Computing&AI Connect
ISSN: 3104-4719
2025, Article ID. 2025.0016, Cite as: https://doi.org/10.69709/CAIC.2025.199112

SCIFINITI PUBLISHING

Research Article

OPEN ACCESS

Exploring a Hybrid Deep Learning Framework to Automatically Discover Topic and Sentiment in COVID-19 Tweets

Khandaker Tayef Shahriar[✉,1,2,*] **Iqbal H. Sarker**[✉,3,*]

[1] Department of Computer Science and Engineering, Chittagong University of Engineering & Technology, Chittagong 4349, Bangladesh
[2] Department of Computer Science and Engineering, International Islamic University Chittagong, Chattogram 4318, Bangladesh
[3] Center for Securing Digital Futures, Edith Cowan University, Perth, WA 6027, Australia

Article History
Received: October 25, 2024 Accepted: June 24, 2025 Published: July 14, 2025

Abstract
COVID-19 has created a major public health problem worldwide and other issues such as economic crisis, unemployment, mental distress, etc. The pandemic has affected people not only through infection but also by causing stress, worry, fear, resentment, and even hatred. Twitter is a highly influential social media platform and a major source of health-related information, news, opinions, and public sentiment, with content shared by both individuals and official government sources. Therefore, an effective analysis of COVID-19 tweets is essential for policymakers to make informed decisions. However, identifying relevant and useful content from large volumes of text is challenging when trying to understand public sentiment on COVID-19-related topics. In this paper, we propose a deep learning framework for analyzing topic-based sentiments by extracting key topics with significant labels and classifying positive, negative, or neutral tweets on each topic to quickly find common topics of public opinion and COVID-19-related attitudes. While building our model, we take into account the hybridization of Bidirectional Long Short-Term Memory (BiLSTM) and Gated Recurrent Unit (GRU) structures for sentiment analysis to achieve our goal. The experimental results show that our topic identification method extracts better topic labels, and the sentiment analysis approach using our proposed hybrid deep learning model achieves the highest accuracy compared to traditional models.

Keywords:
COVID-19; deep learning; sentiment analysis; topic identification

1. Introduction

Social networks play a crucial role during major crises, as individuals use these online platforms to share feedback and ideas, generating valuable information and insights related to disaster response [1]. Twitter is regarded as one of the most important social media platforms to explain the behavior and predict the outcomes of the pandemic [2]. At the end of December 2019, a novel coronavirus outbreak that caused COVID-19 was reported [3]. Due to the rapid spread of the virus, the World Health Organization (WHO) declared a state of emergency [4]. The impact of the COVID-19 pandemic on the Twitter platform has been significant. Thus, it is very necessary to know the topics related to the COVID-19 tweets and the associated sentiment polarities users are producing on this platform regularly to get an overview of the pandemic situation, human needs, and steps to reduce the harmful impact of the pandemic. The Twitter platform can be considered a source of useful information to highlight user conversations and better understand public perception towards the COVID-19 pandemic situation. Therefore, by focusing on the context of the COVID-19 pandemic, in this work, we analyze people's tweets on the basis of extracting meaningful topics in an unsupervised manner and predicting sentiment classes in a supervised manner.

Consider a dataset of Twitter that contains user-generated tweets having the sentiment polarities as the

* **Corresponding Author:**
Khandaker Tayef Shahriar, Department of Computer Science and Engineering, Chittagong University of Engineering & Technology, Chittagong 4349, Bangladesh; International Islamic University Chittagong, Chattogram 4318, Bangladesh;
Iqbal H. Sarker, Center for Securing Digital Futures, Edith Cowan University, Perth, WA 6027, Australia

© 2025 Copyright by the Authors.
Licensed as an open access article using a CC BY 4.0 license.

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/CAIC.2025.199112

target label about COVID-19-related issues. It is very difficult to parse all the tweets and express the internal context manually [5]. Therefore, finding the internal topics automatically of all tweets will help policymakers in the relevant departments to implement mandatory measures to reduce the negative impact of the pandemic. However, sentiment analysis is a popular and significant current research paradigm that reflects public perceptions of the event. Sentiment analysis of tweets helps to determine whether a given tweet is neutral, positive, or negative [6]. To achieve our goal, we consider the following research questions:

RQ 1: How to automatically discover the topics associated with the significant labels through analyzing COVID-19 tweets?

RQ 2: How to design an effective framework for predicting the sentiments of tweets with associated topic labels?

To answer these questions, we develop a hybrid deep learning framework by considering the sentiment terms and aspect terms of tweets. However, it is a challenging task to extract meaningful topic labels automatically by machine instead of following the manually annotated topic labeling approach with the diverse human interpretations [7]. Therefore, in this paper, we use Latent Dirichlet Allocation (LDA) [8], which is an unsupervised probabilistic algorithm to discover topics from tweets. Thus, for the topic identification purpose, we do not need any labeled dataset. A collection of topics found in the tweets is generated by LDA. In our earlier paper [9], we proposed "SATLabel" for identifying topic labels, comparing the effectiveness with the manual labeling approach only. In this paper, we expand our previous work by changing the LDA-based optimal topic selection process and evaluating the effectiveness with respect to other topic labeling techniques. We also integrate the sentiment detection approach to build a complete framework for topic-based sentiment analysis. Our experimental results show that the label produced by the approach in the proposed framework has the highest soft cosine similarity score with tweets of the same topic compared to other topic labeling methods. In recent years, deep learning models have become very successful in many fields. Deep learning models automatically extract features from various neural network models and learn from their errors [10]. Deep learning models are extensively used in the field of sentiment analysis by Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), BiLSTM, and GRU [11]. GRU is good at capturing order details and long-distance dependency in a sequence from aspect terms and sentiment terms [12]. On the other hand, BiLSTM can read the input sequences in the forward and the backward direction to increase the amount of information and improve context successfully [13]. To improve the performance of sentiment prediction, we propose a hybrid deep learning model based on the combination of the above GRU and BiLSTM features with the use of the Global Average Pooling layer.

Conventional sentiment analysis methods typically classify raw text directly, whereas our proposed framework incorporates latent topic extraction using LDA as additional relevant features, enabling deeper insights through topic-aware sentiment analysis. Our proposed hybrid model incorporates both long-term contextual patterns (via BiLSTM) and short-term dependencies (via GRU) to improve sentiment classification performance. Our proposed method also provides structural analysis by integrating topic modeling with hybrid deep learning models to identify sentiment variations on pandemic-related themes. The primary contributions of this paper can be summarized as follows:

- We effectively use sentiment terms and aspect terms to identify topic labels in COVID-19 tweets.
- We propose a hybrid deep learning model combining GRU and BiLSTM
- features with the use of the Global Average Pooling layer.
- We have conducted extensive experiments utilizing real-world COVID-19 tweets to show the effectiveness of our proposed framework.

The rest of this paper is organized as follows. Section 2 reviews work on related topics, modeling and sentiment analysis. The methodology and architecture of the proposed framework are depicted in Section 3. After that, the results of the experiments are analyzed in Section 4. Next, we present the discussion and the conclusion section of our work that provide directions for future work.

2. Related Work

The suffering caused by COVID-19 is severe [14]. COVID-19 news has been propagated widely on social media, surpassing other news. The Twitter remains on top to broadcast pandemic-related news [15]. Tweets related to COVID-19 can help identify meaningful topic labels to highlight user conversations and understand ideas about people's needs and interests. Many researchers use the LDA algorithm to extract hidden themes from texts. Hingmire et al. [16] proposed a paper to create an LDA-based topic model, but expert annotation is needed to assign the topic to class labels. Wang et al. [17] presented a paper that minimizes the problem of data sparsity without labeling the main topics particularly. Hourani et al. [18] pre-

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/CAIC.2025.199112

sented a technique of text classification according to their topics that required a labeled dataset. Asmussen et al. [19] presented a paper where labeling the topic depends on the researcher's point of view, without having any machine-dependent automatic method. Maier et al. [20] introduced the applicability and accessibility of communication researchers using LDA LDA-based topic labeling system based on extensive knowledge of the context. Guo et al. [21] compared LDA analysis with dictionary-based analysis by implementing a manual topic labeling technique. Ordun et al. [22] labeled each LDA-generated topic with the first three terms having the highest associated probabilities in the corresponding single document of each topic to interpret. Since the Mallet wrapper of the LDA algorithm tends to offer a better division of topics than the inbuilt Gensim version, the LDA Mallet models were implemented in the study [23] to generate topics where each topic is an integration of keywords and each keyword provides a certain weighted percentage to the topic and the topic labels were inferred from keywords.

