International Journal of Advances in Applied Sciences (IJAAS)
Vol. 14, No. 2, June 2025, pp. 562~579
ISSN: 2252-8814, DOI: 10.11591/ijaas.v14.i2.pp562-579

Sentiment analysis of vaccine data using enhanced deep learning algorithms

Monika Verma, Sandeep Monga
Department of Computer Science and Engineering, Oriental University, Indore, India

Article Info

Article history:

Received Oct 15, 2024
Revised Apr 25, 2025
Accepted May 10, 2025

Keywords:

Bi-LSTM model
Deep learning algorithms
Dropout rates
Hierarchical SoftMax
Negative sampling
Sentiment analysis
Vaccine dataset

ABSTRACT

This paper investigates and experiments with an approach to improve sentiment analysis on vaccine datasets with deep learning. It evaluates random forest (RF), naïve Bayes (NB), and recurrent neural network (RNN) models across a variety of configurations, i.e., vector dimensions, pooling techniques, as well as evaluation methods, hierarchical SoftMax vs negative sampling. The results show that the model we proposed prevailed with an accuracy of 99.05% on a learning rate equal to 0.001, outperforming all other models based on metrics including precision, recall, and F1-score for benign/malignant cases. The results suggest that higher vector dimensions, average pooling, lowering the dropout rate, and employing hierarchical SoftMax for output significantly improve model performance. Hierarchical SoftMax performs better than negative sampling, whereas a lower dropout rate decreases overfitting and leads to improved generalization. Our results demonstrate the necessity to apply more sophisticated deep-learning tools around capturing nuances of public vaccine-related sentiment, which may be crucial for informing communication strategies and supporting decision-making in a real-world health emergency. The findings indicate that the performance of sentiment analysis with regard to COVID-19 vaccine deployment policy design and public monitoring could be enhanced by advanced deep learning algorithms.

Corresponding Author:

Monika Verma
Department of Computer Science and Engineering, Oriental University
Gate No.1, Sanwer Rd, opposite Revati Range, Jakhya, Indore, Madhya Pradesh 453555, India
Email: monikaverma03@rediff.com

1. INTRODUCTION

The rapid development and deployment of vaccines, particularly in response to global pandemics like COVID-19, have sparked extensive public discourse and varied sentiments across social media, news platforms, and other communication channels. Understanding these sentiments is crucial for public health officials, policymakers, and researchers to gauge public opinion, address concerns, and craft effective communication strategies. Sentiment analysis, a field within natural language processing (NLP), involves determining the emotional tone behind texts and is increasingly being applied to study vaccine-related sentiments. However, analyzing sentiments in vaccine datasets poses unique challenges due to the complexity and variability of the text data. This study focuses on enhancing sentiment analysis of vaccine datasets using advanced deep learning algorithms, leveraging state-of-the-art techniques such as hierarchical SoftMax, negative sampling, and novel model architectures to improve accuracy and reliability.

The effectiveness of deep learning models in sentiment analysis is significantly influenced by various factors, including vector dimensions, pooling techniques, dropout rates, and evaluation methods. For example, increasing vector dimensions from 100 to 300 has been shown to enhance model performance, as it

allows the representation of more semantic nuances in the text. Similarly, using advanced pooling techniques, such as average pooling over max pooling, can result in better model accuracy by effectively capturing the most representative features from the data. This study also highlights the importance of fine-tuning model parameters, such as dropout values and learning rates, to prevent overfitting and improve generalization on unseen data.

This research demonstrates the potential of advanced deep-learning algorithms for enhancing sentiment analysis of vaccine datasets. By leveraging state-of-the-art techniques such as hierarchical SoftMax, average pooling, and optimized parameter settings, the proposed model achieves unprecedented accuracy and robustness. These findings not only contribute to the field of sentiment analysis but also provide valuable insights for public health communication and policy-making, particularly in understanding and responding to public sentiments about vaccines. Future work can explore the integration of other deep learning architectures, such as transformers, and extend the analysis to multilingual vaccine datasets to further enhance the generalizability and impact of these models. The contributions of this study are summarised as follows: i) We propose an innovative deep-learning model for sentiment analysis, specifically designed for vaccine-related datasets. The model integrates bidirectional long short-term memory (Bi-LSTM) with attention mechanisms, optimizing sentiment classification by capturing contextual information from the text; ii) Hierarchical SoftMax is employed as a superior evaluation method over negative sampling, significantly improving model performance and accuracy; iii) We systematically investigate the impact of various configurations, including vector dimensions, pooling techniques (average pooling versus max pooling), and dropout rates, revealing their influence on model accuracy and generalization; and iv) The study demonstrates that higher-dimensional embeddings and lower dropout rates enhance the model's ability to learn complex sentiment patterns and prevent overfitting, offering crucial insights for improving sentiment analysis in public health communication.

The rest of the paper is divided into five sections: section 2 presents a literature review of existing sentiment analysis techniques for vaccine datasets; section 3 details the proposed methodology, including the advanced deep learning algorithms employed; section 4 discusses the results, providing an in-depth analysis of the findings and their implications; and section 5 concludes the study, summarizing the key outcomes.

2. LITERATURE REVIEW

Social media and online forums allow individuals to debate public health topics like COVID-19 and spread correct and inaccurate information. This study introduces Bi-LSTM-based NLP. Classifying sentiments and identifying issues related to COVID-19 public comments is the goal. To improve traditional LSTMs, Bi-LSTM uses outputs from previous and subsequent data contexts. Our performance metrics exceeded traditional LSTM models and prior research after analyzing Twitter and Reddit datasets. This notion helps governments reduce negative communication and understand public opinion during health crises. Our work also underlines the need to employ NLP to analyze public mood, which may guide health policy [1].

After the 2019 COVID-19 pandemic, which is assumed to have started in Wuhan, China, people globally practiced hand hygiene, face masks, and physical distancing. COVID-19 vaccinations were introduced in early 2021 in several countries, including the US, bringing relief but also polarizing debate. Vaccine reluctance became a major problem after these conversations. Since conventional data is limited, live-streamed tweets from API queries provide a viable way to study public opinion on vaccine issues. Azure machine learning (ML), VADER, and TextBlob were employed in this work together with five machine learning algorithms and three text vectorization methods for sentiment analysis. Our comprehensive model evaluation showed public sentiment on COVID-19 immunization is improving. The best public sentiment prediction metrics were achieved using TextBlob sentiment score, term frequency-inverse document frequency (TF-IDF) vectorization, and LinearSVC classification [2].

Search engines and social networks have been used in internet-based syndromic monitoring to predict epidemic tendencies for 20 years. Recently, the focus has shifted to understanding public emotional responses to health crises, notably pandemics. This study examined Greek Twitter's attitude towards COVID-19 cases in real-time. Two emotion lexicons one translated from English to Greek and one original Greek were used to analyze 153,528 tweets from 18,730 persons. The study measured positive and negative attitudes and six emotions. We also examined attitudes, tweet volumes, and COVID-19 cases. The most common emotions were surprise (25.32%) and disgust (19.88%). However, mood did not significantly correlate with COVID-19 dissemination, indicating that interest may decline with time [3].

Indonesia's COVID-19 answer split Twitter users. The examination of these tweets may inform policymaking and government assessment. Sentiment analysis uses tweets to assess public opinion. This research examines public opinion on Indonesia's epidemic management from general and economic aspects. The Twitterscraper library collects tweets. Sentistrength_id and expert assessment classified these tweets as favorable, negative, or neutral. Data pre-processing removed extraneous tweets. Confusion matrices and

K-fold cross-validation confirmed the accuracy of machine learning models that assessed fresh data sentiment. Support vector machine (SVM) research yielded impressive accuracy, precision, recall, and f-measure scores of 82.00, 82.24, 82.01, and 81.84%. Civilians tended to like the epidemic's economic policies, but many were dissatisfied with the government's performance. The SVM approach, notably with the Normalized Poly Kernel, predicts Twitter sentiment rapidly and accurately [4].

The COVID-19 pandemic caused health and economic problems. Big data helps logistics businesses find profitable solutions to these difficult situations. This study analyzed logistics consultancy websites using text mining. The main objectives were to detect frequent phrases, undertake sentiment analysis using the NRC lexicon, stress common word combinations, and provide cost-effective shipping and inventory management solutions throughout the pandemic. Important phrases were "supply," "chain," and "COVID-19." Trust-related emotions were present, showing a desire for reliable remedies throughout the crisis [5].

Information on social media, particularly on the COVID-19 pandemic, is frequently false. This research analyses Malaysian COVID-19 news sentiment on Twitter. We collected largely Malay, English, and Chinese tweets since Malaysian tweets are multilingual. After analyzing the byte-pair encoding-text-to-image-convolutional neural network (BPE-Text-to-Image-CNN) and BPE-multilingual bidirectional encoder representations from transformers (BPE-M-BERT) models, we chose the pre-trained M-BERT network for our multilingual dataset [6].

