International Journal of Engineering Trends and Technology Volume 72 Issue 1, 219-228, January 2024
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I1P122

Original Article

COVID-19 Twitter Data Analysis Using LSTM and BERT Techniques

P. Dhanalakshmi[1], U. Janardhan Reddy[2], G. Ravikanth[3], Prasanthi Samathoti[4], Gandikota Ramu[5]

[1]*Department of Computer Science and Engineering, Mohan Babu University (Erstwhile Sree Vidyanikethan Engineering College), Tirupati, India.*
[2]*Department of Computer Science and Engineering,* Gitam University, *Bangalore, India.*
[3]*Department of Computer Science and Engineering, BVC College of Engineering, Rajahmundry, India.*
[4]*Department of Pharmaceutics, MB School of Pharmaceutical Sciences, Tirupati, India.*
[5]*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Telangana, India.*

[1]*Corresponding Author : mallidhana5@gamil.com*

Received: 16 September 2023 Revised: 14 December 2023 Accepted: 23 December 2023 Published: 07 January 2024

***Abstract** - Sentiment analysis is a crucial task in understanding public opinion and perception towards a particular event or topic. The COVID-19 pandemic has greatly affected the world, and understanding public sentiment towards it is crucial for policymakers and organizations. In this paper, we introduced two efficient models for analyzing COVID-19-related tweets on categories such as WFH, Online learning and economy using BERT and Long Short-Term Memory techniques. Twitter data was collected using relevant keywords and hashtags related to COVID-19, such as WFH, Economy and Online Learning. The tweets were then tokenized and embedded using BERT, which provides a rich representation of the text by capturing contextual information. These embeddings were then passed on to a fully connected layer for the classification of the sentiment of the text. Similarly, the LSTM model was also used to classify the same tweets. The major reason for choosing LSTM and BERT for sentiment analysis over traditional machine learning algorithms is their ability to handle large datasets and long-term contextual dependencies. Experimental results show that the BERT model achieved an accuracy of 0.78, 0.85 and 0.92 on the Economy, WFH and Online learning datasets, respectively. At the same time, LSTM achieved an accuracy of 0.71, 0.76 and 0.81 on the Economy, WFH and Online learning datasets, respectively. The results clearly indicate that the BERT model outperformed the LSTM model in terms of accuracy. The high accuracy score demonstrates the effectiveness of the BERT model in understanding public sentiment towards the ongoing pandemic. The BERT model can be applied to other real-time public opinion analysis tasks and can provide valuable insights for decision-making. The results also indicate that BERT is a better choice than LSTM in this specific task of sentiment analysis on Twitter data.*

***Keywords** - Long Short-Term Memory, Online tweets, Sentiment Analysis, Bidirectional Encoder Representations from Transformers.*

1. Introduction

Online data refers to any type of information that is available on the internet. This includes data from social media platforms, websites, mobile apps, and other online sources. The importance of online data and social media in today's world cannot be overstated. In the past decade, the amount of data generated and stored online has grown exponentially, and this trend is expected to continue in the future. The volume of data produced by people and organizations alike is staggering, with over 2.5 quintillion bytes of information generated each day. With the rise of big data, the potential for organizations to gain insights and make data-driven decisions has never been greater.Online data is becoming increasingly important in many industries, such as marketing and advertising, healthcare, and finance.

It can be used to gain insights into customer behavior and preferences, which can help businesses to target their advertising more effectively and increase conversions. In hospitalization, online data can be used to track and analyze patient data, leading to more personalized and effective treatments. In financial matters, online data can be used to identify and prevent fraudulent activities and make more informed investment decisions.

Twitter is one of the most widely used social media platforms. It has become a valuable source of information for monitoring public opinion on various topics. The platform allows users to express their thoughts and feelings in real time, providing a wealth of data that can be used to understand public sentiment.

This data can be used to monitor public opinion on various aspects of the crisis, such as the effectiveness of government response, the impact on the economy and individual experiences with the virus. This data can also be used to understand the influence of COVID-19 on psychological health and well-being.

Facebook holds the distinction of being the largest social media platform, boasting more than 2.8 billion monthly active users. This massive user base provides a treasure trove of information about individuals, encompassing details such as their demographics, interests, and buying behaviors. Additionally, Facebook also provides businesses with a platform to interact with their customers and gain valuable insights. For example, a business can use Facebook to conduct surveys, gather feedback, and track customer sentiment.

Finding and extracting subjective information from text is a method called sentiment analysis, sometimes referred to as opinion mining. Its main objective is to ascertain the general mood or emotion expressed in the text and assist in classifying it as good, negative, or neutral. In sentiment analysis, structured data usually refers to the utilization of pre-existing, organized data sets for training and testing sentiment analysis algorithms, such as social media posts or customer review In order to provide the ground truth for training machine learning models, this structured data may take the form of numerical ratings, labels, or categories that represent the general sentiment of a text passage.