Twitter sentiment analysis has become a major source of interest for over a decade [24]. In history, many works were performed in the field of sentiment analysis using deep learning models [11,25]. Behera et al. [26] presented a hybrid model called Co-LSTM, which highly fits with big social data. Poria et al. [27] considered different aspects of analysis in order to investigate multimodal sentiment analysis, implementing three deep learning-based structures. Kaur et al. [28] examined emotions and sentiments using TextBlob and IBM Tone analyzer by translating 16,138 tweets into English. Nemes and Kiss [29] scrutinized tweets using TextBlob and RNN. Jiang et al. [30] proposed a paper based on aspect-level sentiment conversion modules by extracting aspect-specific sentiment words. Imran et al. [31] proposed a multi-layer LSTM-based model for predicting both emotions and sentiment polarities. Authors in [32] explored sentiments of Indians in COVID-19-related tweets after the lockdown. For analysis purposes, they collected 24,000 tweets from March 25th to March 28th, 2020, using the keywords: #IndiaLockdown and #IndiafightsCorona. The results showed that although there were negative sentiments about the closure, tweets possessing positive sentiments were actually present. Rustam et al. [33] created the feature set by combining the term frequency-inverse document frequency and bag-of-words for performance comparison of different machine learning techniques by classifying Twitter data as positive, neutral, or negative sentiment. They investigated the dataset using the LSTM deep learning model. Naseem et al. [34] compared conventional machine learning techniques and deep learning methods using the application of various word embeddings such as Glove [35], fastText [36], and Word2Vec [37] by categorizing COVID-19-related tweets into negative, positive, and neutral classes. Their results suggested that deep learning (DL) methods performed better than traditional ML techniques.

Baired et al. [38] used SenticNet for fine-tuning various features from the dataset. They investigated short microblogging (e.g., Twitter), messaging, and sentiment analysis by combining knowledge-based and statistical methods. A study [39] suggested a way to classify sentiments using a deep learning model. It used NLP to model the topic to identify key issues related to COVID-19 expressed on social media. The classification was performed by applying the LSTM Recurrent Neural Networks (LSTM RNN) model. Ahmed et al. [40] used the opinion mining approach to identify sets of users with similar attitudes. At the beginning of the COVID-19 pandemic, Xue et al. [41] performed sentiment analysis on Twitter data, expressing that fear of the unknown environment was prevalent on Twitter. Medford et al. [42] explored sentiments and investigated topics using LDA by collecting the Twitter data from January 14th to 28th, 2020. Xiang et al. [43] collected 82,893 tweets for performing sentiment analysis and topic modeling by applying the NRC Lexicon and LDA, respectively. Satu et al. [44] presented "TClustVID," which performs a cluster-based division of tweets and topic modeling. They produced themes of positive, negative, and neutral topics for the cluster investigation and the realization of the COVID-19 pandemic status. Chandrasekaran et al. [45] analyzed key topics and corresponding sentiments by identifying the patterns and trends of COVID-19-related tweets before and after the outbreak. They used LDA to extract 26 topics from the COVID-19-related tweets and also clustered them into 10 broad themes to analyze the effect of COVID-19 on various domains. By analyzing the above works, we can conclude that most of the works do not consider a complete framework for the automatic topic identification and sentiment analysis of COVID-19-related tweets. Moreover, most of the works followed a way of labeling topics manually, which is expensive, time-consuming, and dictates difficult human interpretations of the tweets. However, a topic-based analysis of sentiments of the COVID-19 tweets categorizes the predicted multi-class sentiment polarities according to the extracted topics, to gain an overview of broader public opinion. Therefore, in this paper, we propose a framework for automatically identifying key topics associated with specific topic labels and predicting sentiment polarities to present a topic-based sentiment analysis of the COVID-19-related tweets.

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/CAIC.2025.199112

3. Methodology

Topic identification and sentiment analysis (also called opinion mining) play an important role in obtaining information from social media platforms. In this paper, we implement two methods for in-depth analysis of the COVID-19-related tweets on Twitter. The first one is the LDA-based automatic topic modeling and labeling approach, and the second one is multiclass sentiment analysis based on a hybrid deep learning model. Our proposed framework aims to illustrate public sentiments in tweets towards some specific topic labels. In this section, we present our proposed framework for conducting a topic-based sentiment analysis of COVID-19-related tweets, as shown in Figure 1. In Figure 1, the "topic identification" module generates a topic label, and the "deep learning model" simultaneously predicts the sentiment polarity of each tweet. Therefore, we can perform topic-based sentiment analysis to classify the sentiment polarities of all tweets into corresponding topic labels to provide insight into the COVID-19 pandemic situation. Text pre-processing is one of the most important steps to analyze and extract features from textual data for further processing. After performing normalization of noisy, non-grammatical, and unstructured tweets, we follow a set of processing steps to produce the expected result. By following the topic identification step, our framework separates positive, neutral, or negative sentiments from COVID-19 tweets by implementing a GRU-BiLSTM-based hybrid deep learning model. The final summary generated by the proposed framework provides decision support to the policymakers based on the topic-based sentiment analysis, which can be presented in any format, either text or charts.

3.1. Preparing the Dataset

TextBlob, VADER, and other sentiment analysis tools are often used to detect sentiments from COVID-19-related tweets [29,30,33,40]. However, in this work, we only consider the publicly available manually tagged dataset with sentiment polarities of the COVID-19-related tweets from the Kaggle repository[1] (accessed on 04 July 2025). The dataset consists of two CSV files: one named Corona_NLP _train.csv and the other Corona_NLP_test.csv. There are 41,157 tweets in the Corona NLP train.CSV file and 3798 tweets in the Corona NLP test.CSV file. The five sentiment categories—Positive, Extremely Positive, Neutral, Negative, and Extremely Negative—are consolidated into three classes: Positive, Neutral, and Negative. This transformation serves as the target label to enhance classification accuracy. There are 7713 Neutral, 15,398 Negative, and 18,046 Positive tweets available in the dataset, which are marked with 0, 1, and 2, respectively. From the training dataset, we have kept 20% of the data as a validation set. We prepare the Twitter dataset to get the standard form by using several text processing functions.

(1) *Conversion of words into lowercase:* We transform all words in the tweets into lowercase.
(2) *Deletion of URLs and links:* We replace all the URLs and hyperlinks in the tweets with empty strings.
(3) *Elimination of mentions:* Normally, people use @ in tweets to mention users. In our work, we eliminate the mentions.
(4) *Removal of hashtags:* We identify hashtags and remove them with empty strings because hashtag words do not carry useful meaning in sentiment analysis [46].
(5) *Handling contractions:* Generally, people use common English contractions in social media. We apply the regular expression to handle common contractions. For example, we replace won't with would not.
(6) *Elimination of punctuation:* There is no importance of punctuation in tweets. We withdraw punctuation symbols such as &, -, etc.
(7) *Managing words less than two characters:* We remove the words whose length is less than two characters because words such as 'I' do not carry much influence in determining sentiment.
(8) *Stop words elimination:* Stop words in the English language, such as pronouns, articles, prepositions, etc., have no emotional impact on the sentiment classification process. We remove the stop words.
(9) *Dealing with Unicode and Non-English words:* To develop a clean and noise-free dataset, we remove the tweets that are not in English, and Unicode like "\u018e".
(10) *Tokenization:* We vectorize the corpus by transforming each text into a sequence of integers based on word count.

3.2. Extracting Sentiment Terms and Aspect Terms

Sentiment terms capture the tone or perception of a sentence. Usually, adjectives and verbs are regarded as sentiment terms in a sentence that reflects the expressed opinion of the text. Noun phrases and nouns are regarded as aspect terms. Aspect terms are often considered as features that describe the event, product, or entity [47]. We follow the precise parts of speech tagging, which is an effective way to extract sentiment terms and aspect terms from texts. Examples of sentiment terms and aspect terms are shown in Table 1.