Twitter has made it easier to follow global events and public viewpoints, especially amid health crises. The COVID-19 and MPox pandemics boosted Twitter communication. This study analyses tweets about two conditions simultaneously to fill a research gap. Mainly negative opinions and numerous allusions to President Biden and Ukraine give a complete understanding of popular discourse in this age [7].

COVID-19 prevention measures were implemented, but the community's ignorance and lack of self-control rendered them ineffective. This oversight increased viral exposure. In addition to health, the pandemic had major impacts on social, political, religious, economic, and population resilience. Social media also sheds light on these effects, especially socioeconomic ones. This study uses Twitter API data to assess Indonesians' pandemic views. We created a sentiment analysis tool using TF-IDF, Lexical, and NB for feature extraction and classification. Indonesian Twitter tweets on "COVID-19" were classed by emotion: fear, rage, love, grief, and happiness. Our proposed transcription factor binding site (TFBS) algorithm outperformed others with an accuracy of 0.85. Accuracy, recall, and F-score demonstrated their superiority [8]. Sentiment analysis of Greek National Public Health Organization (EODY) Facebook posts during the pandemic was performed using Microsoft Azure Machine Learning Studio. An examination of 300 reviews identified opinions as positive, negative, or neutral. The research examined reactions to EODY's Facebook daily COVID-19 surveillance reports. The sentiment classifications were automated using machine learning. Government, immunizations, and COVID-19 were often mentioned, indicating disapproval of these reports. Neural network (NN) and NB point machine classification methods have high accuracy and F1-scores. In pandemics, machine learning can measure public opinion and aid public health decision-making [9].

Palliative care became increasingly important as the disease spread. This study investigated over 26,000 English tweets from 2020 to 2022 to determine palliative care views during the pandemic. Four themes emerged from web scraping. Notably, although many tweets highlighted the pandemic's negative effects on palliative care, many also highlighted its positive effects. Machine learning categorized several of these tweets, focusing on COVID-19's negative effects [10]. Twitter's wide public opinion may bring fresh and essential information, especially on vaccine hesitancy. The current research preprocessed and categorized subject-related tweets by various attitudes and emotions using NRC Lexicon. Statistical testing confirmed emotional correlations. After training many neural networks for sentiment multi-classification, the BERT model obtained 96.71% accuracy [11].

This study analyses Twitter data to examine epidemic-related mobility mode preferences. January 2020 to January 2022 tweets concerning New York City transit modalities were gathered. We created NLP-based travel mode classifiers to categorize tweets into multiple transportation modes using this data. During the outbreak, public sentiment changed towards transportation choices. Buses, bicycles, and private cars were well-regarded since commuters chose them over subways. Twitter posts on the tube and bus mask violations raised concerns. A regression study employing user demographics indicated the factors that influence public transport attitudes, particularly the service sector's vulnerability to metropolitan transportation authority (MTA) subway performance [12].

COVID-19's global pandemic caused anxiety and drove many South Africans to religious rites. Social media data is analyzed to determine South Africans' views on religion and well-being throughout the pandemic. We analyzed tweets about COVID-19, religion, life's purpose, and personal experiences to determine sentiment. This research shows that religious beliefs and COVID-19 attitudes affect life events.

Also, we propose a novel sentiment analysis threshold of depression measure that provides useful insights into collective emotional circumstances during crises [13].

Twitter became a global conversation venue during the COVID-19 pandemic. There is little study on the potential usefulness of sentiment analysis of epidemic tweets, despite various studies on this platform. This research shows how sentiment analysis paired with behavioral and social science might help pandemic management. After reviewing machine learning approaches for sentiment analysis of pandemic tweets, we conclude that ensemble models, such as BERT and RoBERTa, are best for Twitter data [14].

Vaccinations are key to fighting the pandemic. Social media analytics and health surveillance data are used to study complicated public perceptions concerning COVID-19 immunizations. Deep learning was used to assess millions of tweets from 2020 to 2022 for emotions. This research examines the pregnant population's emotional trajectory and trends. We identify ways to increase vaccine outreach by correlating sentiments with global vaccination patterns [15].

Social media platforms saw a rise in instructive and false content during the COVID-19 pandemic. This study studies Twitter comments on COVID-19 vaccination to help create a vaccine acceptance policy. After testing numerous methods, the extra tree classifier (ETC) using the bag of words (BoW) is the most efficient sentiment analysis framework for COVID-19 tweets. Our analysis constantly shows a growing endorsement of vaccination [16].

Twitter showed the synchronized public attitude changes generated by COVID-19. This research dives into Mexican beliefs during a particularly difficult pandemic time. To measure attitudes, we trained multiple models, two of which were trained in Spanish, using a hybrid semi-supervised technique. Comparing the Spanish-centric model to SVM and Decision Trees showed its higher accuracy. Mexican Twitter users' COVID-19 moods were assessed using this tailored approach [17].

The research examined how daily TikTok case films by Public Health Organizations (PHAs) affected COVID-19 acceptance and understanding. The goal was to understand public opinion and anxiety during the 2022 Shanghai lockdown. The crisis and emergency risk communication model divides the lockdown into five parts [18]. This research divided the Shanghai shutdown into stages. User reaction to Healthy China's daily TikTok case videos was extensively analyzed throughout these periods. The pre-trained ERNIE model classified user comment emotions. Sentiment classification also informed semantic network investigations. Given the high cost of controlling the epidemic, the public was unwilling to take prophylactic measures in the beginning. Shanghai's unilateral definition of "asymptomatic patients" affects control efforts in other municipalities and daily life. During maintenance, individuals focused on disease-related areas of their lives. Interest in daily case update videos dropped after resolution. The apparent divergence of local government strategy from central government orders caused widespread displeasure. The worldwide health landscape is threatened by COVID-19. Vaccine development and distribution are crucial to preventing this. As vaccine dissemination increases, viral immunity should rise. In the digital age, Twitter is essential for public opinion research, especially for vaccination efforts. This study analyzed COVID vaccination tweets using advanced AI and geo-spatial methods. TextBlob analyzed these tweets' sentiment polarity. The data was graphed using word clouds and emotion was identified using BERT. Geocoding placed and visually displayed emotion data on a global map. Advanced approaches include hotspot analysis and kernel density estimations that identify regions with pleasant, negative, or neutral sentiments. The model's accuracy, recall, and F-score for positive and negative sentiment categorizations were compared to well-established approaches. The main purpose is to study public opinion on vaccination programs during global health emergencies using sentiment and geographical analytics [19].

The worldwide spread of SARS-CoV-2 has accelerated since March 2020. Millions have been infected since the outbreak, prompting global reactions on Twitter. The main goal of this research is to analyze Hindi-speaking Twitter users' views on the pandemic. Data processing began with NLP on Hindi COVID-19 tweets. Afterward, a unique Grey Wolf optimization approach improved feature selection. The major analytical tool was a hybrid model that combined the benefits of CNNs and LSTMs. A comparison of the proposed model with current machine learning paradigms showed superior accuracy, precision, recall, and F-score [20].

The COVID-19 epidemic has had a major economic effect on worldwide financial markets. Investors are interested in accurate forecasts since shares are volatile, especially during epidemics. This research introduces sentiment analysis of COVID-19 news to predict stock market changes. Daily stock movements were forecasted using COVID-19 stock news headlines. Machine learning classifiers also estimated the epidemic's impact on high-value stocks including Tesla, Inc. (TSLA), Amazon.com, Inc (AMZN), and Alphabet Inc. (GOOG). To boost forecast accuracy, features were refined and spam tweets were removed. The system's textual analysis of social media and data mining of historical stock data make it unique. The approach predicted stocks accurately [21].

Social media has been utilized globally to share ideas, feelings, and viewpoints on the COVID-19 pandemic from its start. Twitter and other digital platforms save publicly communicated material, allowing

individuals to debate the pandemic at any time or location. The fast rise in cases worldwide has caused people to worry, dread, and feel uncomfortable. This study proposes a novel sentiment analysis method for recognizing emotions in Moroccan COVID-19 tweets from March to October 2020. Our system classifies tweets as positive, bad, or neutral using recommendations. Our method outperforms well-established machine learning methods with 86% accuracy, according to experiments. Temporal changes in sentiments show that the shifting COVID-19 scenario in the country affected public feelings [22].

Twitter data from March 2023 was used to analyze Indian public opinion on the COVID-19 vaccine. Before sentiment analysis using NLP, tweets were preprocessed and sanitized using relevant hashtags and keywords. Numerous tweets supporting immunization show a positive outlook. However, questions were raised concerning immunization reluctance, side effects, and government and pharmaceutical mistrust. A more detailed analysis of views by gender, age, and regional indicators revealed different feelings across population groups. Our analysis illuminates the vaccination situation in India and emphasizes the necessity for personalized communication to minimize vaccine uncertainty and boost acceptability among chosen populations [23].

