*Nanotechnology Perceptions*
*ISSN 1660-6795*
*www.nano-ntp.com*

# Improving The Classification Accuracy Of Opinion For Big Data Decision-Making In Social Media

**Nilam Deepak Padwal[1*], Dr. Kamal Alaskar[2]**

[1*]*Asst. prof, Research scholar, Computer application,Bharati Vidyapeeth (Deemed to be university) Institute of management Kolhapur. Nilam17aug@gmail.com.*
[2]*Professor, Computer application, Bharati Vidyapeeth (Deemed to be university) Institute of management Kolhapur.Kamal.alaskar@bharatividyapeeth.edu*

The surge of technology and the Internet globally has led to the ubiquitous use of social media. Opinion mining facilitates a crucial role in analyzing the context of text by recognizing the sentiment. Twitter has become a popular platform for analyzing sentiment,and numerous studies have focused on textual sentiment analysis and opinion mining of Twitter data. The deep neural network has become a vital solution in addition to traditional big data technologies to handle the abundance of social data evolving in social networks.Recently, deep learning models have proved promising results in opinion mining in handling sequences of arbitrary data lengths.Despite the impressive outcomes observed in the earlier studies, these approaches face difficulty in analyzingthe opinions of the users due to linguistic and grammatical errors, lack of aspect-level consideration of the words or phrases, ignoring the effect of emojis and emoticons in text and vanishing gradient problem in transformer models. In addition, human-generated language comprises various noises, which misleads the opinion-mining decision-making in a big data environment. Hence, this work aims to developa novel big data decision-making strategy for Twitter data involving denoising, pretrained model-assisted aspect-level feature extraction, and deep neural network-aided opinion classification. Several fundamental natural language preprocessing tasks are applied to improve the tweets' quality, particularly the emoji lexicon analysis that converts all emojis and emoticons into textual content to improvethe polarity recognition performance. Subsequently, a denoising strategy in the proposed feature extraction phase plays a significant role in removing unwanted entities through key entity recognitionandhandlingword ambiguity with the assistance of semantic knowledge sources. By providing the denoised texts as input to the RoBERTa model, the proposed approach generates the embedding representation with the aspect-level scorefor the input tweet. The aspect-level tweet embeddings are fed into Bi-LSTM with a self-attention mechanism to classify the tweets into positive and negative classes. Thus, the experimental results illustrate that the proposed RoBERTa-self-attention-based Bi-LSTM model was evaluated on two widely used opinion mining datasets: Coronavirus Tweets and Sentiment140, obtaining 95.51% and 94.05% classification accuracy.

**Keywords:** Aspect-level, Deep learning, Denoising, Opinion mining, RoBERTa, Self-Attention, Social Network, and Twitter data.

## 1. Introduction

The advancement of Internet technology has greatly influenced the proliferation of social media platforms. In the age of technology, social media is becoming increasingly popular and significant due to the enormous number of users who engage in different forms of social interaction on various platforms. Social media imparts a platform for individuals to express their opinions, feedback, and emotions on various subjects such as products, services, events, and topics via social media platforms, blogs, channels, forums, and review sites. Social media platforms are crucial in facilitating individuals to express their views and access their opinions [1].Therefore, the increasing amount of information on social media has made sentiment analysisa relevant area for discovering the opinions of others [2]. Owing tothe many potential impacts created by online social networks from economic, political, and social perspectives, opinion mining has become a hot research topic in the academic field.In day-to-day life, the rapid growth of social information has reached the “5V” behaviors of ‘Big Data’, including Volume, Variety, Velocity, Value, and Veracity.Organizations and institutions monitor multiple social media platforms in real-time and respond to decision-making accordingly benefit greatly from opinion mining [3]. The objective of opinion mining relies on Natural Language Processing (NLP) for gathering and analyzing sentiment words and opinions [4]. As a subfield of NLP, opinion mining is incredibly useful and helps to accomplish various applications such as monitoring public mood for market intelligence, predicting movie sales, measuring customer satisfaction, creating recommendation systems, determining patient healthcare coverage, and movie sales prediction and tracking political movements [5].