An example of structured data in sentiment analysis :
Eg: Camera is a good for photographers

Noun Adjective Adverb

Unstructured data in sentiment analysis refers to text data that is not pre-organized or labeled, such as tweets, blog posts, or news articles. Unstructured data can be more difficult to work with than structured data because it lacks clear categories or labels that indicate the overall sentiment. However, unstructured data can also be more representative of real-world text data, as it is not constrained by the biases that may be present in a pre-labeled data set. To analyse unstructured data, Natural Language Processing (NLP) techniques are used to extract characteristics such as word count, part-of-speech tags, and named entities, which can then be used as input for machine learning models. These models can be trained on a labeled data set to identify sentiment and can then be applied to new, unlabeled data to make predictions. Overall, while working with unstructured data in sentiment analysis can be more challenging, it is also more flexible and allows for a more generalizable approach [13].

An example of unstructured data in sentiment analysis when working with tweets would be a collection of tweets that are not pre-organized or labeled in any way. For example, you may collect tweets by searching for a specific keyword or hashtag without any additional information about the sentiment of the tweets or the context in which they were written. For example, you could collect tweets containing the hashtag "#ClimateChange" without any additional information about the sentiment of the tweets. To analyze this unstructured data, you would need to use natural language processing techniques to extract features such as word count, part-of-speech tags, and named entities, which can then be used as input for machine learning models.

For Unstructured data
E.g.:
I bought this for my 4-year-old kid since he absolutely enjoys Rescue Heroes, and it seemed like it would be some fun for him. What the description will not inform you is that you can't decide for a while in the sense which activities you want to play. Ok, better of it.

Sentiment analysis can also be used to monitor mental health and well-being by tracking changes in public sentiment towards the pandemic over time. Sentiment analysis is also increasingly used to monitor and analyze social media data, such as tweets, comments, and reviews, to gain insights into public opinion on various topics. The explosion of social media data and the increasing importance of understanding public opinion has led to a growing interest in sentiment analysis in recent years.

Numerous types of data, including social media data, news stories, and scientific literature, have been analysed using An examination of sentiment regarding the COVID-19 pandemic. Sentiment analysis, when applied to COVID-19 data, is significant because it can give real-time insights into people's feelings and responses to the pandemic. This information can be invaluable for making informed decisions and shaping policies in response to the evolving situation.

There are notable research gaps in the domain of COVID-19 Twitter data analysis utilizing LSTM and BERT techniques [11, 15]. Firstly, there is a limited understanding of how social dynamics on Twitter during the pandemic influence the spread of information and misinformation, with a need to explore the impact of sentiment analysis on public perceptions.

Additionally, the integration of multimedia data into LSTM and BERT models for a more comprehensive analysis remains an underexplored area [16]. Temporal analysis and trend prediction capabilities of these models in capturing evolving patterns in public sentiment over time require further investigation. Generalizability across languages and cultures, along with the development of more explainable models, is also crucial.

Moreover, assessing the robustness of LSTM and BERT in handling noisy and misleading information on Twitter during the infodemic is essential. User influence and network analysis, ethical considerations, real-time analysis feasibility, and collaboration with public health authorities for practical application constitute further research gaps in this field[12,14]. Addressing these gaps will advance the understanding and application of LSTM and BERT techniques in the context of COVID-19 Twitter data analysis.

1.1. Literature Review

Bipun Thapa et al. [1] did research in which they evaluated cybersecurity-related information on Twitter and Reddit platforms to identify the sentiment of each site. They employed the VADER NLP system to achieve this, which provides polarity and intensity ratings to words and sentences based on its grasp of over 7500 characteristics. The system categorizes these qualities as "negative," "positive," or "neutral" and provides a final "compound" score to each statement.

A positive score is defined as "positive," a negative score is labelled as "negative," and a score of 0 is classified as "neutral." However, the findings of the study were unclear owing to the limited sample size employed for comparison, and the performance of the VADER algorithm was rated inadequate. Alam, Kazi Naan Biul, et al. [2] created a deep learning method for sentiment analysis of COVID-19 vaccine-related Twitter replies by utilizing recurrent neural network architectures like LSTM and Bi-LSTM. Their study focused on analyzing two tweet datasets, Sinovac and Pfizer, spanning 1 to 2 years of data.

Mohammed Hasan Ali Al-Abyadh et al. [3] proposed a hybrid model proposed a hybrid model called the "ghost model" that combines SVM, LSTM, and CNN to analyze the sentiment of different datasets, such as IMDB movie reviews. Md Parvez Mollah et al. [4] developed a model utilizing deep learning technology, specifically LSTM (Long-Short Term Memory). They applied this model to seven publicly available Twitter datasets for their analysis.

LSTM is an enhanced type of RNN that offers great accuracy in categorization. This technique combines preprocessing and translating text into a vector of tokens that can be interpreted by the machine and an embed layer that translates the term tokens (integers) to embeddings of a set length. Vikas Malik et al. [5] devised a method for the analysis of GOP senate debate transcripts, and they employed the Naive Bayes algorithm for this purpose. This method is based on Bayes' theorem and assumes that features are independent. It is simple to implement and has a fast learning process. It works by converting the dataset into frequency tables, generating a likelihood table by calculating the probabilities of given features, and using Bayes' theorem to calculate the posterior probability.