[1] https://www.kaggle.com/datatattle/covid-19-nlp-text-classification

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/CAIC.2025.199112

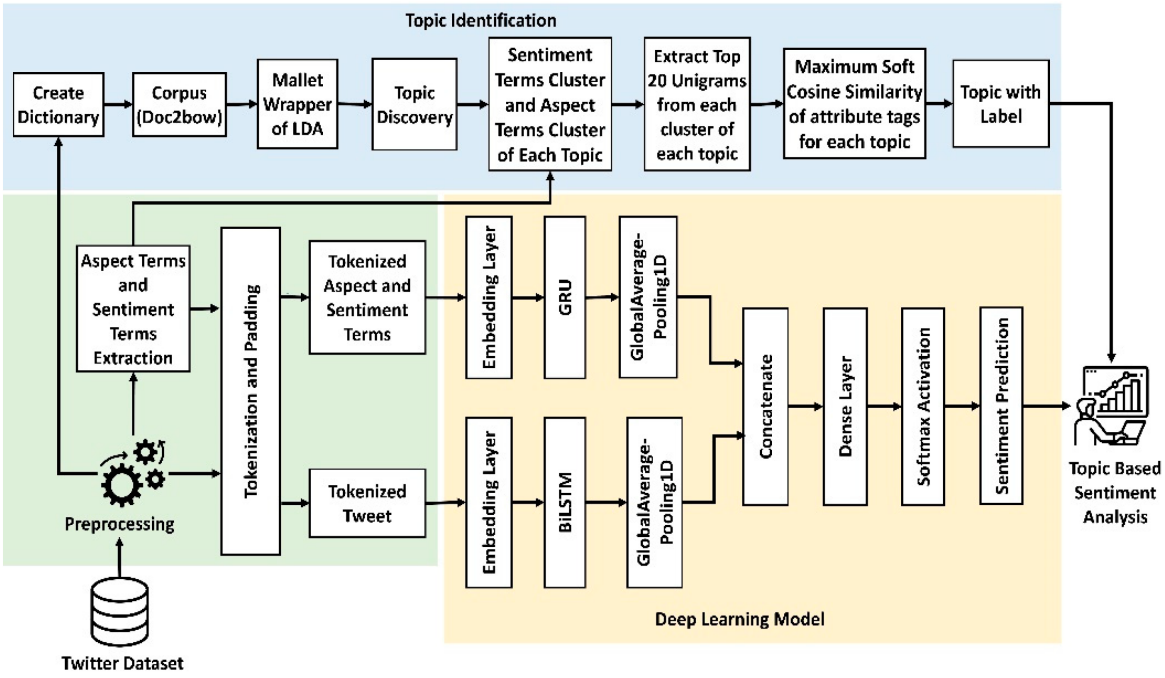


Figure 1: Architecture of the proposed framework.

Table 1: Examples of sentiment terms and aspect terms.

Sample Tweet	Sentiment Terms	Aspect Terms
Please read the thread	read	thread
To enjoy and relax with your dinner It is a great place	enjoy, relax, great	dinner, place
Links with info on communicating with children regarding COVID-19	communicate, covid	links, info, children
The retail store owners right now	retail, right	owners, store

3.3. Feature Extraction

To implement the machine learning and deep learning models, feature extraction is essential from the tweets [48]. Two methods of extracting features, a bag of words (BoW) and Word2vec, are utilized in the proposed framework.

The BoW is an easy way to extract features, and it is used to retrieve information and perform NLP tasks [49]. The Bag of Words (BoW) model generates features for training machine learning models based on the presence of individual words in the text and is widely used in text classification tasks. BoW produces vocabulary for all unique words and their frequency of occurrence in the documents to train models. In our work, we implement Doc2bow to extract numerical features from tweets to train the LDA model, which is a simple technique in Gensim [50] that converts documents into a representation of the bag of words, as shown in the topic identification section of Figure 1. In this paper, we use Word2vec (skip-gram version) as embedding layers to extract numerical features from tokenized texts and develop our framework for sentiment analysis purposes [37,51].

3.4. Topic Identification

LDA is a well-performed topic modeling algorithm for finding hidden themes available in the corpus on an unlabeled dataset. However, the challenge is how to assign important labels to topics generated by the LDA. The working principle for generating automatic topic labels from the Twitter dataset is presented in Algorithm 1. The steps for the identification of topics and the generation of significant topic labels automatically as output in the proposed framework are discussed below:

Algorithm 1: Automatic topic labeling

Input: T: Number of Tweets in the dataset

Output: Topic Label (T_{Label})

// Preprocess Tweets
for ***each*** $t \in \{1, 2, \ldots\ldots, T\}$ **do**
  $T_p \leftarrow Preprocess(t)$
// Initialize Corpus
Initialize an empty list *Corpus*
// Corpus Development, Sentiment, and Aspect Terms Extraction
for ***each*** $t_p \in \{1, 2, \ldots\ldots, T_p\}$ **do**
  Append dictionary representation of Doc2Bow(t_p) to *Corpus*
  $S_{T_p} \leftarrow$ *Sentiment Terms* (t_p) -
  $A_{T_p} \leftarrow$ *Aspect Terms* (t_p) -
// Topic Discovery using LDA
$K \sim Mallet\ (LDA(Corpus))$
// Cluster Creation Based on Dominant Topics
for ***each*** $k \in \{1, 2, \ldots\ldots, K\}$ **do**
  for ***each*** $t_p \in \{1, 2, \ldots, T_p\}$ **do**
    $k_{dominant,t_p} \sim dominant\ topic\ (t_p, k)$
    // Create Clusters from Topics
    $C_S \sim Cluster\ (S_{T_p} \rightarrow k_{dominant,t_p})$
    $C_A \sim Cluster\ (A_{T_p} \rightarrow k_{dominant,t_p})$
// Extract Unigrams and Generate Topic Labels
for ***each*** $k \in \{1, 2, \ldots, K\}$ **do**
  // Extract Top 20 Unigrams from Sentiment Terms Cluster
  $U_S \sim max\ count\ (Top_Unigrams\ (C_S \rightarrow k))$
  // Extract Top 20 Unigrams from Aspect Terms Cluster
  $U_A \sim max\ count\ (Top_Unigrams\ (C_A \rightarrow k))$
  // Create Attribute Tags
  $A_{tag} \leftarrow U_S + U_A$
  // Assign Topic Label
  $T_{Label} \leftarrow max\ soft\ cosine\ similarity\ (A_{tag}, k)$

3.4.1. Creating Dictionary and Corpus

The systematic generation of multiple language lexicons is supported by a dictionary, while the corpus typically illustrates an arbitrary sample of the language. The corpus of a document is made up of words or phrases. In the Natural Language Processing (NLP) paradigm, the language corpus plays an important role in creating a knowledge-based system and text mining. In the proposed framework, we construct a dictionary from the preprocessed text and then create a corpus. Documents in the dictionary are converted to the Bag of Words (BoW) format using Doc2bow embedding [50] for the corpus development. Corpus comprises the word id and its frequency in all documents. Each word is considered a normalized and tokenized string.

3.4.2. Topic Discovery

The BoW corpus is transferred to the LDA Mallet wrapper. The presence of a set of topics in the corpus is discovered by the LDA. LDA Mallet wrapper works fast and provides a better classification of topics using the Gibbs Sampling [52] method. LDA produces the most prominent words in the themes. So, by using the weightage of keywords, one can infer prevalent topics in texts. To overcome the difficult manual labeling method, our framework produces automated topic labels using sentiment terms and aspect terms without human interpretation. Based on the result of the topic coherence score, we select an LDA model that gets a total of 14 topics itself. Then we get the main topic as the dominant topic in each tweet from the LDA model to identify the distribution of topics across all tweets in the dataset.

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/CAIC.2025.199112

3.4.3. Labeling Topic Using Top Unigrams

We generate sentiment terms clusters and aspect terms clusters by experimenting with tweets corresponding to each LDA-generated topic. Therefore, we find 14 sentiment terms cluster and 14 aspect terms cluster from LDA-generated topics. Unigram is a single n-gram word sequence. The use of unigrams can be seen in NLP, cryptography, and mathematical analysis. However, the soft cosine similarity takes into account the similarity of the features in the vector space model [53]. For each topic, we select the top 20 unigrams as the threshold value because of high frequencies of these unigrams and extract the top 20 unigrams from the sentiment terms cluster and the top 20 unigrams from the aspect terms cluster. Then we create all the possible combinations of the top 20 unigrams of sentiment terms with the top 20 unigrams of aspect terms for each topic, respectively, to generate an appropriate attribute tag. We choose an attribute tag that has the highest soft cosine similarity value concerning the tweets of that topic to assign a significant topic label. We use the sentiment term and aspect term blended attribute tag to label each topic because this feature tag describes the topic of the tweets. To classify a tweet with a specific topic label from test data, we detect the topic number that has a significant percentage impact on that tweet.