US COVID-19 statistics. African Americans had a disproportionate share of viral infections and deaths, according to change data capture (CDC) data through June 2020. This mismatch highlights the need to understand African American COVID-19 experiences and views. Our study examines African American pandemic narratives using aspect-based sentiment analysis of 2020 Twitter data. Our machine learning technology filters tweets that are unrelated to COVID-19 or not from African Americans. Approximately 4 million tweets were analyzed. The tweets were mostly gloomy, with volume spikes coinciding with US pandemics. This research shows how pandemic language changed during the year and highlights crucial challenges including food shortages and vaccine rejection. To understand how the pandemic affected African American Twitter users' conversations, we wish to stress word linkages and attitudes [24]. The media must aggressively promote health concerns to raise public awareness and reduce health hazards to improve society. Deep neural networks are becoming increasingly popular in textual sentiment analysis, allowing real-time health monitoring and insights. Cov-Att-BiLSTM is an innovative artificial intelligence model for sentiment analysis of COVID-19 news headlines. The model uses deep neural networks. Our work uses attention processes, embedding methods, and semantic data labeling to improve prediction accuracy. Our model performed well versus other machine learning classifiers, attaining 0.931 test accuracy. We also applied it to 73,138 pandemic tweets from six foreign sources to accurately depict global COVID-19 news and vaccination discussions [25].

Progress is driven by knowledge. Data collection techniques have evolved largely due to technological advancement. Many technological mediums are fast, reliable, and efficient, yet others are lacking for various reasons. This research will analyze Nigerian COVID-19 tweets and public opinion using TextBlob and VADER models. It measures emotional responses and illuminates Nigeria's societal, ecological, and economic effects. This work might be useful for data science, machine learning, and deep learning researchers. After preprocessing 1,048,575 'COVID-19' tweets, TextBlob and VADER were used to assess sentiment. Our research found a variety of perspectives, demonstrating the efficacy of social media analysis in assisting major companies battle COVID-19's effects and misinformation [26].

Twitter allows people globally to communicate their ideas, especially during major events like the current pandemic. This research uses tweets to analyze Indian perceptions concerning COVID-19 and immunization. We classified tweet emotions using deep learning and lexicon-centered methods. The lexicon approach utilized VADER and NRCLex, whereas the deep learning method used Bi-LSTM and gated recurrent unit (GRU), yielding amazing accuracy. Our models may help healthcare professionals and decision-makers in future pandemics [27].

Social distancing measures did not deter 20% of commuters from using public transport during the COVID-19 epidemic. Traditional urban transportation data collection techniques miss passengers' complex psychological experiences. Thus, understanding the transportation-dependent impoverished populations' situation is tough. We used machine learning to segment Twitter data to properly depict the travel patterns of 120,000 metro Vancouver transit riders before and throughout the pandemic. Our research found a considerable rise in unfavorable views, notably across demographic categories, during the early COVID-19 epidemic. Our goal is to identify transit users' inequities and risks during the crisis to improve public health and transport planning in future disruptions [28].

This study uses NLP and opinion analysis (SA) to assess Italian public opinion about COVID-19 immunization. This research solely includes Italian vaccine tweets from January 2021 to February 2022. From 1,602,940 tweets including "vaccine," 353,217 were examined. We divide users into four categories: common users, media, medicine, and politics, which is novel. This categorization evaluates user profiles using NLP and domain-specific lexicons. Our sentiment analysis uses Italian's polarized and heated words to reveal each user group's tone. The study found generally negative attitudes throughout the review period,

especially among common users. In particular, post-immunization fatalities altered attitude patterns after 14 months [29].

Sentiment analysis in NLP may extract useful information from online COVID-19 data, helping China fight the pandemic. Deep learning-based sentiment analysis algorithms have improved; however, dataset limits typically hinder them. We present a federated learning system that blends BERT with a multi-scale convolutional neural network (Fed_BERT_MSCNN) to address this problem. The suggested model uses transformer-derived bidirectional encoder representations and multiscale convolution layers. The federated architecture trains datasets independently using a central server and local deep learning systems. Edge networks simplify model integration and parameter exchange. This new network addresses data scarcity, increases communication efficiency, and improves data privacy during training. The Fed_BERT_MSCNN model outperformed its peers on six social sites [30].

The extraordinary COVID-19 pandemic affected financial markets, especially in its early stages. The current research examines how COVID-19 news flows affect market expectations. We examined 203,886 COVID-19 related internet articles from January to June 2020 on MarketWatch.com, NYTimes.com, and Reuters.com. We used a financially-tuned BERT model, which understands contextual word semantics, to analyze these articles' sentiments using machine learning. Our research shows a substantial positive correlation between sentiment indicators and S&P 500 market performance. It's interesting that NYTimes.com's attitude and news kinds affected market performance differently [31].

3. PROPOSED METHODOLOGY

3.1. Proposed flowchart

Figure 1 presents a comprehensive workflow for analyzing textual data using an advanced Bi-LSTM model. The process begins with a raw text input that is first subjected to processing and cleaning to remove noise, and irrelevant information, and standardize the text for analysis. Next, the cleaned text undergoes a computational phase where sentiment scores are calculated, and the objectivity of the text is assessed to determine the subjective or objective nature of the content. These sentiment and objectivity scores are then prioritized and ranked in order of their significance to refine the analysis. Following this, the scores are utilized to initialize a Bi-LSTM model, which incorporates a contextual prediction method, enhancing the model's ability to understand context and dependencies in the text data. This initialized Bi-LSTM model is then employed to accurately identify the review category of the text, categorizing it based on predefined criteria or themes. The entire process culminates in producing a final result, which effectively categorizes the text and provides meaningful insights derived from the sentiment and context of the input data.

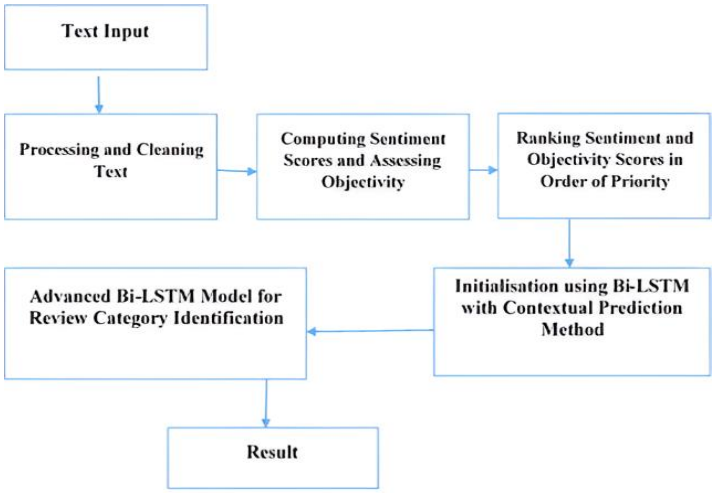


Figure 1. Proposed methodology workflow

3.2. Proposed algorithm

The preprocessing algorithm takes textual data as input and produces preprocessed text as output. As shown in Algorithm 1. It begins by iterating through each line of text in the file, where the text is first tokenized to break it down into individual words or tokens. Next, common stopping words are removed to eliminate unnecessary or irrelevant words that do not contribute significantly to the analysis. The text then undergoes stemming, which reduces words to their root form, followed by lemmatization, which further

refines the text by converting words to their base or dictionary form. The preprocessed text is then returned for further analysis.

The proposed Bi-LSTM model algorithm takes preprocessed text as input and outputs a class label by systematically processing the text through multiple steps to enhance its understanding and classification as shown in Algorithm 2. It begins by iterating over each token in the text, which represents a list of feature maps. First, the text is embedded using a Bi-LSTM embedding layer to capture contextual information. The embedded text is then processed through a Bi-LSTM layer to understand the sequence and capture long-range dependencies between words. An attention mechanism is applied to focus on the most relevant parts of the text, improving the model's ability to identify crucial features. Sentiment analysis is conducted to determine the polarity and objectivity (Pl, Ol) of the text, highlighting significant sentiment features by ranking these scores. The context of the text is extracted using context extraction techniques to understand the underlying scenario and topic. Finally, a Bi-LSTM-based classifier uses the sentiment and context features (Pl, Ol, and Ct) to make a final classification, and the corresponding class label is returned as the output. This approach leverages the combined strength of Bi-LSTM and attention mechanisms for effective text classification based on both sentiment and contextual information.

The proposed final Bi-LSTM in-depth model is designed to classify sentiment in text data about vaccination by leveraging multiple advanced natural language processing techniques as shown in Algorithm 3. It begins with a preprocessing step where raw text data is cleaned by removing noise such as URLs and special characters and normalized by converting text to lowercase and applying stemming. Next, the preprocessed text is transformed into a numerical format using word embeddings like GloVe, converting each token into a dense vector representation. The sequence of word vectors is then passed through a Bi-LSTM layer, which captures both forward and backward contextual dependencies to understand the sequence and meaning within the text. An attention mechanism is applied to the hidden states generated by the Bi-LSTM, highlighting the most important parts of the text for analysis. Sentiment analysis is performed on these text features to predict sentiment polarity and objectivity scores, determining whether the text expresses positive or negative sentiment and its level of objectivity. Simultaneously, context extraction techniques are applied to understand the topic and thematic elements of the text. Finally, the sentiment scores and context features are fed into a Bi-LSTM-based classifier, which outputs a sentiment label, categorizing the text as positive or negative regarding vaccination. This comprehensive approach combines preprocessing, embedding, deep learning, attention mechanisms, sentiment analysis, and context extraction to achieve precise sentiment classification.