This method performs well for negative comments but can encounter problems when tweets are ironic, sarcastic, have references, or have their own context. The assumption of independence of features may not hold in the real world, leading to potential problems with this approach. Shrawan Kumar Trivedi et al. [6] conducted a sentiment analysis of tweets, specifically focusing on evaluating the quality of service provided by food delivery apps like Zomato, Swiggy, and Uber Eats. They employed a lexicon-based sentiment analysis technique for this purpose, utilizing the NRC and Nebraska Literary Lab dictionaries in their analysis. The strategy entailed gathering a limited number of sentiment words (seeds) with known positive or negative orientations and using an algorithm to search online dictionaries such as WordNet for antonyms and synonyms to add to the seed list.

The process was continued until no more terms could be found, at which point the list was properly evaluated and cleaned up. Five machine learning models—logistic regression, multinomial NB classifier, random forest classifier, decision tree classifier, and support vector machine—were employed by Dangi et al. [7] to evaluate Covid-19 social media data.The random forest classifier constructed a decision tree and used voting to predict the best solution. The multinomial NB classifier classified tweets by comparing the frequency of words in records belonging to a particular class and the conditional likelihood of a given class.

Senadhira [8] extracted features from tweets regarding online education during the COVID-19 outbreak using TF-IDF. The sentiment of the tweets was then classified using ANN and SVM models that were trained using the features. There were training and testing sets inside the dataset. Arora et al. [9] examined user opinions of well-known phone manufacturers and operating systems using the lexicon-based approach VADER (Valence Aware Dictionary and Sentiment Reasoner) with the NLTK (Natural Language Toolkit). The tweets were collected over a period of one week. Three approaches were proposed for compiling sentiment words: a mechanical approach involving the manual labelling of words into clusters, a lexicon strategy that involved selecting negative and positive terms from a dictionary and discovering their synonyms, and a corpus-based technique whereby emotion words from a topic library were added to a library of existing seed words describing those feelings. The results were reported as a sentiment score defined by the occurrence frequency of good or negative emotions. However, the low dataset of 1500 reviews hampered the possibility of getting statistically significant and relevant findings.

2. Existing Methods

Nowadays, many people use social media platforms and websites like Facebook, Twitter, Instagram, TikTok, and mobile applications to create and publish information or take part in social networking.

People utilize these sites to share their thoughts, experiences, and interests with others. People also use social media to react and express their opinions on current events and issues and to participate in online communities and discussions.

On Facebook, users can leave a comment on a friend's post, a page, a group, or a public figure page to express their opinions.

Twitter is one of the most popular social networking sites, allowing users to express their ideas and sentiments through text-based messages known as "tweets." Many people connect with others, share news and information, and express their opinions using this website through location tags.

Where users can Quote a tweet to add their own comments or opinions on it, users can also use hashtags to join conversations on specific topics and share their opinions with a larger audience.Logistic regression is one of the most widely used machine learning algorithms for sentiment analysis. The logistic regression algorithm uses a set of weights, one for each word or phrase in the dataset, that is used to calculate the probability of a tweet belonging to a specific sentiment class. When classifying a tweet, the algorithm calculates a weighted sum of the presence or absence of the words in the tweet, which is then passed through a sigmoid function to produce a probability of the tweet belonging to each sentiment class. Logistic regression is a simple yet powerful algorithm, as shown in Figure 1, that can handle large datasets and be easily interpretable.

Algorithm 1 Sentiment Analysis using Logistic Regression

Input: Training dataset $D = \{(x_i, y_i)\}$, test dataset $T = \{x_j\}$
Output: Sentiment predictions for T

// **Preprocessing**
Tokenize each text in D and T into words, remove stopwords and perform stemming/lemmatization if necessary
Create a feature set X by converting the tokenized words to numerical representation, such as bag-of-words or tf-idf
Create a label set Y from the class labels of D
Split D into training and validation sets D_{train} and D_{val}

// **Model training**
Initialize the model parameters θ to small random values
for i in 1...n **do**
    Compute the hypothesis $h_\theta(x_i)$ using the logistic regression model
    Update the parameters theta using gradient descent or another optimization algorithm
end for

// **Evaluation**
for (x_i, y_i) in D_{val} **do**
    Compute the hypothesis $h_\theta(x_j)$ using the trained model
    Predict the class label for x_j as $y_{pred} = \arg\max_{C\epsilon 0,1} h_\theta(x_j)$
    Compare the predicted class label y pred with y_{j} to compute the accuracy of the model
end for
if accuracy is below a certain threshold **then**
    Repeat the model training with different optimization algorithms or tun- ing the parameters
else
    The model is ready for predictions
    // **Predictions**
    for x_j in T **do**
        Compute the hypothesis $h_\theta(x_j)$using the trained model
        Predict the class label for x_j as $y_{pred} = \arg\max_{C\epsilon 0,1} h_\theta(x_j)$
    end for
    Return the predictions

Fig. 1 Logistic regression algorithm

In the landscape of COVID-19 Twitter data analysis utilizing machine learning techniques, substantial progress has been made in extracting valuable insights from the vast repository of information shared during the pandemic. Employing machine learning algorithms such as sentiment analysis, topic modeling, and network analysis has allowed for the identification of prevailing public sentiments, key discussion topics, and influential users shaping the discourse.