3.5. Building a Hybrid Deep Learning Model for Sentiment Analysis

In our proposed framework, we implement a hybrid deep learning model for multiclass sentiment classification using two powerful feature extractors. Gated Recurrent Unit (GRU) and Bidirectional Long Short Term Memory (BiLSTM) are used for feature extraction in two branches of the hybrid model in the proposed framework, as shown in Figure 1. Both extractors are followed by individual Global Average Pooling layers. In the embedding layers, we apply word embeddings generated using the Word2Vec skip-gram model to capture the semantic relationships between word pairs. The GRU branch takes tokenized sentiment terms and aspect terms as input, and the BiLSTM branch takes the tokenized tweet as input. The concatenation layer merges the features coming from two branches, followed by the fully connected dense layer with a SoftMax activation function for multiclass sentiment classification, i.e., positive, neutral, or negative, as shown in Figure 1. Below, we describe the main models used to build the GRU-BiLSTM-based hybrid model in the proposed framework:

Long Short Term Memory (LSTM) is a kind of RNN architecture that handles the vanishing gradient problem using exclusive units [54]. Data is stored for a long time in the memory cell of the LSTM unit, and three gates control the access to information inside and outside the cell. For instance, which information from the previous state cell will be memorized and which information will be deleted is determined by “Forget Gate”, whereas which information should enter the cell state is determined by “Input Gate”, and the output is controlled by the “Output Gate”. Bidirectional LSTM, more commonly known as BiLSTM, is a standard LSTM extension that enhances the model performance in the order of sequence [55]. It is a type of sequence processing model combining two LSTMs: one receives the input approaching in the forward direction, and the other receives it in the backward direction. In natural language processing activities, Bidirectional LSTM is considered a popular option to carry on.

Gated Recurrent Unit (GRU) is another kind of recurrent network where information flow is managed and controlled between cells by implementing gating techniques in the neural network. GRU contains an update gate and a reset gate without having an output gate. The GRU structure makes it possible to take on the dependence of large sequences of data adaptively, without losing any data from earlier segments of the sequence. Moreover, GRU is a relatively simple type and much faster to calculate, which often provides comparable functionality. The performance of the GRUs is better on certain smaller and less frequent datasets.

Global Average Pooling (GAP) does not slide in the structure of a small window but is measured over the entire output feature map in the previous layer. GAP typically regularizes the whole network structure, and each output channel corresponds to the features of each class, making the relationship between the output and the feature class more intuitive. Moreover, the GAP layer contains no data parameter. Therefore, using GAP helps to improve network performance and increase cognitive accuracy [56]. In addition, the GAP layer avoids overfitting because it does not need parameter optimization. It serves as a flattened layer for transforming a multi-dimensional feature space into a one-dimensional feature map. In addition, it takes less time for computation. Our hybrid deep learning-based model in the proposed framework effectively utilizes the advantages of BiLSTM, GRU, and GAP models for multi-class sentiment classification. To predict the sentiment polarities, we apply a fully connected dense layer to extract the features from the feature map of the concatenation layer, as shown in Figure 1. The SoftMax classifier acts as its input and takes the output to the final step. Table 2 shows the hyperparameters used in the proposed hybrid model. As summarized in Table 2, the hyperparameter characteristics include word embedding size of

word2vec, number of hidden neurons in BiLSTM, GRU, and batch size. Due to the different lengths of tweets in the dataset, we take the maximum length of the tweets when we input tweets to the model. RMSProp optimizer is used because RMSProp incorporates the best properties to adjust the learning rate adaptively and optimize model parameters [57]. The pseudo-code of the proposed hybrid deep learning model in the framework is provided in Algorithm 2. The model summary for the hybrid deep learning model is shown in Table 3. The shape of the output of the hybrid deep learning model is 3, as seen in the last row of Table 3.

3.6. Topic-Based Sentiment Analysis

After predicting the sentiments of the tweets, the final step of our proposed framework is to classify the sentiment polarities for each of the 14 topics. We visualize and count the number of positive, negative, and neutral tweets to provide a summary to support decision-making for policymakers. Our proposed framework effectively conducts topic-based sentiment analysis to provide a comprehensive overview of the broader public opinion. In the experimental section, we demonstrate the performance and effectiveness of our framework compared to other benchmark models.

Algorithm 2: Pseudo-code of the proposed hybrid model

Input: COVID-19 Training Set T_1, Testing Set T_2
Output: Predicted sentiment classes: s_{pos}, s_{neu}, s_{neg}
// Preprocess the data
Preprocess Tweets
Extract tokenized sentiment terms and aspect terms as tok_{sa}
Extract tokenized tweets as tok_{tweets}
// Train Word2Vec model
Initialize hyperparameters for the Word2Vec model
Train the Word2Vec embedding model using T_1
// Build the hybrid model
Initialize hyperparameters for GRU and BiLSTM layers
// Branch 1 for tokenized sentiment terms
Create an embedding layer using the trained Word2Vec model for tok_{sa}
Add a GRU layer
Add a GlobalAveragePooling layer
// Branch 2 for tokenized tweets
Create an embedding layer using the trained Word2Vec model for tok_{tweets}
Add a BiLSTM layer
Add a GlobalAveragePooling layer
// Concatenate features from both branches
Concatenate the output features from the two branches
// Add the output layer
Add a dense layer with softmax activation for final classification
// Train the hybrid model
Train the hybrid model using T_1 to compute the initial weights
for *epoch = 1 to max epochs* **do**
  // Select a mini-batch for training
  Select a mini-batch from T_1
  // Forward and backward propagation
  Perform forward propagation and compute the loss
  Perform backpropagation and update weights using RMSProp optimization
// Evaluate the model
Use the trained hybrid model to predict sentiment classes for the testing set T_2
Output the predicted sentiment classes: s_{pos}, s_{neu}, s_{neg}

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/CAIC.2025.199112

Table 2: The model hyperparameters.

Hyper-Parameter	Initial Value	Hyper-Parameter Space	Optimal Value
Word embedding size for word2vec	100	50, 100, 150	100
The number of hidden neurons in BiLSTM	128	64, 128, 256, 512	256
The number of hidden neurons in GRU	128	64, 128, 256	128
Batch Size	32	16, 32, 64	32

Table 3: Model Summary for hybrid deep learning model.

Layer (Type)	Output Shape	Param #	Connected to
embedding 1 input (InputLayer)	(None, 23)	0	[]
embedding input (InputLayer)	(None, 37)	0	[]
embedding 1 (Embedding)	(None, 23, 100)	3,565,900	['embedding 1 input[0][0]']
embedding (Embedding)	(None, 37, 100)	3,565,900	['embedding input[0][0]']
gru (GRU)	(None, 23, 128)	88,320	['embedding 1[0][0]']
bidirectional (Bidirectional)	(None, 37, 512)	731,136	['embedding[0][0]']
global average pooling1d 1	(None, 128)	0	['gru[0][0]']
global average pooling1d	(None, 512)	0	['bidirectional[0][0]']
concatenate (Concatenate)	(None, 640)	0	['global average pooling1d 1[0][0]', 'global average pooling1d[0][0]']
dense (Dense)	(None, 3)	1923	['concatenate[0][0]']
activation (Activation)	(None, 3)	0	['dense[0][0]']

4. Experimental Results

In this section, we present the outcomes and findings of our experiments. First, we search trending topics with specific labels. Then we perform sentiment analysis of the tweets using a hybrid deep learning model. We also make a comparison of the results of the proposed framework with the corresponding benchmark models.

4.1. LDA-Based Topic Identification

4.1.1. Finding The Optimal Number of Topics for LDA

We built a function to extract several LDA models with multiple coherence values for the number of topics to obtain the optimal number of topics. LDA produces common co-occurred words and builds them into separate themes. We choose the optimal number of topics on the basis of the Gensim coherence model [41]. We select the LDA model that returns the number of 14 topics for this dataset because it generates the highest coherence value. This is illustrated in Figure 2, which presents the coherence scores for various topic numbers across multiple LDA models.

4.1.2. Selection of Top Unigram Features from Clusters

We generate sentiment and aspect term clusters for each topic. We identified the top 20 unigrams in each cluster based on their frequency. Figures 3 and 4 show the top 20 unigrams from the sentiment terms cluster and aspect terms cluster of the topic no. 3 respectively. Then we get the topic label depending on the highest soft cosine similarity value of the attribute tag about the tweets for that particular topic, which is generated by the combination of top unigrams of sentiment terms and aspect terms.

4.1.3. Qualitative Evaluation of Topic Labels

In Table 4, we present part of a set of tweets assigned by the proposed framework with generated topic labels.

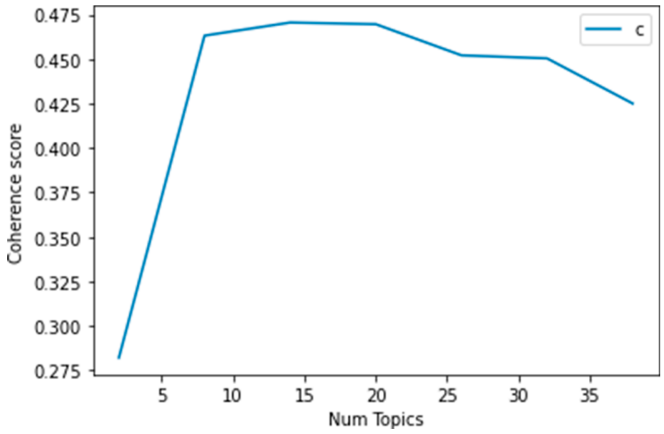


Figure 2: Coherence score for the number of topics.