Algorithm 1. Preprocessing
Input: Textual data
Output: Preprocessed text
Step 1: Begin ()
{
Step 2: For each text line in the file:
Step 3: Text ← Extract tokens
Step 4: Text ←Remove stopping words
Step 5: Text ← Stemming
Step 6: Text ← Lamination
Step 7: return text
}

Algorithm 2. Proposed Bi-LSTM model
Input: Preprocessed text
Output: Class label
{
Step 1: For token in text: //Text contains a list of feature maps
Step 2: Textf ← Embedding //Apply a Bi-LSTM embedding layer to capture contextual information from the text
Step 3: Textf ← Bi-LSTM //Process the text using a Bi-LSTM layer to understand the sequence and capture dependencies
Step 4: Textf ← Attention //Apply an attention mechanism to focus on relevant parts of the text
Step 5: Pl, Ol ← Apply sentiment analysis over Textf to get polarity and objectivity //Use an appropriate sentiment analysis technique to obtain the sentiment features
Step 6: Pl, Ol ← Rank the Pl, Ol //Rank the polarity and objectivity scores to highlight significant sentiment features

Step 7: Ct ← Extract context from text //Use context extraction techniques to understand the scenario and topic of the text
Step 8: Label ← Bi-LSTM classifier makes classification over text by receiving Pl, Ol, and Ct as input //Use a Bi-LSTM-based classifier that considers sentiment and context features for text classification
return label
}

Algorithm 3. Proposed final Bi-LSTM in-depth model
Step 1. Preprocessing:
Input: Raw text data about vaccination.
- Goal: Clean and prepare text for analysis.
- Process:
  - Remove noise (e.g., URLs, special characters).
  - Normalize text (e.g., lowercase, stemming).

Step 2. Embedding:
Input: Preprocessed text.
- Goal: Convert text into numerical format.
- Process:
  - Use word embeddings like GloVe to transform each token into a dense vector.
- Equation: $\vec{w} = Embedding(token), where\ \vec{w}\ is\ the\ word\ vector.$

Step 3. Bi-LSTM layer:
Input: Sequence of word vectors.
- Goal: Capture context and dependencies in the text.
- Process:
  - Pass the embeddings through a Bi-LSTM layer.
- Equation:
  - Forward pass: $\overrightarrow{h_t} = LSTM(\overrightarrow{w_t}, \overrightarrow{h_{t-1}})$
  - Backward pass: $\overleftarrow{h_t} = LSTM(\overrightarrow{w_t}, \overleftarrow{h_{t+1}})$
  - Combined state: $h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}]$

Step 4. Attention mechanism:
Input: Hidden states from Bi-LSTM.
- Goal: Highlight important parts of the text.
- Process:
  - Apply an attention mechanism to the sequence of hidden states.
- Equation: $\propto_t = \frac{\exp(score(h_t))}{\sum_t \exp(score\ (h_t))}$, where the score is a function measuring the importance of each hidden state.

Step 5. Sentiment analysis (polarity and objectivity):
Input: Text features from the attention mechanism.
- Goal: Determine sentiment polarity and objectivity.
- Process:
  - Use sentiment analysis tools or models to predict sentiment scores.
- Equation: Pl=SentimentModel (h_t) , O1 = ObjectivitModel (h_t)

Step 6. Context extraction:
Input: Preprocessed text.
- Goal: Understand the context and topic of the text.
- Process:
  - Apply NLP techniques to extract thematic elements.
- Equation: C_t=ContextExtractor(*text*)

Step 7. Classification:
Input: Sentiment scores (Pl, Ol) and context features (Ct).
- Goal: Classify the sentiment of the text regarding vaccination.
- Process:
  - Input the features into a Bi-LSTM classifier.
- Equation: Label=Bi-LSTMClassifier (Pl, Ol, C_t)

Step 8. Output:
Output: Sentiment label (Positive, Negative).

4. RESULTS AND DISCUSSION

4.1. Dataset

The dataset has recent tweets about the COVID-19 vaccines used in the entire world on a large scale, as follows: Pfizer/BioNTech, Sinopharm, Sinovac, Moderna, Oxford/AstraZeneca, Covaxin, and Sputnik V. Initial data was merged from tweets about Pfizer/BioNTech vaccine. I added then tweets from Sinopharm, Sinovac (both Chinese-produced vaccines), Moderna, Oxford/Astra-Zeneca, Covaxin, and Sputnik V vaccines. The collection was first days twice a day until I identified approximately the new tweets quota, and then collection for all vaccines stabilized at once a day, during morning hours (GMT).

Figure 2 displays the distribution of different sentiment labels within a training dataset, showcasing the number of samples for each sentiment category. The labels include "strongly negative," "negative," "neutral," "positive," and "strongly positive." The "neutral" category has the highest representation, with a significant number of samples, suggesting a large portion of the dataset consists of neutral sentiments. The "positive" label follows as the second most frequent category, indicating a considerable number of samples with positive sentiment. The "negative" label has a moderate number of samples, while "strongly positive" has fewer samples, and "strongly negative" has the least number of samples, indicating a relatively small representation of strong negative sentiment in the dataset. This distribution suggests that the dataset is skewed towards neutral and positive sentiments, which could influence the model's performance during training, potentially favoring these more frequent categories.

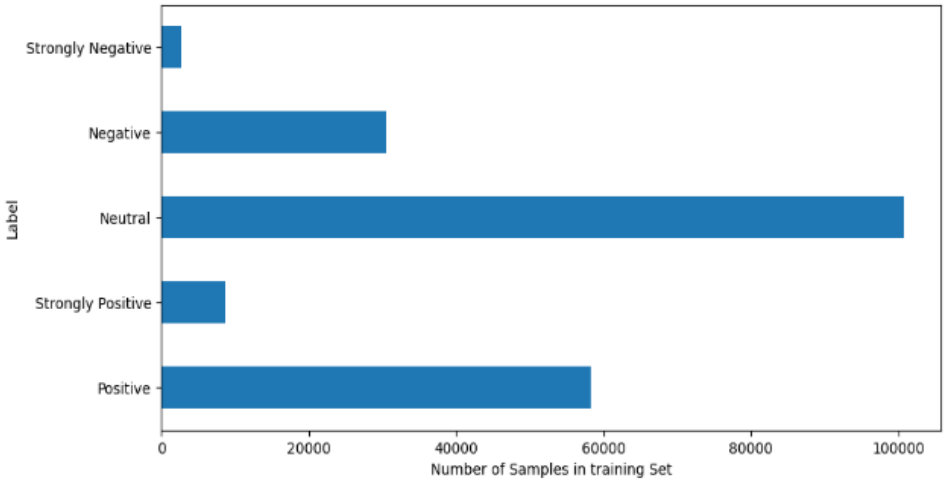


Figure 2. Number of samples in the training set

Figure 3 correlation matrix displayed in the heatmap visualizes the relationships between different features in a dataset, including id, user_followers, user_friends, user_favourites, user_verified, retweets, and favorites. The color gradient ranges from dark purple indicating a low or no correlation, closer to 0 to bright yellow indicating a high correlation, closer to 1. Notably, the matrix shows a strong positive correlation between retweets and favorites, highlighted by a bright yellow color, suggesting that tweets with more retweets tend to have more favorites. There is also a moderate correlation between user_followers and user_favourites, as indicated by a lighter shade of blue, implying that users with more followers are likely to have more favorites. However, most other correlations appear weak or non-existent, as shown by the darker purple shades. For example, the id feature has almost no correlation with any other variables. Additionally, user_verified does not show a strong correlation with other attributes, indicating that verification status does not significantly impact the number of followers, friends, favorites, or retweets. This matrix provides a comprehensive overview of how these variables are interrelated within the dataset.

Figure 4 confusion matrix illustrates the performance of a multi-class classification model by comparing the true labels against the predicted labels across five sentiment categories, represented by indices 0 to 4. The diagonal elements from top left to bottom right indicate the number of correct predictions for each category: 11,454 for class 0, 39,929 for class 1, 22,924 for class 2, 1,075 for class 3, and 3,046 for class 4, reflecting strong model accuracy in these categories. Off-diagonal elements represent misclassifications, such as 450 samples of class 0 incorrectly predicted as class 2, and 462 samples of class 4 misclassified as class 2. The matrix shows relatively high accuracy for classes 1 and 2, with minimal misclassifications across other

classes. The most frequent misclassifications occur in adjacent classes, indicating that the model sometimes confuses similar sentiment categories, such as mistaking "neutral" (class 2) for "positive" (class 1) or vice versa. Overall, the confusion matrix reveals that the model performs well, with the majority of predictions aligning with the true labels, although there are areas where further refinement could reduce misclassification rates.