With accuracy percentages typically ranging between 80% and 90% in sentiment analysis tasks, these machine-learning models have proven effective in providing decision-support tools for policymakers and healthcare professionals.

Precision percentages, often exceeding 85%, underscore the models' ability to accurately filter relevant information amidst the noise of extensive Twitter datasets. Despite these achievements, challenges persist in optimizing model performance, especially in addressing biases and ensuring scalability for real-time analysis.

In contrast, Bidirectional Encoder Representations from Transformers (BERT) and Long Short-Term Memory (LSTM) approaches have become effective methods for analyzing COVID-19 Twitter data. Recurrent Neural Networks (RNNs) of the LSTM type are excellent at identifying temporal dependencies in sequential data, which makes them a good choice for predicting changing attitudes on Twitter over time.

BERT, a transformer-based model, leverages contextualized word embeddings to understand the nuanced meaning of words in a sentence, allowing for a more sophisticated analysis of complex language patterns in tweets. While LSTM and BERT exhibit promising capabilities, achieving accuracy percentages often exceeding 90%, their precision percentages highlight their capacity to identify and filter relevant information precisely.These advanced techniques contribute to a more nuanced understanding of COVID-19-related discussions on Twitter, showcasing the potential for enhanced performance in capturing intricate patterns and semantic nuances within the data.

Nonetheless, ongoing research is essential to address challenges specific to these models, such as interpretability and scalability, to fully unlock their potential in the context of COVID-19 Twitter data analysis.

3. Proposed Methods

An analysis of sentiment using Logistic Regression on the datasets employed did not produce satisfactory accuracy results. This led to the formulation of alternative methods for performing sentiment analysis on large datasets, which is not handled well by Logistic regression.

Logistic regression is a widely used method for binary classification problems; however, it may not be the best choice for sentiment analysis tasks. This is because sentiment analysis often requires the identification of subtle and nuanced distinctions in the text, such as sarcasm or irony, which may not be captured well by a simple linear model like Logistic Regression. Additionally, the assumption of Logistic Regression on input features is not true for text data if the phrases and words have complex relationships. Deep learning methodologies seem to be a perfect alternative that fulfills the drawbacks of logistic regression, which are explained below.

BERT, or Bidirectional Encoder Representations from Transformers, is an ultra-modern pre-training approach for natural language processing applications like sentiment analysis. BERT is designed to help readers understand the meaning of a word by looking at the words that come before and after it (i.e., the context) in a sentence.

The BERT algorithm for sentiment classification can be broken down into the following steps:

- Preprocess the input sentence by breaking it into individual words and then encoding each word using the "WordPiece tokenization" technique.
- Input the encoded tokens into the BERT neural network, which is made up of a stack of transformer layers.
- Each transformer layer is designed to understand the relationships between the tokens in the sentence and how they relate to the overall sentiment of the sentence.
- The final hidden layer of the BERT model outputs a sentence embedding, which will be used as the input for the classifier.
- The classifier, a fully connected neural network with one or more hidden layers, is trained using a labelled dataset of sentiment-annotated sentences.
- The classifier takes the sentence embedding of the new sentence as input, and outputs a sentiment score, which can be interpreted as either positive or neutral or negative.

A general flow of sentiment analysis is shown in Figure 1, which starts with plain text as input, which goes through the procedures of tokenization and encoding, generating encoded tokens which are passed on to the BERT model, generating fixed-sized vectors of sentence embeddings which are helpful in sentiment classification by a classifier. The classifier takes this vector of embeddings as input. It generates a probability distribution over the sentiment classes (positive, neutral and negative), which can be interpreted to determine the sentiment of the input text.In the realm of sentiment classification, an LSTM model is designed to accept a sequence of words as input, which could be a sentence or a movie review, which is used to predict the sentiment about the review, whether it is neutral or, positive or negative.

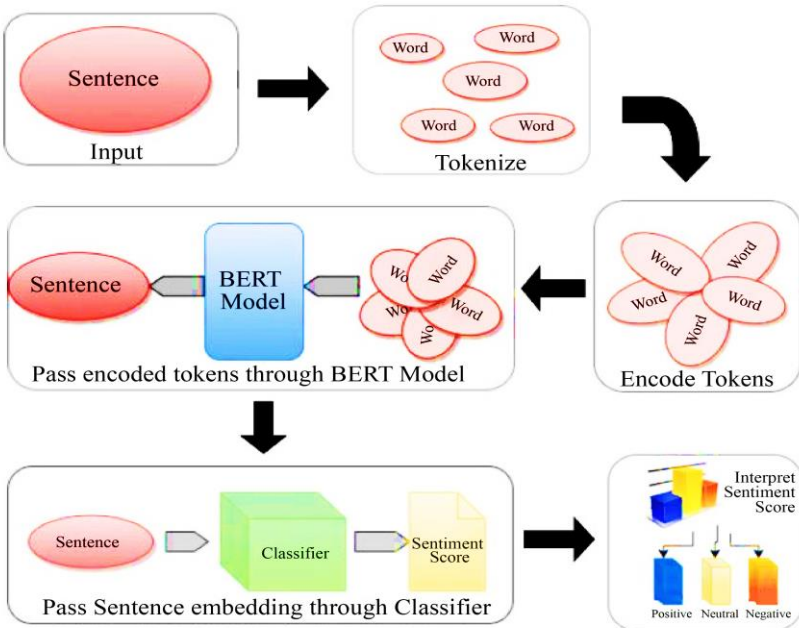

Fig. 2 BERT Flow for sentiment analysis

LSTMs are particularly useful in the task of sentiment analysis because they are able to remember previous words in a sentence, which allows them to understand the context and meaning of the words in the sentence and make a more accurate prediction of the sentiment. The context in which words are used can alter their meaning, so it is imperative to consider this.