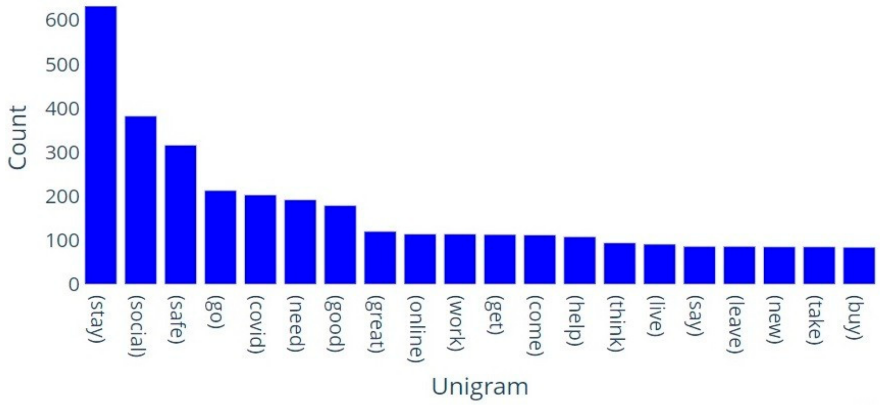


Figure 3: Top 20 unigrams from the sentiment terms cluster of topic no. 3.

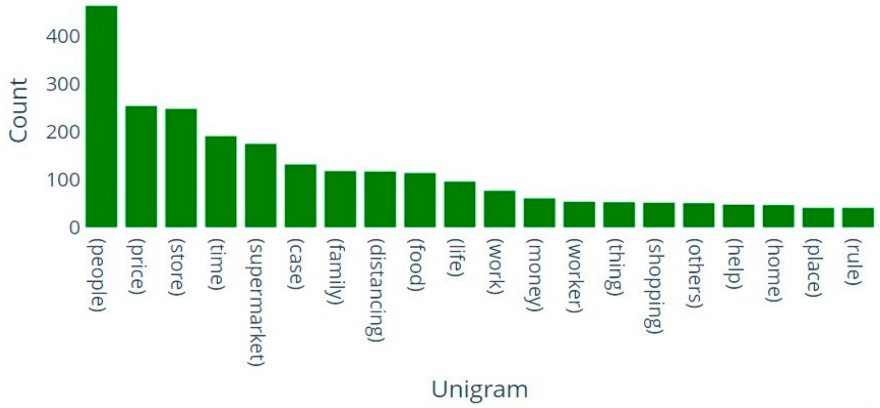


Figure 4: Top 20 unigrams from the aspect terms cluster of topic no. 3.

Table 4 presents that the detected topic labels are well-aligned and closely coherent with the descriptions of the tweets. We can categorize tweets and extract useful information related to a topic by simply separating the tweets using the key label generated by the proposed method for that topic.

4.1.4. Effectiveness Analysis

In this experiment section, we compare the Soft Cosine Similarity (SCS) values of topic labels obtained by our proposed method with the other two approaches, as shown in Figure 5. One approach involves using the top three

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/CAIC.2025.199112

Table 4: Example of topics detected on tweets.

Sample Tweet	Topic No	Topic Label
Create a contact list with the phone numbers of the neighbors' schools The employer chemist GP, set up online shopping accounts If possible adequate supplies of regular meds, but not over-order	5	online slot
Coronavirus Australia: Woolworths to give elderly, disabled dedicated shopping hours amid the COVID-19 outbreak	1	thank store
My food stock is not the only one that is empty... Please, don't panic, there will be enough food for everyone if you do not take more than you need. Stay calm, stay safe	2	hoard food
Dear Coronavirus, I've been following social distancing rules and staying home to prevent the spread of you. However, now I've spent an alarming amount of money shopping online. Where can I submit my expenses to for reimbursement? Let me know. #coronapocolypse #coronavirus	5	safe distancing
The cashier at the grocery store was sharing his insights on # Covid-19, to prove his credibility, he commented, "I'm in Civics class, so I know what I'm talking about"	10	learn behavior
Due to COVID-19, our retail store and classroom in Atlanta will not be open for walk-in business or classes for the next two weeks, beginning Monday, March 16. We will continue to process online and phone orders as normal! Thank you for your understanding!	6	small event
Airline pilots offering to stock supermarket shelves in #NZ lockdown #COVID-19	12	covid supermarket
@TartiiCat Well, new/used Rift S are going for $700.00 on Amazon rn, although the normal market price is usually $400.00. Prices are crazy right now for VR headsets since HL Alex was announced, and it's only been worse with COVID-19	11	crude market
Response to complaint not provided, citing COVID-19 related delays. Yet prompt in rejecting policy before the consumer TAT is over. Way to go?	7	learn scam

significant keywords of each topic for topic labeling [22], while another approach assigns manual topic labels inferred from the keywords of the 14 LDA-generated topics [23], as LDA provides the importance percentage of each keyword within a topic. SCS is used to find semantic text similarities between two documents. The high value of SCS provides a high degree of similarity, and the low SCS value provides less similarity to unrelated documents.

From Figure 5, we find that SCS values generated by the proposed method for all topics are higher than other approaches except topics no. 1, 10, and 12. In the case of these three topics, the approach followed by Ordun et al. [22] finds slightly higher SCS values. For topics no. 9 and 13, we find the same SCS values for our proposed method and the method followed by Ordun et al. [22]. In case of topic no. 3, we find the same SCS value of the method followed by Prabhakaran [23] with our proposed method. From Figure 5, we can see that the topic labels produced by the proposed method in the framework illustrate higher SCS values for a maximum number of topics compared to other labeling methods. Therefore, our proposed framework works better and traces better topic labels from the dataset in an unsupervised manner to reduce the complex workload of human labeling.

4.2. Sentiment Analysis Using a Hybrid Deep Learning Model

4.2.1. Effect of Model Iterations

Multiple iterations of model training significantly influence model performance [58]. The validation accuracy of the model first rises and then falls with increasing model iterations. It can be seen in Figures 6 and 7, when the number of iterations of the model increases, the model gradually overfits, and performance decreases. These figures illustrate the accuracy and loss graph of the proposed hybrid deep learning model in the framework. Based on these plots, we find the best model at epoch no. 4. The fourth epoch is selected as the best epoch for model se-

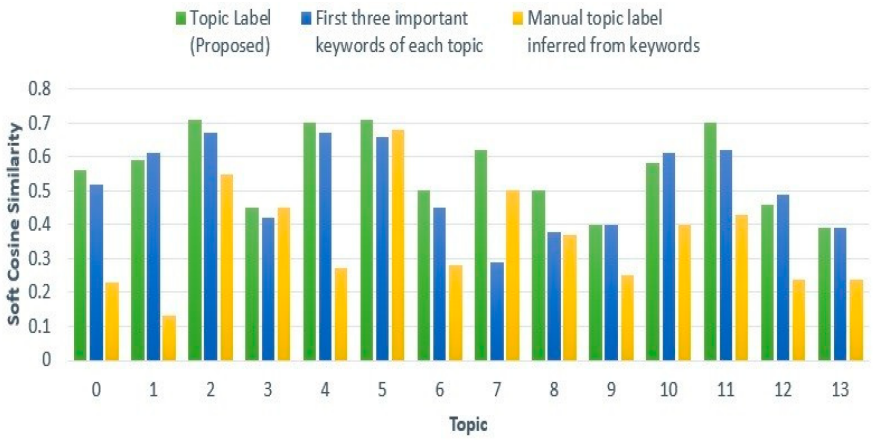


Figure 5: Comparison of the proposed topic label approach with other approaches.

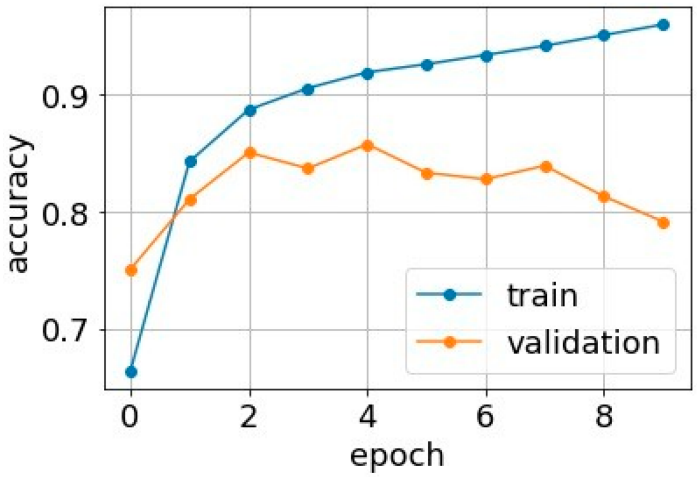


Figure 6: Model accuracy graph (adopted from [58]).