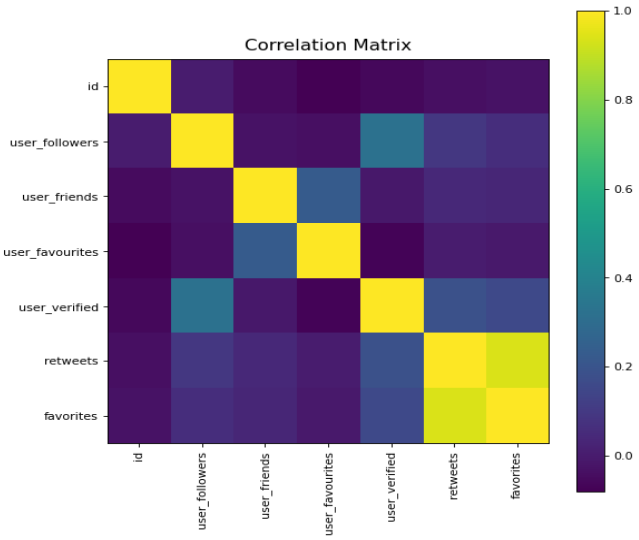


Figure 3. Correlation matrix

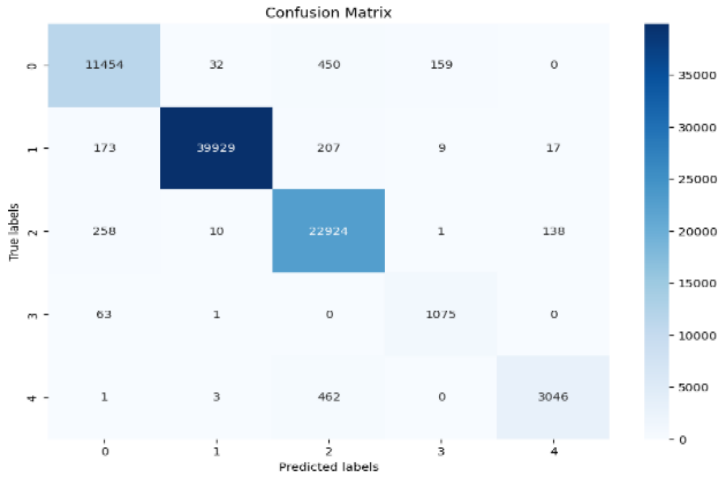


Figure 4. LSTM confusion matrix

Figure 5 confusion matrix provides a detailed evaluation of the performance of a multi-class classification model by comparing actual versus predicted labels across five sentiment categories, denoted by indices 0 to 4. The diagonal entries represent the correctly classified instances for each category: 8,685 for class 0, 29,936 for class 1, 17,227 for class 2, 776 for class 3, and 2,363 for class 4, indicating that the model performs well overall with a high number of correct predictions in each category. Off-diagonal entries highlight the misclassifications, such as 233 samples of class 0 incorrectly predicted as class 2 and 278 samples of class 4 misclassified as class 2. The most frequent misclassifications occur between similar or adjacent sentiment classes, like "neutral" (class 2) being confused with "positive" (class 1) and vice versa. The matrix indicates that while the model generally achieves good classification performance, particularly for the more populated classes, there are certain areas where confusion between adjacent sentiment classes could be reduced to further improve the model's accuracy and reliability.

The two-line Figure 6 shows the training and validation accuracy (Figure 6(a)) and loss (Figure 6(b)) of a model over five epochs. In the accuracy plot, the training accuracy (green line) shows a rapid increase from around 0.89 to just over 0.98 within the first epoch and continues to improve slightly, reaching nearly 1.0 by the fourth epoch. The validation accuracy (red line) also increases steadily, peaking near 0.98 around the second epoch, but then flattens and shows a slight decline towards the end, indicating potential overfitting. In the loss plot, the training loss (green line) decreases sharply from approximately 0.39 to under 0.1 within the first epoch and continues to decrease to around 0.02 by the fourth epoch, suggesting effective learning. Meanwhile, the validation loss (red line) also declines initially from about 0.1 to under 0.05 but starts to increase slightly after the second epoch, which could signal the onset of overfitting where the model performs well on the training data but starts to lose generalization capability on unseen data. Together, these plots suggest that while the model learns effectively during training, it might begin to overfit as training progresses.

Figure 7 presents the model's evaluation metrics, specifically its accuracy and loss, in a single visual representation. The blue bar on the left represents the accuracy of the model, which is very high, approaching 1.0, indicating that the model correctly predicted the outcomes for nearly all cases in the dataset. In contrast, the red bar on the right represents the loss, which is relatively low, close to 0, suggesting that the model's predictions are generally accurate and that it minimizes the difference between the predicted and actual values effectively. This combination of high accuracy and low loss demonstrates that the model performs exceptionally well, indicating effective learning and potentially strong generalization to new data. However, without additional context on validation performance, it is unclear whether these results are robust against overfitting or data bias.

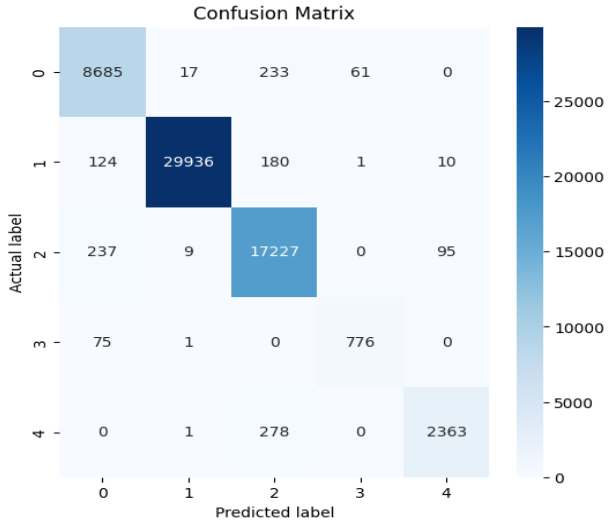
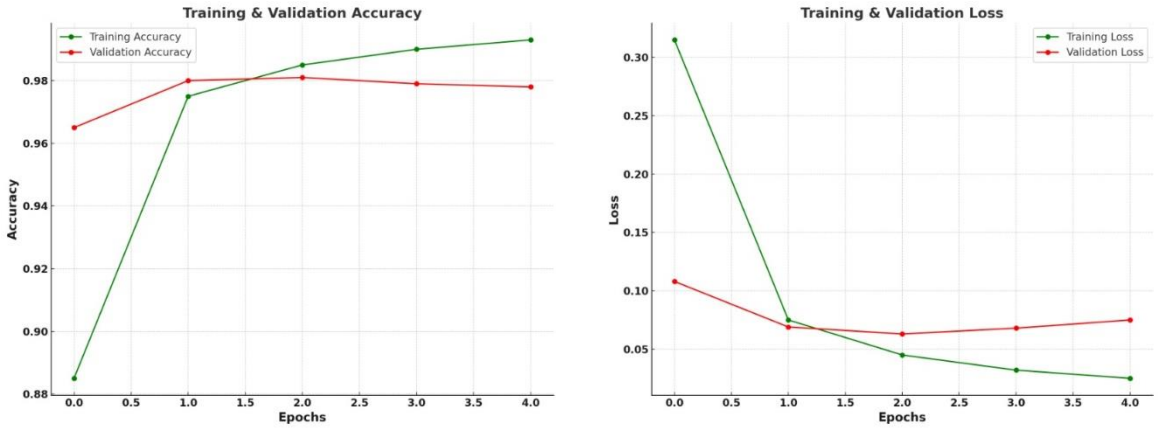


Figure 5. BI-LSTM confusion matrix



(a) (b)

Figure 6. Model over five epochs of (a) training and validation accuracy and (b) training and validation loss

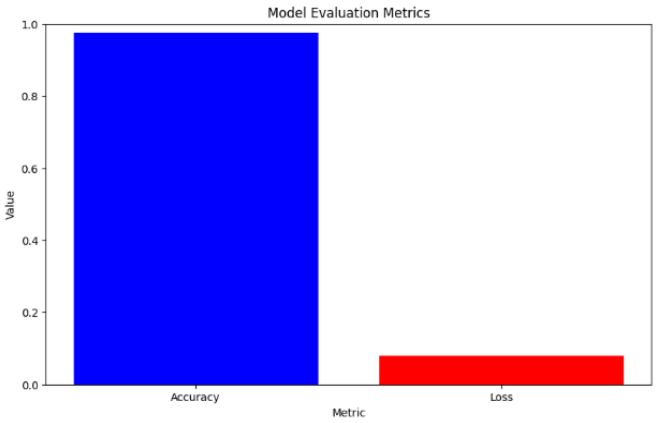


Figure 7. The model's evaluation metrics, specifically its accuracy and loss

4.2. Result

Table 1 and Figure 8 illustrate the accuracy of different Word2Vec architectures CBOW, Skip-gram, and GloVe across various models: random forest (RF), NB, LSTM, and the proposed model. The proposed model achieves the highest accuracy for CBOW at 99.02%, Skip-gram at 96.48%, and GloVe at 98.64%, outperforming all other models in each architecture. The LSTM model shows strong performance with GloVe (89.72%) and CBOW (85.36%), while skip-gram records a slightly lower accuracy (84.96%). NB performs better with GloVe (89.69%) than with CBOW (77.18%) and Skip-gram (83.2%). The RF model has the lowest accuracy across all architectures, with its highest accuracy for GloVe (80.75%). The graph highlights that the proposed model consistently outperforms others, while LSTM also performs well, particularly with GloVe embeddings.