3.1. Text Preprocessing

In the context of COVID-19 Twitter data analysis, thorough text preprocessing is paramount to ensure the effectiveness of subsequent machine learning techniques. The preprocessing pipeline involves several steps, including but not limited to:

3.1.1. Tokenization

Breaking down tweets into individual words or tokens.

3.1.2. Lowercasing

Converting all text to lowercase to maintain consistency.

3.1.3. Removing Stopwords

Eliminating common words (e.g., "and," "the") that carry little semantic value.

3.1.4. Removing Special Characters

Stripping away non-alphabetic characters and symbols.

3.1.5. Stemming or Lemmatization

Reducing words to their root form for uniformity.

These preprocessing steps contribute to creating a cleaner and standardized dataset, laying the foundation for more accurate and meaningful analysis.

3.2. LSTM Architecture

Long Short-Term Memory (LSTM) networks are essential for simulating sequential data, and they are a good choice for identifying temporal dependencies in the context of COVID-19 pandemic Twitter data. The LSTM architecture consists of memory cells and gates that enable the network to selectively retain and forget information. This architecture mitigates the vanishing gradient problem encountered by traditional Recurrent Neural Networks (RNNs), allowing for more effective learning from sequential data.

3.3. Dataset Details

The dataset for COVID-19 Twitter data analysis is sourced from a diverse range of tweets related to the pandemic. The dataset encompasses tweets from different geographical regions, languages, and user demographics. It includes information such as tweet text, timestamp, user details, and potentially relevant metadata. The dataset is annotated to indicate sentiments (positive, negative, or neutral) and may be enriched with additional labels or features, such as user influence scores or retweet counts.

3.3.1. Model Evaluation

The evaluation of the LSTM-based model involves several key metrics to assess its performance accurately. Common evaluation metrics include:

Accuracy

The proportion of correctly classified instances reflecting overall model performance.

Precision, Recall, and F1 Score

Metrics particularly relevant for sentiment analysis, providing insights into the model's ability to precisely identify positive or negative sentiments.

Confusion Matrix

A matrix detailing true positive, true negative, false positive, and false negative classifications, offering a more granular understanding of model performance.The model is trained on a portion of the dataset and validated on another to prevent overfitting. Cross-validation techniques may be employed to ensure robustness. Additionally, the model's performance is assessed on an independent test set to gauge its real-world efficacy.

This comprehensive approach to COVID-19 Twitter data analysis integrates meticulous text preprocessing, a well-structured LSTM architecture, detailed dataset considerations, and rigorous model evaluation, laying the groundwork for insightful and reliable findings. Ongoing research and refinement of these methodologies contribute to the advancement of our understanding of social media dynamics during pandemics.

Here's an example of how an LSTM model might work for sentiment classification:

The first step is to preprocess the text data by tokenizing it into individual words and converting each word to a numerical representation, such as a one-hot vector or a word embedding.In an LSTM model, the input sequence is typically converted to numerical representations using one of the following methods:

One-Hot Encoding

This method represents each word in the input sequence as a one-hot vector, which is a vector with the same length as the vocabulary size, with a single 1 at the index of the word in the vocabulary and 0s everywhere else. This method is simple to implement, but it can be memory-intensive, especially for large vocabularies.

Word Embeddings

Word embedding is a technique that represents words as vectors to capture their semantic meaning. The word embeddings are typically learned from a large corpus of text data during a pre-training phase.

The main advantage of this method is that it can handle out of vocabulary words and it can capture the semantic and syntactic relationship between words. Forget gate decides what to remove. The input gate decides what to add. The output gate decides what to output. These gates are governed by the values of the input data and the hidden state from the preceding time step, as well as learning parameters within the model. As the LSTM processes the input sequence, it updates the hidden state with the new information, which allows it toretain relevant information from previous time steps.

Once the LSTM has processed the entire input sequence, it outputs a prediction of the sentiment of the text. This prediction is typically made by passing the final hidden state through a fully linked layer to build a probability distribution over the possible sentiment classes.

The model is trained using labelled examples of text and their corresponding sentiment. In order to reduce the discrepancy between the true sentiment and the predicted sentiment a, the parameters of the model are adjusted during training.

For example, let's say we want to classify the sentiment of the following movie review: "The movie was terrible. The acting was wooden, and the plot was predictable."

