imbalances, the training batches were balanced with equal numbers of positive and negative reviews. Model Architecture: Eight deep learning models were tested: three CNN models (one-layer, nine-layer, and 29-layer) and five RNN models (Vanilla RNN, LSTM, GRU, bidirectional LSTM, and bidirectional GRU), using both word-level and character-level inputs. Model Training and Optimization: The models were trained using the Adam optimizer for CNN models and the RMSProp optimizer for RNN models to ensure efficient learning and performance across different datasets.	The study also notes high computational demands, especially for complex models like 29-layer CNNs and bidirectional RNNs, which require extensive resources and long training times, making them less accessible for smaller organizations or individual researchers. Additionally, the use of multiple models can add complexity and hinder scalability for real-time applications due to long processing times. Finally, the generalization of these models to other types of text or languages is uncertain, particularly when sentiment is expressed in different tones or structures, limiting their applicability beyond the review datasets tested.
--	---------------------	--	--	---	--	---

					Hyperparameter Tuning: Hyperparameters like the number of layers, learning rate, and batch size were tuned for each model to optimize their performance and ensure fair comparison. Performance Evaluation: Models were evaluated using AUROC scores across the 13 datasets. RNN models performed better with word-level input, while CNN models worked better with character-level input. Accuracy was also considered, but AUROC was the primary evaluation metric due to its robustness in handling class imbalance.	
Emotion Recognition from Social Media Text Using Deep Learning	Chen Li and Fanfan Li	2023	Experimental	The study found that the proposed CNN-BiLSTM model, combined with LDA-based feature extraction and an emotion dictionary, significantly improved accuracy in identifying emotions in social media comments. The model achieved an F1 score of 0.89 and an accuracy of 0.87,	Data Collection: The dataset for this study was gathered from Weibo, a Chinese microblogging platform, using a custom-built web crawler implemented in Python with XPath. The dataset consisted of 200,000 comments	The limitation for this study is its focus on just three general emotional categories—joy, anger, and sadness—which may not encompass the full spectrum of emotions expressed in social media content. This narrow focus limits the model's ability to

				performing better than multiple machine learning and deep learning models. By integrating LDA and an emotion dictionary, the model effectively extracted emotional words, leading to an increase of 3.66% in F1 score over machine learning models and 1.84% over deep learning methods. These results indicate that using both local and contextual feature extraction enhances sentiment classification, allowing for more precise identification of joy, anger, and sadness in microblog discussions.	related to public opinion events, categorized into three primary emotional labels: joy, anger, and sadness. The dataset was split into training (60%), testing (25%), and validation (15%) sets to ensure balanced model evaluation. Data Preprocessing: The collected data underwent extensive cleaning and preparation. Chinese word segmentation was performed using the Jieba tool to split the text into meaningful words. Afterward, stop words and special characters were filtered out, and duplicate comments were removed. Manual annotation was applied to ensure the accuracy of emotional labels, ensuring that only relevant and clean data was used for the modeling process. Feature Extraction (LDA Model & Emotion Dictionary):	capture more subtle or complex emotions such as fear or surprise. Additionally, the model is based solely on text analysis, meaning it might not perform well when analyzing multimodal content, like images or videos, which are often part of social media discussions. The reliance on LDA and emotion dictionaries for feature extraction could also present challenges, as these methods may not account for the full range of emotional expressions found in the informal and often noisy language used on social platforms. Expanding emotional categories and incorporating more varied data sources could enhance the model's capability in future research.
--	--	--	--	---	--	--

To extract meaningful emotional features, the study employed the Latent Dirichlet Allocation (LDA) model to identify emotion-related terms in the text. In addition, an emotion dictionary was used to refine the feature selection process by filtering out irrelevant words. The LDA model, combined with the emotion dictionary, ensured that the emotional features extracted from the dataset were both relevant and accurate.