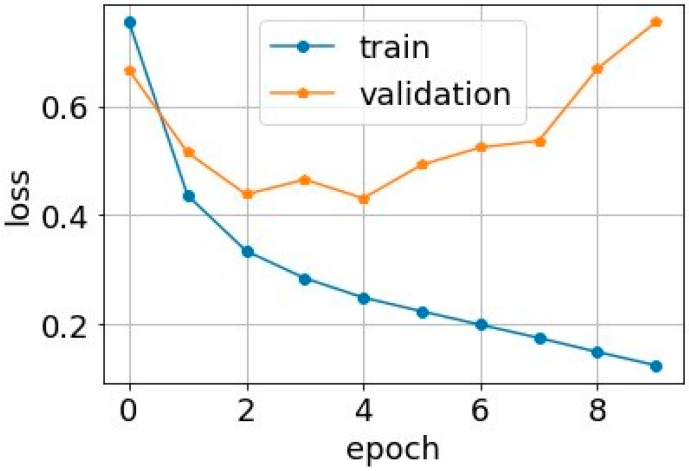


Figure 7: Model loss graph (adopted from [58]).

lection due to the model's performance on validation data without exposure to test data. To mitigate the overfitting problem, we have saved the best model at epoch 4, which

2025, Vol. 2, Article ID. 2025.0017
https://doi.org/10.69709/CAIC.2025.199112

is similar to the early stopping strategy, and then we have loaded the best model for further performance analysis.

4.2.2. Performance Comparison of Different Models

We compare the sentiment analysis method in the proposed framework with the basic deep learning models such as SimpleRNN, CNN, GRU, BiLSTM, and LSTM using standard parameters in the dataset [58]. By implementing the word2vec (skip-gram version), we also train different deep learning models for the dataset and compare their results with the proposed hybrid model in the framework. Figure 8 represents the results of accuracy comparisons, and Table 5 shows the precision, recall, and F1-score comparisons of different models. For simplicity, we have rounded the metrics to two decimal places to gain clarity on the differences in model performance. These differences show meaningful improvements in the context of the proposed hybrid model. We have performed multiple training runs to generate results to ensure robustness. From Figure 8, it can be seen that almost all deep learning models provide good results with an accuracy level of more than 80%. The values reveal that our proposed approach outperforms other models and provides about 86% accuracy. From the results of Table 5, it can be seen that the performance of the proposed hybrid model is higher compared to other basic deep learning models in terms of precision, recall, and F1-score for the weighted average of negative, neutral, and positive sentiment predictions. The experimental results show that the classification performance of the proposed hybrid deep learning classifier is better with the input of aspect terms and sentiment terms in the GRU branch of the model.

4.2.3. Error Analysis

The confusion matrix is used to perform the detailed error analysis as shown in Figure 9. It is clear from the matrix that some data are incorrectly classified. For example, 45 out of 619 cases of the neutral class were predicted positive. In the negative class, 66 of the 1633 data points were mistakenly classified as neutral. Error analysis shows that the positive class achieved the highest rate of accurate classification (87.45%) while the neutral class achieved the lowest (81.42%). A possible reason for the high incorrect predictions in the neutral class may be the presence of a small number of neutral class tweets in the training dataset. Here, we consider natural data distributions to evaluate the performance of hybrid deep learning models, without introducing any data augmentation techniques that may disrupt the natural properties of the dataset and compromise model interpretability. However, sentiment analysis is widely subjective, depends on individual perceptions, and people may consider a tweet in many ways. Therefore, by creating a balanced dataset with a variety of high amounts of data, the incorrect predictions can be reduced to some degree.

5. Discussion

We present topic-based sentiment analysis to extract useful information from COVID-19-related tweets. We plot and count the number of positive, neutral, and negative tweets for each of the fourteen topics as shown in Figure 10. The chart reports that in each topic, most of the tweets are positive. In topics such as COVID-19, supermarkets, crude markets, food hoarding, and public perception, a higher number of tweets were found to express negative sentiment. The topic-based sentiment analysis related to the COVID-19 tweets helps to understand people's perceptions and emotions for taking necessary steps to reduce suffering and compensate for the loss of resources. Government and policymakers can therefore take the necessary steps to understand the details of the human problems and their needs on the basis of public sentiments that reflect tweets on social media in order to minimize the detrimental effects of the COVID-19 pandemic.

It is difficult for the sentiment classifier to infer the correct polarity due to the limited textual context present in tweets. An interpretable topic distribution provided by LDA helps with the larger thematic structure of the data, along with the sequential word dependencies. Thus, domain-specific context changes for crises like COVID-19 are ignored by traditional sentiment models. Our proposed framework ensures alignment of sentiment classifications with thematic diversity in the dataset. Overall, in this paper, we present a data-driven [7] framework for the topic-based sentiment analysis to extract beneficial information from the COVID-19-related Twitter dataset. We use a laptop with a 2.0 GHz Core i3 processor and 4 GB of RAM to execute our proposed technique. We write code in the Python programming language and run our code within the Google Collaboratory platform. We firmly believe that the proposed framework can be used effectively in other areas of application, such as agriculture, health, education, business, cybersecurity, etc. Training a transformer model from scratch, such as BERT, ALBERT, or RoBERTa, requires a substantial dataset and significant computational resources [59]. However, due to infrastructure limitations and the size of our dataset, this approach was not feasible for the present study. However, our objective is to measure the performance of the hybrid configuration of BiLSTM and GRU rather than focusing on the benchmark architecture. Hence, in this paper, we use clas-

Table 5: Evaluation of metrics of different models.

Model	Precision	Recall	F1-Score
Simple RNN	0.82	0.82	0.82
CNN	0.82	0.81	0.81
GRU	0.85	0.85	0.85
BiLSTM	0.85	0.85	0.85
LSTM	0.85	0.85	0.85
Hybrid Model (Proposed)	0.86	0.86	0.86

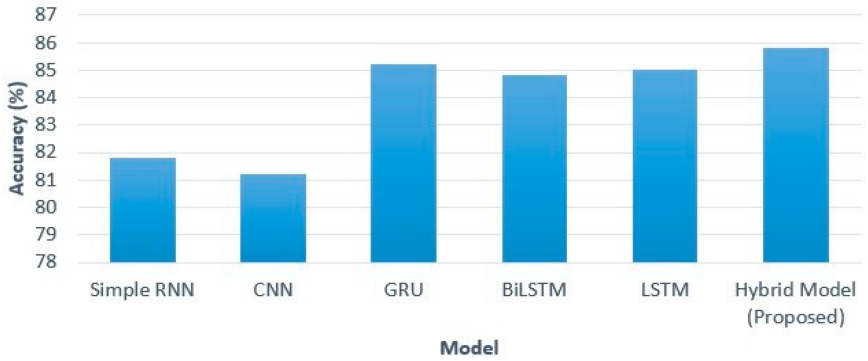


Figure 8: Performance comparison of different models.

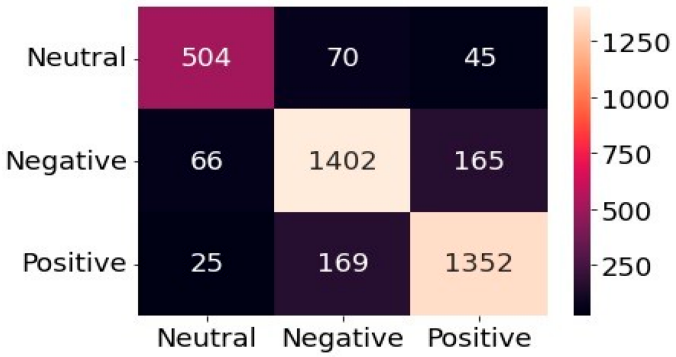


Figure 9: Confusion matrix of the proposed hybrid deep learning model.

sical deep learning classifiers such as BiLSTM and GRU to build the hybrid model [60]. To better understand the decision-making process, in the future, we will incorporate explainable AI (XAI) techniques [48]. While our goal is to develop a methodological framework to balance interpretability and performance, we will analyze the potential and awareness of recent popular methods such as large language models (LLMs) [59] in our future work.