Table 2 and Figure 9 compare the accuracy of different models RF, NB, LSTM, and a proposed model using two evaluation methods: hierarchical SoftMax and negative sampling. Across all models, the hierarchical SoftMax method consistently results in higher accuracy than negative sampling. The proposed model achieves the highest accuracy with both methods, reaching 98.64% with hierarchical SoftMax and 96.83% with negative sampling, indicating its superior performance. LSTM also shows strong performance, with an accuracy of 93.56% using hierarchical SoftMax and 91.75% with negative sampling. NB has an accuracy of 83.1% with hierarchical SoftMax and 78.8% with negative sampling, while RF records the lowest accuracies at 82.76% for hierarchical SoftMax and 72.49% for negative sampling. Overall, the data suggests that hierarchical SoftMax is a more effective evaluation method for achieving higher accuracy across all models.

Table 1. The accuracy of different Word2Vec architectures

Word2Vec architectures (accuracy)			
Models	CBOW	Skip-gram	GloVe
RF	72.48	79.48	80.75
NB	77.18	83.2	89.69
LSTM	85.36	84.96	89.72
Proposed	99.02	96.48	98.64

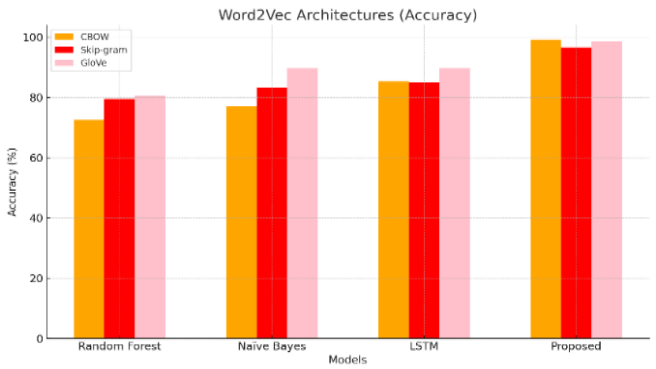


Figure 8. The accuracy of different Word2Vec architectures

Table 2. The accuracy of different evaluation methods

	Evaluation methods (accuracy)	
Models	Hierarchical SoftMax	Negative sampling
RF	82.76	72.49
NB	83.1	78.8
LSTM	93.56	91.75
Proposed	98.64	96.83

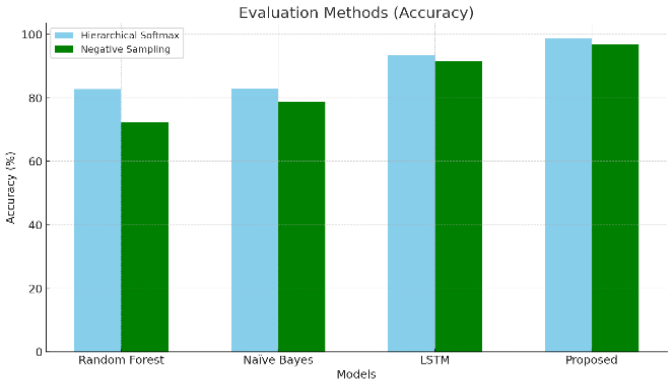


Figure 9. The accuracy of different evaluation methods

Table 3 and Figure 10 present the accuracy of different models RF, NB, LSTM, and a proposed model across three vector dimensions: 100, 200, and 300. As the vector dimension increases, all models show an improvement in accuracy. The proposed model demonstrates the highest accuracy at each dimension, achieving 94.86% at 100 dimensions, 97.38% at 200 dimensions, and 98.46% at 300 dimensions, reflecting significant gains as the dimension size increases. LSTM also shows a notable upward trend, with accuracy rising from 86.37% at 100 dimensions to 89.46% at 200 dimensions and reaching 91.85% at 300 dimensions. NB exhibits a modest increase in accuracy, from 79.4% at 100 dimensions to 79.1% at 200 dimensions and 80.34% at 300 dimensions. RF shows the least accuracy but still improves from 72.36% at 100 dimensions to 76.48% at 200 dimensions and 78.42% at 300 dimensions. Overall, the table indicates that higher vector dimensions generally enhance the performance of all models, with the proposed model achieving the most significant accuracy gains.

Table 3. The accuracy of different vector dimensions

	Vector dimension (accuracy)		
Models	100	200	300
RF	72.36	76.48	78.42
NB	79.4	79.1	80.34
LSTM	86.37	89.46	91.85
Proposed	94.86	97.38	98.46

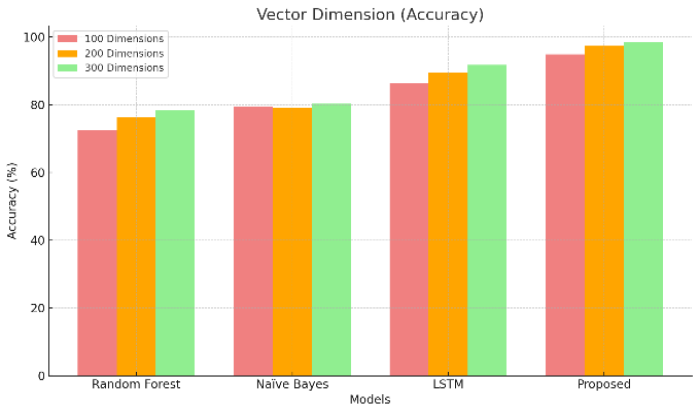


Figure 10. The accuracy of different vector dimensions

Table 4 and Figure 11 illustrate the impact of different dropout values (0.2, 0.5, and 0.8) on the accuracy of various models: RF, NB, LSTM, and a proposed model. The results show that a lower dropout value of 0.2 generally yields the highest accuracy across all models. The proposed model achieves the highest accuracy of 98.36% with a 0.2 dropout, which decreases to 96.14% at a 0.5 dropout and further drops to 92.49% at a 0.8 dropout, indicating that higher dropout values reduce accuracy. Similarly, LSTM shows a decrease in accuracy from 83.45% at a 0.2 dropout to 82.36% at 0.5 and 81.95% at 0.8. NB also performs best with a 0.2 dropout at 80.85% accuracy, declining to 80.1% at 0.5 and 77.69% at 0.8. RF follows the same trend, with its accuracy reducing from 78.65% at 0.2 to 77.36% at 0.5 and 76.39% at 0.8. Overall, the data suggests that lower dropout values are more effective for maintaining higher accuracy across all models.

Table 5 and Figure 12 compare the accuracy of different models RF, NB, LSTM, and a proposed model using two pooling techniques: max pooling and average pooling. For all models, the average pooling technique results in slightly higher accuracy than max pooling. The proposed model shows the most significant improvement, achieving an accuracy of 99.03% with average pooling compared to 96.56% with max pooling. LSTM also benefits from average pooling, reaching an accuracy of 83.56%, up from 81.36% with max pooling. NB shows a minor increase in accuracy from 80.14% with max pooling to 80.21% with average pooling, while random forest's accuracy improves from 78.24% to 79.36%. Overall, the table indicates that average pooling generally provides better accuracy across all models, with the proposed model achieving the highest performance using this technique.

Table 6 and Figure 13 show the accuracy of different models RF, NB, LSTM, and a proposed model at two learning rates: 0.0001 and 0.001. At a lower learning rate of 0.0001, the proposed model achieves the highest accuracy at 94.53%, followed by LSTM at 78.1%, NB at 75.29%, and RF at 72.37%. When the learning rate is increased to 0.001, all models show improved accuracy, with the proposed model reaching the highest accuracy of 99.05%. LSTM also improves to 82.5%, NB to 81.82%, and RFt to 79.38%. These results suggest that increasing the learning rate generally enhances the performance of all models, with the proposed model consistently outperforming the others by a significant margin at both learning rates.