Word Embedding (Word2Vec):
The study utilized Word2Vec, a word embedding technique, to convert the extracted emotional features into word vectors. Word2Vec captures semantic relationships between words and helps the model understand the contextual meanings of words within social media comments. This embedding technique provided a better

					representation of text data for the subsequent model classification tasks. Model Architecture (CNN-BiLSTM): The study implemented a hybrid deep learning model combining Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) networks to classify emotions in the comments. The CNN was responsible for extracting key local features from the comment text, while the BiLSTM captured long-range dependencies and semantic relationships in the text. The final emotion classification was performed using a Softmax classifier to predict whether the text expressed joy, anger, or sadness. Performance Evaluation: The model's performance was evaluated using F1 score and accuracy as	
--	--	--	--	--	--	--

					the primary metrics. The results were compared against several traditional machine learning models, including Logistic Regression, SVM, Random Forest, KNN, Naïve Bayes, and AdaBoost, as well as deep learning models such as LSTM, BiLSTM, GRU, BiGRU, CNN, and TextCNN. The proposed model demonstrated superior performance, outperforming these models in emotion recognition tasks.	
Aspect-Based Sentiment Analysis Using Deep Learning	Mengqi Zhang, Jiazhao Chai, Jianxiang Cao, Jialing Ji, Tong Yi	2024	Experimental	The study found that the proposed U-ASGCN model, which combines User Comment Feature Extraction (UCFE) with Aspect-Sentence Graph Convolutional Networks (ASGCN), significantly improved the performance of aspect-level sentiment analysis. By integrating user-specific sentiment tendencies, derived from historical user comments, the model achieved higher accuracy and F1 scores compared to traditional sentiment analysis models	Data Collection: The dataset used in this study was primarily sourced from the Yelp and SemEval datasets. The Yelp dataset was used for user-level sentiment analysis, while the SemEval datasets provided diverse examples for aspect-level sentiment classification. The user-aspect dataset was specifically created by extracting aspect words from comments, along	The limitation of this study is its reliance on user comment data to extract user sentiment tendencies. This approach assumes that a user's past comments accurately reflect their current sentiment, which may not always be the case, especially if a user's opinions or behavior have changed over time. Additionally, the model is focused entirely on textual data, which means it may not perform well when

like BERT and BiLSTM. The results demonstrated that embedding user features at key stages in the analysis pipeline enhanced the model's ability to capture both contextual relationships between aspect words and user-specific sentiment tendencies, ultimately leading to more accurate sentiment predictions. This suggests that incorporating user behavior along with advanced deep learning models can effectively improve the precision of aspect-based sentiment classificatio

with sentiment labels and user information.

Data Preprocessing:
We started by cleaning up the data, getting rid of unnecessary stuff like stopwords, punctuation, and irrelevant content. Then, we used a BiLSTM-CRF model to pick out key aspect words. Each of these words was assigned a sentiment label based on how it was used in the sentence. On top of that, we factored in user information—like their past comment trends—to enrich the feature vectors and make the sentiment analysis more insightful

Feature Extraction (UCFE):
The User Comment Feature Extraction (UCFE) method was used to extract user-specific features from the historical comments of users. These features were transformed into user feature vectors to capture the sentiment tendencies of each user,

dealing with other types of content like images or videos that are often present in online reviews. The datasets used in the study, such as Yelp and SemEval, provide valuable insights but may not cover the full range of user experiences or sentiment expressions across different platforms or industries. Including more diverse datasets in future research could help improve the model's generalizability. Lastly, while the model incorporates user features effectively, there may be room for further refinement in how these features are integrated into the analysis, particularly in more complex or nuanced contexts.

					which were integrated into the sentiment classification model. Model Architecture (U-ASGCN): The proposed model combines the UCFE model, which focuses on extracting user features, and the ASGCN model, which captures relationships between aspect words and their syntactic dependencies. Three embedding methods were proposed to incorporate user features into the ASGCN model, allowing the integration of user sentiment tendencies at various stages of the model pipeline. This hybrid approach enables the model to capture both contextual information and user-specific sentiment. Performance Evaluation: The performance of the U-ASGCN model was evaluated using accuracy and F1 scores on the Yelp and	
--	--	--	--	--	--	--