6. Conclusions

In this paper, we have presented a hybrid deep learning framework that effectively and automatically identifies key topic labels and corresponding sentiments from COVID-19 tweets. We have used the popular topic modeling algorithm LDA to extract hidden topics from tweets. Then we have used unigrams of the sentiment terms and aspect terms of the tweets to produce significant and meaningful topic labels on the basis of soft cosine similarity (SCS) values. Our proposed topic labeling method performs better and helps to categorize a huge number of tweets corresponding to semantically similar topic labels to highlight user conversations. In our framework, we utilize the advantages of GRU, BiLSTM, and GAP models explained in the methodology section of this paper. Sequential dependencies are effectively captured by BLSTM

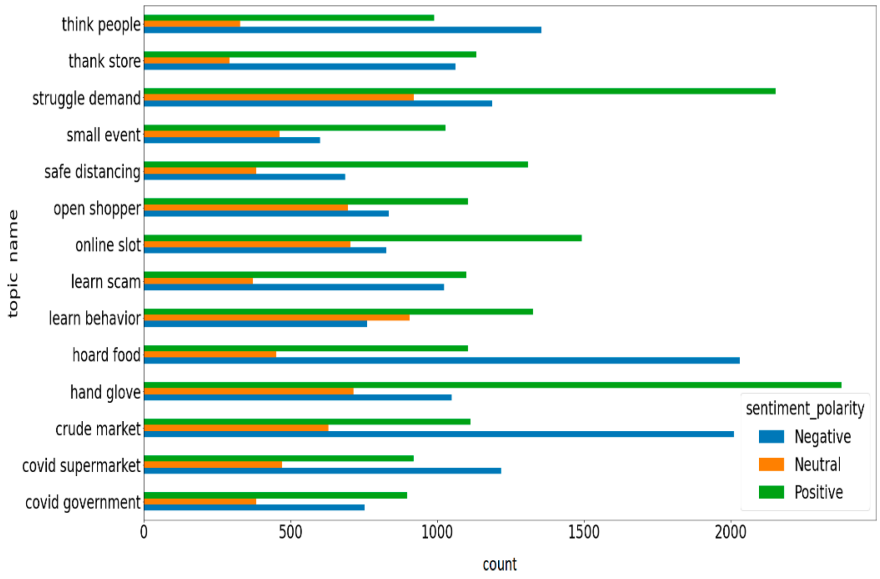


Figure 10: Topic-based sentiment analysis.

and GRU architectures; however, they face challenges in handling ambiguous text without adequate context modeling, which is crucial for interpreting distant contextual information. We have also categorized sentiment polarities towards each topic. Overall, our proposed framework effectively facilitates topic-based sentiment analysis based on COVID-19 tweets and detects various issues related to the COVID-19 pandemic. We believe that the proposed framework will help policymakers identify the problems associated with COVID-19 by analyzing tweets.

List of Abbreviations

BiLSTM	Bidirectional Long Short Term Memory
BoW	Bag of Words
CNN	Convolutional Neural Network
GAP	Global Average Pooling
GRU	Gated Recurrent Unit
LDA	Latent Dirichlet Allocation
LSTM	Long Short Term Memory
RNN	Recurrent Neural Network
SCS	Soft Cosine Similarity
WHO	World Health Organization
XAI	Explainable AI

Author Contributions

K.T.S. performed extensive work of literature review, conceptualization, framework construction, visualization, formal analysis, resources, experimentation, and writing. I.H.S. reviewed, validated, conceptualized, and provided the necessary guidance and supervision throughout this work. All authors have read and agreed to the published version of the manuscript.

Availability of Data and Materials

Data used in this work can be made available upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

The study did not receive any external funding and was conducted using only institutional resources.

Acknowledgments

This work has been done as part of a Master's thesis at Chittagong University of Engineering and Technology. The authors also confirm that no content in this manuscript was generated using artificial intelligence (AI) tools and that they take full responsibility for the accuracy and integrity of the work.

References

[1] M. O'brien, K. Moore, and F. McNicholas, "Social media spread during Covid-19: The pros and cons of likes and shares," *Ir. Med. J.*, vol. 113, no. 4, 2020, Art. no. 52. [View Online]

[2] K. Jahanbin and V. Rahmanian, "Using twitter and web news mining to predict COVID-19 outbreak," *Asian Pac. J. Trop. Med.*, vol. 13, no. 8, 2020, Art. no. 378. [CrossRef]

[3] M. Malta, A. W. Rimoin, and S. A. Strathdee, "The coronavirus 2019-nCoV epidemic: Is hindsight 20/20?," *EClinicalMedicine*, vol. 20, 2020, Art. no. 100289. [CrossRef]

[4] C. Sohrabi et al., "World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)," *Int. J. Surg.*, vol. 76, pp. 71–76, 2020. [CrossRef]

[5] S. A. El Rahman, F. A. AlOtaibi, and W. A. AlShehri, "Sentiment analysis of twitter data," in *Proc. 2019 Int. Conf. Comput. Inf. Sci. (ICCIS)*, 2019, pp. 1–4. [CrossRef]

[6] P. Mehta and S. Pandya, "A review on sentiment analysis methodologies, practices and applications," *Int. J. Sci. Technol. Res.*, vol. 9, no. 2, pp. 601–609, 2020. [View Online]

[7] I. H. Sarker, "Data science and analytics: An overview from data-driven smart computing, decision-making and applications perspective," *SN Comput. Sci.*, vol. 2, no. 5, pp. 1–22, 2021. [CrossRef]

[8] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003. [View Online]

[9] K. T. Shahriar, M. A. Moni, M. M. Hoque, M. N. Islam, and I. H. Sarker, "SATLabel: A framework for sentiment and aspect terms based automatic topic labelling," in *Machine Intelligence and Data Science Applications*, Berlin/Heidelberg, Germany: Springer, 2022, pp. 1–12. [CrossRef]

[10] D. Zeng, Y. Dai, F. Li, R. S. Sherratt, and J. Wang, "Adversarial learning for distant supervised relation extraction," *Comput. Mater. Contin.*, vol. 55, no. 1, pp. 121–136, 2018. [CrossRef]

[11] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, no. 4, 2018, Art. no. e1253. [CrossRef]

[12] Y. Zhang et al., "Chinese medical question answer selection via hybrid models based on CNN and GRU," *Multimed. Tools Appl.*, vol. 79, no. 21, pp. 14751–14776, 2020. [CrossRef]

[13] A. A. Sharfuddin, M. N. Tihami, and M. S. Islam, "A deep recurrent neural network with bilstm model for sentiment classification," in *Proc. 2018 Int. Conf. Bangla Speech Lang. Process. (ICBSLP)*, 2018, pp. 1–4. [CrossRef]

[14] H. Wang et al., "Phase-adjusted estimation of the number of coronavirus disease 2019 cases in Wuhan, China," *Cell Discov.*, vol. 6, no. 1, pp. 1–8, 2020. [CrossRef]

[15] A. Abd-Alrazaq, D. Alhuwail, M. Househ, M. Hamdi, and Z. Shah, "Top concerns of tweeters during the COVID-19 pandemic: Infoveillance study," *J. Med. Internet Res.*, vol. 22, no. 4, 2020, Art. no. e19016. [CrossRef]

[16] S. Hingmire, S. Chougule, G. K. Palshikar, and S. Chakraborti, "Document classification by topic labeling," in *Proc. 36th Int. ACM Sigir Conf. Res. Dev. Inf. Retr.*, 2013, pp. 877–880. [CrossRef]

[17] B. Wang, M. Liakata, A. Zubiaga, and R. Procter, "A hierarchical topic modelling approach for tweet clustering," in *International Conference on Social Informatics*, Berlin/Heidelberg, Germany: Springer, 2017, pp. 378–390. [CrossRef]

[18] A. S. Hourani, "Arabic topic labeling using Naïve Bayes (NB)," in *Proc. 2021 12th Int. Conf. Inf. Commun. Syst. (ICICS)*, 2021, pp. 478–479. [CrossRef]

[19] C. B. Asmussen and C. Møller, "Smart literature review: A practical topic modelling approach to exploratory literature review," *J. Big Data*, vol. 6, no. 1, pp. 1–18, 2019. [CrossRef]

[20] D. Maier et al., "Applying LDA topic modeling in communication research: Toward a valid and reliable methodology," *Commun. Methods Meas.*, vol. 12, no. 2–3, pp. 93–118, 2018. [CrossRef]

[21] L. Guo, C. J. Vargo, Z. Pan, W. Ding, and P. Ishwar, "Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling," *J. Mass Commun. Q.*, vol. 93, no. 2, pp. 332–359, 2016. [CrossRef]

[22] C. Ordun, S. Purushotham, and E. Raff, "Exploratory analysis of COVID-19 tweets using topic modeling, umap, and digraphs," *arXiv*, 2020. [CrossRef]

[23] S. Prabhakaran, "Topic modeling with gensim (python)," *Machine Learning Plus*, 2018. [Online]. Available: https://www.machinelearningplus.com/nlp/topic-modeling-gensim-python/