Table 4. The accuracy of different dropout values (0.2, 0.5, and 0.8)

	Dropout values (accuracy)		
Models	0.2	0.5	0.8
RF	78.65	77.36	76.39
NB	80.85	80.1	77.69
LSTM	83.45	82.36	81.95
Proposed	98.36	96.14	92.49

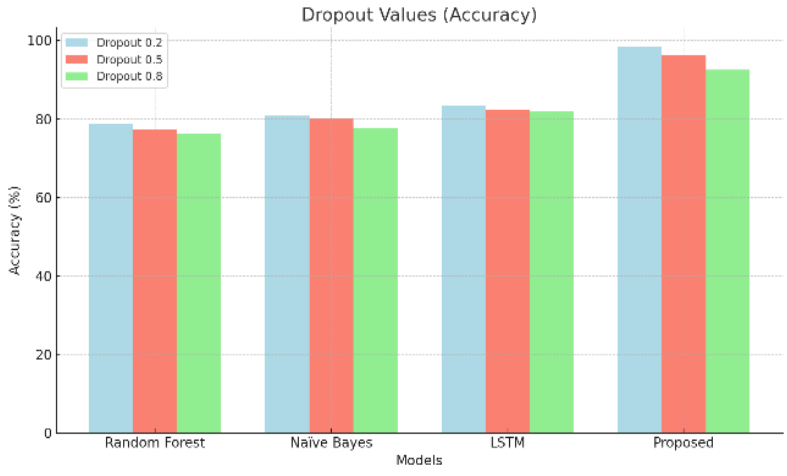


Figure 11. The accuracy of different dropout values (0.2, 0.5, and 0.8)

Table 5. The accuracy of different pooling techniques

	Pooling technique (accuracy)	
Models	Max	Avg
RF	78.24	79.36
NB	80.14	80.21
LSTM	81.36	83.56
Proposed	96.56	99.03

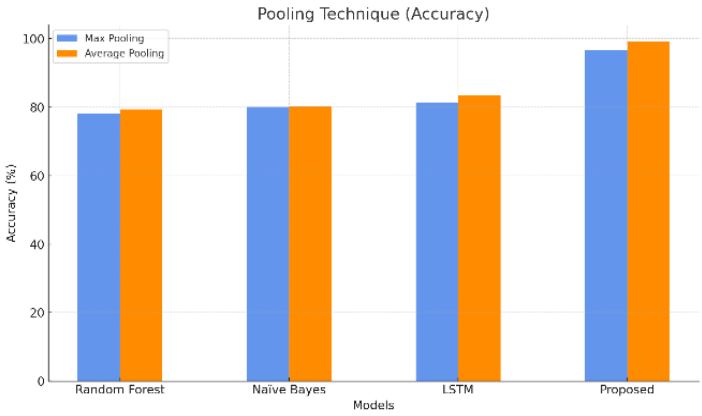


Figure 12. The accuracy of different pooling techniques

Table 6. The accuracy of different two learning rates: 0.0001 and 0.001

Learning rate (accuracy)		
Models	0.0001	0.001
RF	72.37	79.38
NB	75.29	81.82
LSTM	78.1	82.5
Proposed	94.53	99.05

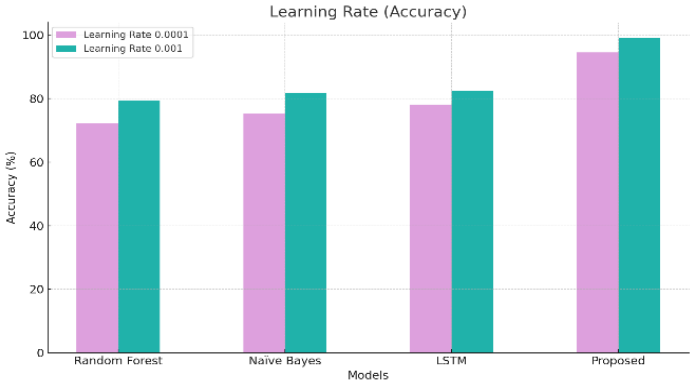


Figure 13. The accuracy of different two learning rates: 0.0001 and 0.001

Table 7 and Figure 14 present the performance metrics of four models RF, NB, LSTM, and a proposed model across four evaluation criteria: accuracy, precision, recall, and F1-score. The proposed model demonstrates the highest performance, achieving an accuracy of 99.23%, precision of 98.85%, recall of 98.64%, and F1-score of 98.17%, indicating superior prediction quality compared to the other models. LSTM follows with a high accuracy of 93.49%, precision of 93.98%, recall of 94.32%, and an F1-score of 92.99%, showcasing strong performance, particularly in recall. NB achieves an accuracy of 91.74%, precision of 92.67%, recall of 92.24%, and F1-score of 91.67%, while RF shows the lowest but still competitive metrics, with an accuracy of 89.65%, precision of 90.45%, recall of 90.67%, and F1-score of 89.15%. Overall, the proposed model outperforms the others in all metrics, reflecting its effectiveness in handling the classification task.

Table 7. The accuracy of across four evaluation criteria

Models	Accuracy	Precision	Recall	F1-score
RF	89.65	90.45	90.67	89.15
NB	91.74	92.67	92.24	91.67
LSTM	93.49	93.98	94.32	92.99
Proposed	99.23	98.85	98.64	98.17

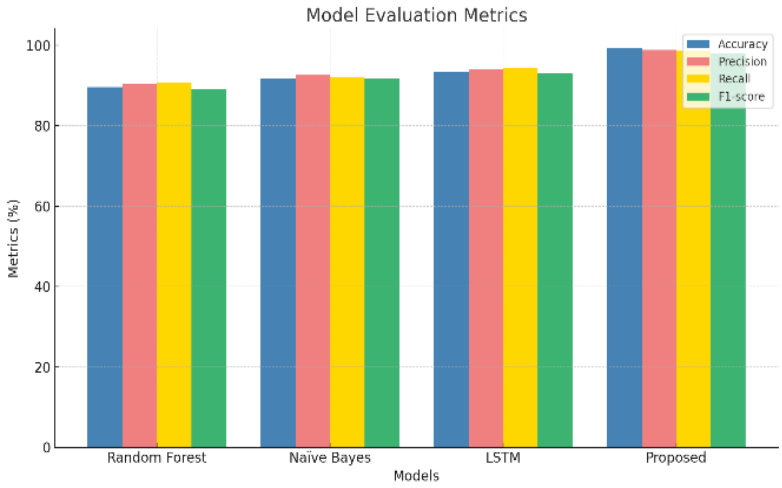


Figure 14. The accuracy of across four evaluation criteria

5. CONCLUSION

Based on the evaluation of various models and techniques for sentiment analysis of the vaccine dataset, the findings reveal significant insights into model performance across different configurations. The proposed model consistently outperforms others in all settings, achieving the highest accuracy, precision, recall, and F1-scores. This model shows superior results, especially with a learning rate of 0.001 with 99.05% accuracy, average pooling with 99.03% accuracy, and hierarchical SoftMax with 98.64% accuracy, indicating its robustness in handling diverse sentiment analysis tasks. LSTM also demonstrates strong performance, particularly with higher vector dimensions of 91.85% accuracy at 300 dimensions and lower dropout rates of 83.45% accuracy at 0.2 dropout, suggesting it is effective in capturing complex dependencies within the text. NB and RF perform moderately well, with NB showing slight improvement using average pooling and lower dropout values, while RF remains less accurate overall but benefits from methods like hierarchical SoftMax. These results suggest that sophisticated models such as the proposed model, leveraging appropriate vector dimensions, pooling techniques, and learning rates, are most effective for sentiment analysis in vaccine-related datasets. Simpler models like NB and RF, while useful, are less accurate, especially with suboptimal configurations.

ACKNOWLEDGMENTS

The authors would like to thank the Department of Computer Science and Engineering, Oriental University, Indore, India, for providing the necessary support and infrastructure for this research.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Monika Verma	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
Sandeep Monga		✓				✓			✓	✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

The research related to animal use has complied with all the relevant national regulations and institutional policies for the care and use of animals.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/datasets/abdulmalik1518/the-ultimate-cars-dataset-2024.

REFERENCES

[1] M. Arbane, R. Benlamri, Y. Brik, and A. D. Alahmar, "Social media-based COVID-19 sentiment classification model using Bi-LSTM," *Expert Systems with Applications*, vol. 212, 2023, doi: 10.1016/j.eswa.2022.118710.

[2] M. Qorib, T. Oladunni, M. Denis, E. Ososanya, and P. Cotae, "COVID-19 vaccine hesitancy: text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset," *Expert Systems with Applications*, vol. 212, 2023, doi: 10.1016/j.eswa.2022.118715.

[3] L. Samaras, E. García-Barriocanal, and M. A. Sicilia, "Sentiment analysis of COVID-19 cases in Greece using Twitter data," *Expert Systems with Applications*, vol. 230, 2023, doi: 10.1016/j.eswa.2023.120577.

[4] P. H. Prastyo, A. S. Sumi, A. W. Dian, and A. E. Permanasari, "Tweets responding to the Indonesian Government's handling of COVID-19: sentiment analysis using SVM with normalized poly kernel," *Journal of Information Systems Engineering and Business Intelligence*, vol. 6, no. 2, p. 112, 2020, doi: 10.20473/jisebi.6.2.112-122.