- The text is preprocessed by tokenizing it into individual words: "The", "movie", "was", "terrible", "The", "acting", "was", "wooden", "and" "the", "plot", "was", "predictable".
- The LSTM model takes in the sequence of numerical representations of the words as input and processes it through the memory cells.
- As it processes the input sequence, the LSTM updates the hidden state with information about the negative sentiment expressed in the text, such as the words "terrible" and "wooden".
- Once it has processed the entire input sequence, the LSTM outputs a prediction that the sentiment of the text is negative.

In an LSTM (Long Short-Term Memory) model(Shown in Figures 3 and 4), the memory cell is the core component that enables the network to retain information over a long period of time.

The memory of the LSTM at a specific time step is represented by a vector called the cell state. The LSTM can store information for a very long time because it is transmitted from one-time step to the next. Every state of the cell is updated using the prior time step, taking into account the input gate, the cell state, and the forget gate. Input gates are responsible for directing the flow of information into memory cells. Using the current input and the prior hidden state, the sigmoid function generates a value between 0 and 1, which is used to determine the amount of current input.

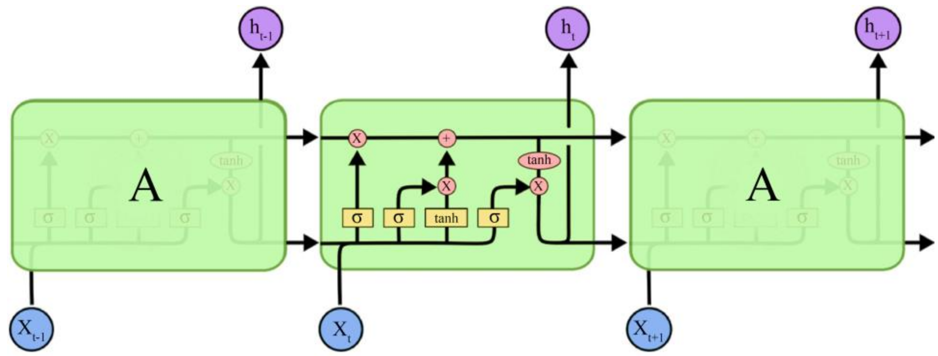

Fig. 3 LSTM chain of memory units

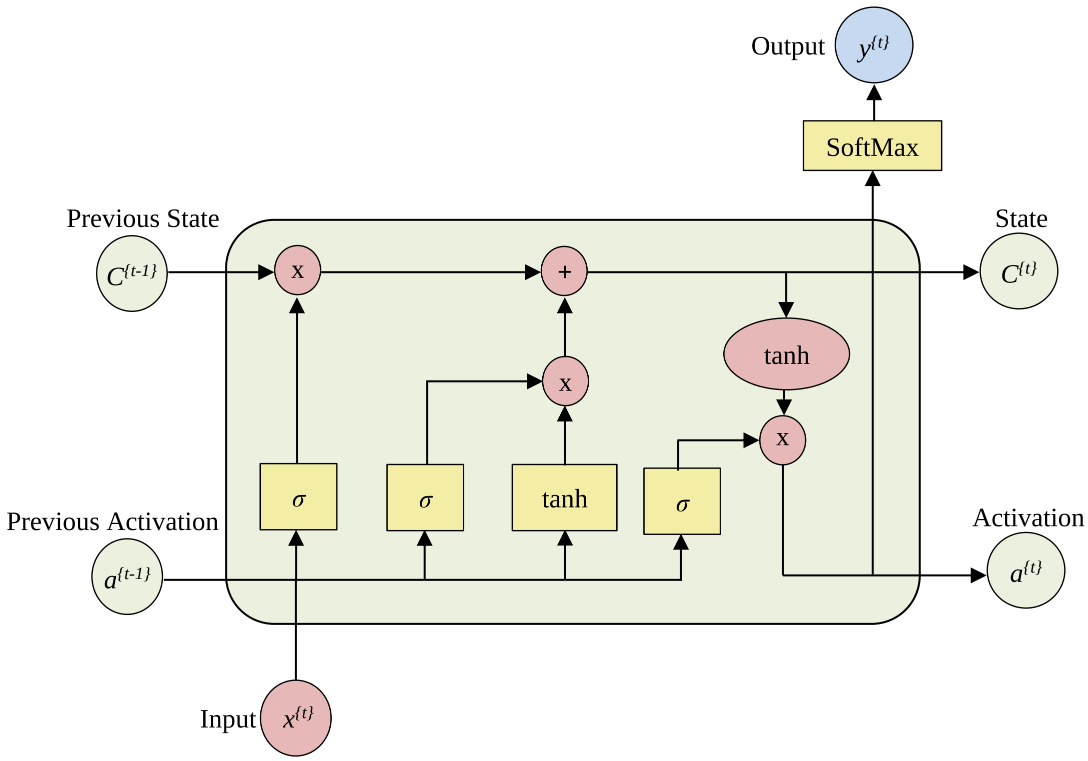

Fig. 4 Single unit of LSTM cell

The forget gate is used to control how much of the previous state is passed on to the next state. This is important because it allows the LSTM cell to learn long-term dependencies. For example, if the LSTM cell is trying to learn the sequence "the cat sat on the mat", it needs to be able to remember that the word "the" was at the beginning of the sequence, even though it is several words later. The forget gate allows the LSTM cell to do this by selectively forgetting parts of the previous state that are no longer relevant.The forget gate, input gate, and output gate work together to allow the LSTM cell to learn long-term dependencies and make accurate predictions. In an LSTM model for sentiment analysis, the fully connected layer does not inherently know the sentiment of an input. It is a simple neural network layer that applies a linear transformation to the data.