[24] H. Khurana and S. K. Sahu, "Bat inspired sentiment analysis of Twitter data," in *Progress in Advanced Computing and Intelligent Engineering*, Berlin/Heidelberg, Germany: Springer, 2018, pp. 639–650. [CrossRef]

[25] M. I. Prabha and G. U. Srikanth, "Survey of sentiment analysis using deep learning techniques," in *Proc. 2019 1st Int. Conf. Innov. Inf. Commun. Technol. (ICIICT)*, 2019, pp. 1–9. [CrossRef]

[26] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data," *Inf. Process. Manag.*, vol. 58, no. 1, 2021, Art. no. 102435. [CrossRef]

[27] S. Poria, N. Majumder, D. Hazarika, E. Cambria, A. Gelbukh, and A. Hussain, "Multimodal sentiment analysis: Addressing key issues and setting up the baselines," *IEEE Intell. Syst.*, vol. 33, no. 6, pp. 17–25, 2018. [CrossRef]

[28] S. Kaur, P. Kaul, and P. M. Zadeh, "Monitoring the dynamics of emotions during COVID-19 using Twitter data," *Procedia Comput. Sci.*, vol. 177, pp. 423–430, 2020. [CrossRef]

[29] L. Nemes and A. Kiss, "Social media sentiment analysis based on COVID-19," *J. Inf. Telecommun.*, vol. 5, no. 1, pp. 1–15, 2021. [CrossRef]

[30] Q. Jiang, L. Chen, W. Zhao, and M. Yang, "Toward aspect-level sentiment modification without parallel

data," *IEEE Intell. Syst.*, vol. 36, no. 1, pp. 75–81, 2021. [CrossRef]

[31] A. S. Imran, S. M. Daudpota, Z. Kastratiand, and R. Bhatra, "Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets," *IEEE Access*, vol. 8, pp. 181074–181090, 2020. [CrossRef]

[32] G. Barkur and G. B. Kamath, "Sentiment analysis of nationwide lockdown due to COVID 19 outbreak: Evidence from India," *Asian J. Psychiatry*, vol. 51, 2020, Art. no. 102089. [CrossRef]

[33] F. Rustam, M. Khalid, W. Aslam, V. Rupapara, A. Mehmood, and G. S. Choi, "A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis," *PLoS ONE*, vol. 16, no. 2, 2021, Art. no. e0245909. [CrossRef]

[34] U. Naseem, I. Razzak, M. Khushi, P. W. Eklund, and J. Kim, "COVIDSenti: A large-scale benchmark Twitter data set for COVID-19 sentiment analysis," *IEEE Trans. Comput. Soc. Syst.*, vol. 8, no. 4, pp. 1003–1015, 2021. [CrossRef]

[35] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proc. 2014 Conf. Empir. Methods Nat. Lang. Process. (EMNLP)*, 2014, pp. 1532–1543. [CrossRef]

[36] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," *Trans. Assoc. Comput. Linguist.*, vol. 5, pp. 135–146, 2017. [CrossRef]

[37] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv*, 2013. [CrossRef]

[38] L. Stappen, A. Baird, E. Cambria, and B. W. Schuller, "Sentiment analysis and topic recognition in video transcriptions," *IEEE Intell. Syst.*, vol. 36, no. 2, pp. 88–95, 2021. [CrossRef]

[39] H. Jelodar, Y. Wang, R. Orji, and S. Huang, "Deep sentiment classification and topic discovery on novel coronavirus or COVID-19 online discussions: NLP using LSTM recurrent neural network approach," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 10, pp. 2733–2742, 2020. [CrossRef]

[40] M. S. Ahmed, T. T. Aurpa, and M. M. Anwar, "Detecting sentiment dynamics and clusters of Twitter users for trending topics in COVID-19 pandemic," *PLoS ONE*, vol. 16, no. 8, 2021, Art. no. e0253300. [CrossRef]

[41] J. Xue, J. Chen, C. Chen, C. Zheng, S. Li, and T. Zhu, "Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter," *PLoS ONE*, vol. 15, no. 9, 2020, Art. no. e0239441. [CrossRef]

[42] R. J. Medford, S. N. Saleh, A. Sumarsono, T. M. Perl, and C. U. Lehmann, "An 'infodemic': Leveraging high-volume Twitter data to understand early public sentiment for the coronavirus disease 2019 outbreak," *Open Forum Infect. Dis.*, vol. 7, no. 7, 2020, Art. no. ofaa258. [CrossRef]

[43] X. Xiang et al., "Modern senicide in the face of a pandemic: An examination of public discourse and sentiment about older adults and COVID-19 using machine learning," *J. Gerontol. Ser. B*, vol. 76, no. 4, pp. e190–e200, 2021. [CrossRef]

[44] M. S. Satu et al., "TClustVID: A novel machine learning classification model to investigate topics and sentiment in COVID-19 tweets," *Knowl.-Based Syst.*, vol. 226, 2021, Art. no. 107126. [CrossRef]

[45] R. Chandrasekaran, V. Mehta, T. Valkunde, and E. Moustakas, "Topics, trends, and sentiments of tweets about the COVID-19 pandemic: Temporal infoveillance study," *J. Med. Internet Res.*, vol. 22, no. 10, 2020, Art. no. e22624. [CrossRef]

[46] R. K. Rastogi, "Performing sentiment analysis using Twitter data!", *Analytics Vidhya Blog*, 2021. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/07/performing-sentiment-analysis-using-twitter-data/

[47] W. Wang, S. J. Pan, D. Dahlmeier, and X. Xiao, "Coupled multi-layer attentions for co-extraction of aspect and opinion terms," *Proc. AAAI Conf. Artif. Intell.*, vol. 31, no. 1, 2017. [CrossRef]

[48] I. H. Sarker, *AI-Driven Cybersecurity and Threat Intelligence: Cyber Automation, Intelligent Decision-Making and Explainability*. Berlin/Heidelberg, Germany: Springer, 2024. [CrossRef]

[49] F. Rustam, A. Mehmood, M. Ahmad, S. Ullah, D. M. Khan, and G. S. Choi, "Classification of shopify app user reviews using novel multi text features," *IEEE Access*, vol. 8, pp. 30234–30244, 2020. [CrossRef]

[50] S. K. Habibabadi and P. D. Haghighi, "Topic modelling for identification of vaccine reactions in twitter," in *Proc. Australas. Comput. Sci. Week Multiconference*, 2019, pp. 1–10. [CrossRef]

[51] B. Jang, I. Kim, and J. W. Kim, "Word2vec convolutional neural networks for classification of news articles and tweets," *PLoS ONE*, vol. 14, no. 8, 2019, Art. no. e0220976. [CrossRef]

[52] S. Boussaadi, H. Aliane, and A. Ouahabi, "The researchers profile with topic modeling," in *Proc. 2020 IEEE 2nd Int. Conf. Electron. Control. Optim. Comput. Sci. (ICECOCS)*, 2020, pp. 1–6. [CrossRef]

[53] G. Sidorov, A. Gelbukh, H. Gómez-Adorno, and D. Pinto, "Soft similarity and soft cosine measure: Similarity of features in vector space model," *Comput. Sist.*, vol. 18, no. 3, pp. 491–504, 2014. [CrossRef]

[54] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997. [CrossRef]

[55] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of LSTM and BiLSTM in forecasting time series," in *Proc. 2019 IEEE Int. Conf. Big Data (Big Data)*, 2019, pp. 3285–3292. [CrossRef]

[56] J. Li, Y. Han, M. Zhang, G. Li, and B. Zhang, "Multi-scale residual network model combined with Global Average Pooling for action recognition," *Multimed. Tools Appl.*, vol. 81, no. 1, pp. 1375–1393, 2022. [CrossRef]

[57] S. Kamila, M. Hasanuzzaman, A. Ekbal, and P. Bhattacharyya, "Resolution of grammatical tense into actual time, and its application in Time Perspective study in the tweet space," *PLoS ONE*, vol. 14, no. 2, 2019, Art. no. e0211872. [CrossRef]

[58] K. T. Shahriar, M. N. Islam, M. M. Anwar, and I. H. Sarker, "COVID-19 analytics: Towards the effect of vaccine brands through analyzing public sentiment of tweets," *Inform. Med. Unlocked*, vol. 31, 2022, Art. no. 100969. [CrossRef]

[59] I. H. Sarker, "LLM potentiality and awareness: A position paper from the perspective of trustworthy and responsible AI modeling," *Discov. Artif. Intell.*, vol. 4, no. 1, 2024, Art. no. 40. [CrossRef]

[60] A. Ezen-Can, "A comparison of LSTM and BERT for small corpus," *arXiv*, 2020. [CrossRef]

Disclaimer/Publisher's Note: The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the publisher or editors. The publisher and editors assume no responsibility for any injury or damage resulting from the use of information contained herein.