[5] M. K. Zuhanda *et al.*, "Supply chain strategy during the COVID-19 terms: sentiment analysis and knowledge discovery through text mining," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 2, pp. 1120–1127, 2023, doi: 10.11591/ijeecs.v30.i2.pp1120-1127.

[6] J. T. H. Kong, F. H. Juwono, I. Y. Ngu, I. G. D. Nugraha, Y. Maraden, and W. K. Wong, "A mixed Malay–English language COVID-19 Twitter dataset: a sentiment analysis," *Big Data and Cognitive Computing*, vol. 7, no. 2, 2023, doi: 10.3390/bdcc7020061.

[7] N. Thakur, "Sentiment analysis and text analysis of the public discourse on Twitter about COVID-19 and MPox," *Big Data and Cognitive Computing*, vol. 7, no. 2, 2023, doi: 10.3390/bdcc7020116.

[8] D. A. Widodo, N. Iksan, and B. Sunarko, "Sentiment analysis of Twitter media for public reaction identification on COVID-19 monitoring system using hybrid feature extraction method," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 1, pp. 92–99, 2023.

[9] C. Stefanis *et al.*, "Sentiment analysis of epidemiological surveillance reports on COVID-19 in Greece using machine learning models," *Frontiers in Public Health*, vol. 11, 2023, doi: 10.3389/fpubh.2023.1191730.

[10] M. Inoue *et al*., "Opinion and sentiment analysis of palliative care in the era of COVID-19," *Healthcare (Switzerland)*, vol. 11, no. 6, 2023, doi: 10.3390/healthcare11060855.

[11] M. Qorib, T. Oladunni, M. Denis, E. Ososanya, and P. Cotae, "COVID-19 vaccine hesitancy: a global public health and risk modelling framework using an environmental deep neural network, sentiment classification with text mining and emotional reactions from COVID-19 vaccination tweets," *International Journal of Environmental Research and Public Health*, vol. 20, no. 10, 2023, doi: 10.3390/ijerph20105803.

[12] X. Chen, Z. Wang, and X. Di, "Sentiment analysis on multimodal transportation during the COVID-19 using social media data," *Information (Switzerland)*, vol. 14, no. 2, 2023, doi: 10.3390/info14020113.

[13] J. C. Mba and M. Biyase, "Threshold of depression measure in the framework of sentiment analysis of tweets: managing risk during a crisis period like the COVID-19 pandemic," *Journal of Risk and Financial Management*, vol. 16, no. 2, 2023, doi: 10.3390/jrfm16020115.

[14] N. Braig, A. Benz, S. Voth, J. Breitenbach, and R. Buettner, "Machine learning techniques for sentiment analysis of COVID-19-related Twitter data," *IEEE Access*, vol. 11, pp. 14778–14803, 2023, doi: 10.1109/ACCESS.2023.3242234.

[15] H. Wang *et al.*, "Patterns of diverse and changing sentiments towards COVID-19 vaccines: a sentiment analysis study integrating 11 million tweets and surveillance data across over 180 countries," *Journal of the American Medical Informatics Association*, vol. 30, no. 5, pp. 923–931, 2023, doi: 10.1093/jamia/ocad029.

[16] S. Ahmed *et al.*, "Temporal analysis and opinion dynamics of COVID-19 vaccination tweets using diverse feature engineering techniques," *PeerJ Computer Science*, vol. 9, pp. 1–29, 2023, doi: 10.7717/PEERJ-CS.1190.

[17] S. C. Hernández, M. P. Tzili Cruz, J. M. E. Sánchez, and A. P. Tzili, "Deep learning model for COVID-19 sentiment analysis on Twitter," *New Generation Computing*, vol. 41, no. 2, pp. 189–212, 2023, doi: 10.1007/s00354-023-00209-2.

[18] S. P. Che, X. Wang, S. Zhang, and J. H. Kim, "Effect of daily new cases of COVID-19 on public sentiment and concern: deep learning-based sentiment classification and semantic network analysis," *Journal of Public Health (Germany)*, vol. 32, no. 3, pp. 509–528, 2024, doi: 10.1007/s10389-023-01833-4.

[19] A. Umair and E. Masciari, "Sentimental and spatial analysis of COVID-19 vaccines tweets," *Journal of Intelligent Information Systems*, vol. 60, no. 1, pp. 1–21, 2023, doi: 10.1007/s10844-022-00699-4.

[20] V. Jain and K. L. Kashyap, "Ensemble hybrid model for Hindi COVID-19 text classification with metaheuristic optimization algorithm," *Multimedia Tools and Applications*, vol. 82, no. 11, pp. 16839–16859, 2023, doi: 10.1007/s11042-022-13937-2.

[21] M. Sharaf, E. E. D. Hemdan, A. El-Sayed, and N. A. El-Bahnasawy, "An efficient hybrid stock trend prediction system during COVID-19 pandemic based on stacked-LSTM and news sentiment analysis," *Multimedia Tools and Applications*, vol. 82, no. 16, pp. 23945–23977, 2023, doi: 10.1007/s11042-022-14216-w.
[22] Y. Madani, M. Erritali, and B. Bouikhalene, "A new sentiment analysis method to detect and Analyse sentiments of COVID-19 Moroccan tweets using a recommender approach," *Multimedia Tools and Applications*, vol. 82, no. 18, pp. 27819–27838, 2023, doi: 10.1007/s11042-023-14514-x.
[23] A. Vishwakarma and M. Chugh, "COVID-19 vaccination perception and outcome: society sentiment analysis on Twitter data in India," *Social Network Analysis and Mining*, vol. 13, no. 1, 2023, doi: 10.1007/s13278-023-01088-7.
[24] M. Chaudhary, K. Kosyluk, S. Thomas, and T. Neal, "On the use of aspect-based sentiment analysis of Twitter data to explore the experiences of African Americans during COVID-19," *Scientific Reports*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-37592-1.
[25] W. Ahmad, B. Wang, P. Martin, M. Xu, and H. Xu, "Enhanced sentiment analysis regarding COVID-19 news from global channels," *Journal of Computational Social Science*, vol. 6, no. 1, pp. 19–57, 2023, doi: 10.1007/s42001-022-00189-1.
[26] O. Abiola, A. Abayomi-Alli, O. A. Tale, S. Misra, and O. Abayomi-Alli, "Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser," *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, 2023, doi: 10.1186/s43067-023-00070-9.
[27] B. S. Ainapure *et al.*, "Sentiment analysis of COVID-19 tweets using deep learning and lexicon-based approaches," *Sustainability (Switzerland)*, vol. 15, no. 3, 2023, doi: 10.3390/su15032573.
[28] M. Tran, C. Draeger, X. Wang, and A. Nikbakht, "Monitoring the well-being of vulnerable transit riders using machine learning based sentiment analysis and social media: lessons from COVID-19," *Environment and Planning B: Urban Analytics and City Science*, vol. 50, no. 1, pp. 60–75, 2023, doi: 10.1177/23998083221104489.
[29] R. Catelli, S. Pelosi, C. Comito, C. Pizzuti, and M. Esposito, "Lexicon-based sentiment analysis to detect opinions and attitude towards COVID-19 vaccines on Twitter in Italy," *Computers in Biology and Medicine*, vol. 158, 2023, doi: 10.1016/j.compbiomed.2023.106876.
[30] W. Liang, X. Chen, S. Huang, G. Xiong, K. Yan, and X. Zhou, "Federal learning edge network based sentiment analysis combating global COVID-19," *Computer Communications*, vol. 204, pp. 33–42, 2023, doi: 10.1016/j.comcom.2023.03.009.
[31] M. Costola, M. Nofer, O. Hinz, and L. Pelizzon, "Machine learning sentiment analysis, COVID-19 news and stock market reactions," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3690922.

BIOGRAPHIES OF AUTHORS

Monika Verma is working as an assistant professor in the Department of Computer Science and Engineering at the Greater Noida Institute of Technology, Greater Noida. Before that, she had more than years of teaching experience in several engineering colleges such as PIEMR, Indore, and SVVV. She completed her B.E. (CSE) from Shri Vaishnav Institute of Technology and Science, Indore (M.P.), and M. Tech (NM&IS) from SCSIT, DAVV, Indore (M.P.). And Ph.D. degree in Computer Science and Engineering from Oriental University, Indore (M.P.). Her research field is machine learning and deep learning. She can be contacted at email: monikaverma03@rediff.com.

Sandeep Monga is an Associate Professor in the Computer Science and Engg department at the Oriental Institute of Science and Technology. He had a Ph.D. from Rajiv Gandhi Proudyogiki Vishwavidyalaya (State Technological University of Madhya Pradesh) and has over 27 years of experience in academia and research. His research contributions include publications in Scopus and peer-reviewed journals. He is a member of professional societies like ISTE, ACM, and CSI. He can be contacted at email: smonga6@gmail.com.