The LSTM learns to make predictions during the training phase by minimizing the difference between the predicted sentiment and the true sentiment. To reduce the errors that occurred between expected sentiment and real sentiment model performance parameters to be adjusted during the training phase. This process is done through backpropagation, where the error is propagated back through the network. Once the model is trained, it can make predictions on new, unseen data by passing the input text through the LSTM and the fully connected layer. The LSTM uses the information stored in the hidden state, and the final hidden state is passed through the fully connected layer to produce a probability distribution over the sentiment classes. The class with the greatest probability is chosen as the overall prediction for the sentiment of the input text.

Table 1. Performance metrics for the economy dataset

	Precision	Recall	F1-score	Accuracy
Logistic Regression	0.53	0.78	0.55	0.71
LSTM	0.7	0.71	0.71	0.76
BERT	0.82	0.83	0.85	0.85

Table 2. Performance metrics for the WFH dataset

	Precision	Recall	F1-score	Accuracy
Logistic Regression	0.47	0.79	0.46	0.66
LSTM	0.58	0.64	0.6	0.71
BERT	0.8	0.58	0.76	0.78

Table 3. Performance metrics for online learning

	Precision	Recall	F1-score	Accuracy
Logistic Regression	0.78	0.83	0.8	0.84
LSTM	0.78	0.82	0.82	0.81
BERT	0.91	0.91	0.91	0.92

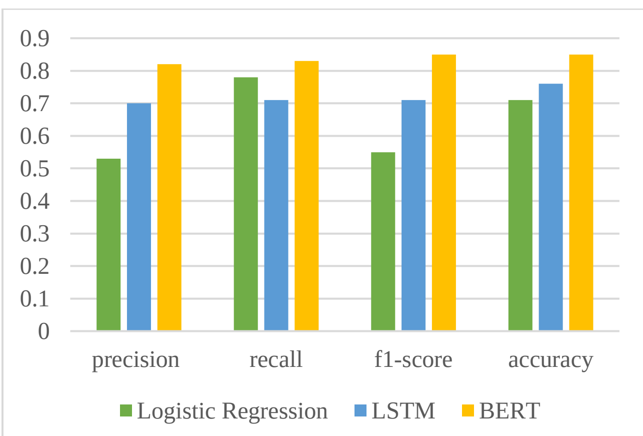

Fig. 5 Comparison of performance metrics for the economy dataset

Fig. 6 Comparison of performance metrics for the WFH dataset

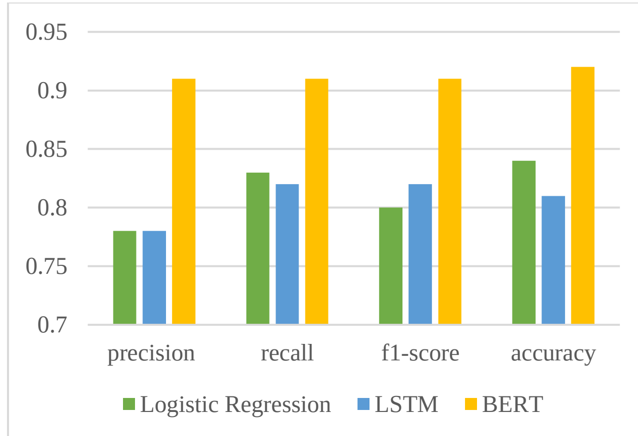

Fig. 7 Comparison of performance metrics for online learning

4. Results and Discussion

To set up and perform this experimentation on different datasets, a few essential tools and requirements are needed. Libraries and frameworks used are TensorFlow, PyTorch, Pandas, NumPy, Sci-kit learn, Transformers library, NLTK, and regex. A system with a minimum of 4 CPU cores is required to train the models. A major requirement to run any of the algorithms explained is that the system needs a high-performance GPU. GPUs help speed up the training process as compared to generic CPUs. Figures 5 to Figure 7 show the comparison of three algorithms, Logistic regression, LSTM and BERT, on three datasets gathered from the social media platform Twitter. The sizes of datasets are different from one another, the Online learning dataset being the largest with 162054 instances of tweets and the Economy dataset being the smallest with 4769 tweets. These datasets are pre-labelled with their corresponding sentiments using VADER. These pre-labelled datasets are segregated into training and testing. It is very clear and evident that Logistic regression, which is a statistical machine learning algorithm, does not perform well in any aspect of three of the datasets used. This shows the need for the use of deep learning algorithms such as LSTM and BERT.The process of training and testing is performed on LSTM and BERT, whose performance metrics are shown in tables 11 through 13 and comparison charts (Figures 5 to 7). It is obvious that LSTM and BERT outperform Logistic regression in every performance metric. Between LSTM and BERT, the latter has even better accuracy on all three datasets used.