					SemEval datasets. The model consistently outperformed baseline models, such as BERT and BiLSTM, showcasing the effectiveness of incorporating user features into aspect-level sentiment analysis.	
Dynamic Light Weight Recommendation System for Social Networking Analysis Using a Hybrid LSTM-SVM Classifier Algorithm	N. S. Kiruthikaa Dr. G. Thailambala	2022	Experimental	This study highlights the advantages of ensemble deep learning models in sentiment classification, demonstrating superior accuracy over traditional approaches. By integrating LSTM and word embeddings, these models effectively capture contextual relationships between words, which is particularly useful for handling evolving terminology, such as language shifts during the COVID-19 pandemic. When tested on datasets from Twitter, Amazon, and Yelp, the model achieved 91.3% accuracy, outperforming CNN (81.5%) and standard LSTM (85.1%). Additionally, it maintained high precision and recall, reducing misclassification errors and enhancing overall reliability. These findings suggest that combining	Data Collection: This study used datasets from Twitter, Amazon, and Yelp for sentiment analysis. The Twitter data focused on COVID-19 tweets, while Amazon and Yelp reviews provided general sentiment insights. COVID-19 tweets were manually gathered using relevant hashtags. Data Preprocessing: The text was cleaned by removing stopwords, punctuation, and special characters. Lemmatization and spelling correction helped normalize words. Word embeddings like Word2Vec, GloVe, and BERT converted text into numerical	The study presents a hybrid LSTM-SVM model for fake news detection on Twitter, specifically targeting misinformation related to COVID-19. While the approach demonstrates high accuracy (90%) and specificity (97%), several limitations exist. The dataset, consisting of only 5000 tweets, is relatively small for training deep learning models, potentially affecting generalizability. Additionally, the model is tested on a static dataset rather than real-time data, raising concerns about its performance in dynamic environments. The binary classification approach oversimplifies the problem, as it does

multiple deep learning techniques creates a more robust and adaptable sentiment analysis system, particularly valuable for dynamic datasets where language continuously evolves.

representations while preserving context.

Feature Extraction:
Count Vectorizer and Word Embeddings structured the text for analysis, capturing word relationships and sentiment patterns to improve classification accuracy.

Model Architecture:
The model combined LSTM with an ensemble classifier to enhance sentiment detection. LSTM captured word context, while CNN and Bi-LSTM refined classification. A softmax activation function assigned sentiment labels: positive, negative, or neutral.

Performance Evaluation:
The model achieved 91.3% accuracy on COVID-19 tweets, outperforming CNN (81.5%) and LSTM alone (85.1%). It maintained high precision, recall, and

not account for partially misleading or satirical content. The reliance on a count vectorizer for feature extraction, instead of advanced word embeddings like Word2Vec or BERT, may limit the model's ability to capture deeper semantic relationships in text. Computational efficiency is another concern, as combining LSTM and SVM increases processing time, making real-time deployment more challenging. Moreover, the study lacks consideration of non-textual features such as user credibility, engagement metrics, or network-based features, which could enhance detection accuracy. Since the dataset is focused on COVID-19 misinformation, the model's ability to generalize to other topics remains uncertain. Lastly, manually curating the dataset may introduce selection bias, further impacting the model's effectiveness when

					F1-scores on Amazon and Yelp reviews, proving its reliability in sentiment analysis.	applied to diverse sources of fake news.
--	--	--	--	--	--	--

References:

Alsayat, Ahmed. "Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model." *Arabian Journal for Science and Engineering*, vol. 47, no. 6, 2022, pp. 2499-2511. <https://doi.org/10.1007/s13369-021-06227-w>.

Seo, Seungwan, et al. "Comparative Study of Deep Learning-Based Sentiment Classification." *IEEE Access*, vol. 8, 2020, pp. 6861-6875. <https://doi.org/10.1109/ACCESS.2019.2963426>.

Li, C., & Li, F. (2023). Emotion recognition of social media users based on deep learning. *PeerJ Computer Science*, 9, e1414. DOI: [10.7717/peerj-cs.1414](https://doi.org/10.7717/peerj-cs.1414)

Zhang, M., Chai, J., Cao, J., Ji, J., & Yi, T. (2024). Aspect-Level Sentiment Analysis Based on Deep Learning. *CMC*, 78(3), 3744-3762. <https://doi.org/10.32604/cmc.2024.048486>.

Dang, Cach N., María N. Moreno-García, and Fernando De la Prieta. "Sentiment Analysis Based on Deep Learning: A Comparative Study." *Electronics*, vol. 9, no. 3, 2020, p. 483. MDPI, <https://doi.org/10.3390/electronics9030483>.