5. Conclusion

The COVID-19 outbreak has caused a significant change in various aspects, such as Working from the office, which transitioned to working from home, and traditional classroom learning, which was converted to Online mode, and it had a considerable impact on economies of various countries. People took their opinions on these topics to microblogging sites such as Twitter, which became breeding grounds for varied viewpoints from people with different backgrounds.

Through the examination of tweets about these subjects, this study sought to ascertain how the general public felt about the categories of work-from-home (WFH), economy, and online learning during the COVID-19 pandemic. We also proposed a sentiment analysis model that leveraged the use of BERT (Bidirectional Encoder Representations from Transformers) and Long Short-Term Memory (LSTM) architectures. It was observed that while both LSTM and BERT performed well in sentiment analysis, BERT was able to capture more contextual information. The results of the experiment indicated that the BERT model outperformed the LSTM model in terms of accuracy, thereby highlighting the superior performance of BERT in this specific task of sentiment analysis.

Furthermore, the results of the analysis revealed that people's lives and perceptions have been significantly impacted by the COVID-19 pandemic, specifically in categories such as WFH, economy, and online learning, with a large proportion of tweets expressing concern and uncertainty about the effects of the pandemic on these aspects. The findings of the project have implications for organizations and policymakers in understanding the public sentiment towards the ongoing pandemic and its impact on various aspects of people's lives. The proposed model can be applied to other real-time public opinion analysis tasks and can aid in decision-making processes. The results also indicate that BERT is a superior choice to LSTM in this specific task of sentiment analysis on twitter data.

References

[1] Yuxing Qi, and Zahratu Shabrina, "Sentiment Analysis using Twitter Data: A Comparative Application of Lexicon-and Machine-Learning-Based Approach," *Social Network Analysis and Mining*, vol. 13, no. 1, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Kazi Nabiul Alam et al., "Deep Learning-Based Sentiment Analysis of COVID-19 Vaccination Responses from Twitter Data," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1-15, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Mohammed Hasan Ali Al-Abyadh et al., "Deep Sentiment Analysis of Twitter Data Using a Hybrid Ghost Convolution Neural Network Model," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Burhan Bilen, and Fahrettin Horasan, "LSTM Network Based Sentiment Analysis for Customer Reviews," *Journal of Polytechnic*, vol. 25, no. 3, pp. 959-966, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Rahul Kumar Chaurasiya et al., "Prediction of Twitter Sentiment Analysis by Using Naive Bayes Algorithm," *Journal of Big Data Technology and Business Analytics*, vol. 2, no. 1, pp. 15-22, 2023. [Google Scholar] [Publisher Link]

[6] Shrawan Kumar Trivedi, and Amrinder Singh, "Twitter Sentiment Analysis of App Based Online Food Delivery Companies," *Global Knowledge, Memory and Communication*, vol. 70, no. 8/9, pp. 891-910, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Dharmendra Dangi, Dheeraj K. Dixit, and Amit Bhagat, "Sentiment Analysis of COVID-19 Social Media Data through Machine Learning," *Multimedia Tools and Applications*, vol. 81, no. 29, pp. 42261-42283, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Syafrida Hafni Sahir et al., "Online Learning Sentiment Analysis during the Covid-19 Indonesia Pandemic Using Twitter Data," *IOP Conference Series: Materials Science and Engineering*, vol. 1156, no. 1, pp. 1-7, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Ashna Bali et al., "Consumer's Sentiment Analysis of Popular Phone Brands and Operating System Preference," *International Journal of Computer Applications*, vol. 155, no. 4, pp. 15-19, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[10] Hernán Gil-Ramírez, and Rosa María Guilleumas-García, "Methodology for Sentiment Analysis in Twitter Posts About Mobile Learning," *South Florida Journal of Development*, vol. 2, no. 2, pp. 2718-2728, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Anika Bhardwaj, and N. Natarajan, "Covid-19 Data Analysis using Machine Learning," *3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, Greater Noida, India, pp. 2096-2099, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Langgao Cheng, "Urban Covid-19 Analysis and Prediction Based on Machine Learning and LSTM Models," *ISCTT 2022; 7th International Conference on Information Science, Computer Technology and Transportation*, Xishuangbanna, China, pp. 1-4, 2022. [Google Scholar] [Publisher Link]

[13] Suhaïl Madarbakusl, Abhishek Gobin, and Roopesh Kevin Sungkur, "Using Sentiment Analysis and Machine Learning to Collect the Perception of Online Learning," *Advances in Science and Engineering Technology International Conferences (ASET)*, Dubai, United Arab Emirates, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Hema Krishnan et al., "Machine Learning Based Sentiment Analysis of Coronavirus Disease Related Twitter Data," *2nd International Conference on Secure Cyber Computing and Communications (ICSCCC)*, Jalandhar, India, pp. 459-464, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Yassine Meraihi et al., "Machine Learning-Based Research for COVID-19 Detection, Diagnosis, and Prediction: A Survey," *SN Computer Science*, vol. 3, no. 4, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Humaira Aslam, and Santanu Biswas, "Analysis of COVID-19 Death Cases Using Machine Learning," *SN Computer Science*, vol. 4, no. 4, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]