Detecting public emotions can be expedited through textual data, which is generously available on social media platforms. Social media platforms, including Twitter and Facebook, proffer a wealth of timely information and facilitate user feedback and discussion in several domains. Posts on social media generate messages, referred to as tweets or microblogs, which may comprise emotional expressions such as happiness, sadness, anger, depression, boredom, and fatigue [6]. However, identifying, continuously monitoring, and filtering the information on social media platforms for sentiment analysis is difficult. Such complications include the occurrence of unstructured data, variations in language, diversity of websites and social media platforms, and heterogeneous data about individual opinions [7]. Thus, it is requisite to have suitable tools and algorithms for sentiment analysis of the data collected from social media.The information shared on Twitter is valuable for investigating crowd emotions, which can help analyze people's moods and behaviors [8]. Due to the limited character count per tweet, users are encouraged to express their thoughts, ideas, opinions, and information compactly. This limitation promotes conciseness and motivates users to condense their messages into the most important elements. Handling misspellings, emoticons, and other inconsistencies in social media language can sometimes be challenging. However, gaining valuable patterns from the tweets is challenging due to their conciseness, context dependency, linguistic nuances and presence of errors, necessitating the use of intelligent text classification methods to automate Twitter data analysis.

In a big data environment, deep learning models have been widely used as a prominent solution to analyze social network data.Over the decades, deep learning techniques are increasingly applied for sentiment or opinion mining, as they effectively understand natural language texts [9]. Deep learning architetures, such as Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN), solve several problems related to opinion mining [10]. In NLP, word embedding techniques are developed to extract the semantic and syntactic information from word representations, which are inadequate for opinion mining. Language models, including Bidirectional Encoder Representations from Transformers (BERT), have achieved exceptional text recognition and classification outcomes, particularly in NLP. However, the conventional word embedding techniques haveseveral issues, such as the need fora large corpus for generating embedding for a single word, being unable to deal with Out of Vocabulary (OOV)words, vector similarity in different sentences with similar words, and sentimental and contextual loss. The transformer-based deep learning models can cope with the issues above, capture contextual relationships between words, and generate high-quality representations of texts. Still, the consideration of the linguistic and grammatical errors, aspect-level features of the words or phrases, the effect of emojis and emoticons and the vanishing gradient problem in tweets are limited. Therefore, developing a robust big data decision-making strategy that can effectively handle the abundance of natural language text and automatically identify and extract opinions from denoised tweet content is essential.

**The key contributions of this work are presented as follows.**

- This work designs the big data decision-making strategy for opinion mining in the social network, incorporating denoising, aspect-level feature extraction, and classification for opinionated tweet recognition.
- In the proposed system, the denoising processinvolvesnatural language preprocessing, emoji and emoticon handling, and key entity recognition. In particular, the denoising strategy removes the unwanted entities from the preprocessed text using key entity recognition and handles ambiguity in the texts by gaining semantic knowledge from the resources such as WordNet and ConceptNet.
- To handle theabundance of data in the social network, the proposed approach employs deep neural networks, such as the pretrained RoBERTa modelfor the embedding representation and Bi-LSTM model for classification.
- Tolearn the training patterns of tweet opinionsat the aspectlevel, RoBERTa-assisted aspect-level tweet embedding is modeledfor the denoised tweet in the subsequence of aspect-level entity recognition and aspect polarity mapping.
- Moreover, the proposed approachdesigns aself-attention mechanism in theBidirectional Long Short-Term Memory(Bi-LSTM) model with implicit aspect term information to capture the potential implicit information to recognize opinion polarity accurately.

## 2. Literature Survey

Deep learning techniques have gained massive popularity over machine learning due to their reliable result in text classification and their ability to model nonlinear and complicated data relationships. Deep learning has gained popularity in NLP, and numerous researchers have explored deep learning for Twitter data-based opining mining. This section different deep learning-based approaches for opinion mining.

A sentiment analysis model [11] was presented to categorize user sentiment using Twitter data and user behavioral information. This model involves preprocessing steps such as normalizing, tokenizing, and part-of-speech tagging and adopts CNN for sentiment classification. This model analyzes the relationship between the user’s tweets and behavior, which is beneficial to improve the performance of sentiment analysis.An efficient sentiment analysis framework [12] was proposed to classify sentiments on a large Twitter corpus. This joint framework utilizes recurrent variants such as Bi-LSTM, Gated Recurrent Unit (GRU) and Bi-GRU on CNN to capture long-term dependencies with domain-specific word embeddingsas input by unsupervised learning. A co-attention networks approach [13] was developed to conduct aspect-based sentiment analysis to reduce noise word effects in social media data. With the Coattention-LSTM network, this approach learns nonlinear representations of context and target and extracts more efficient sentiment features using the co-attention mechanism.

A novel big data approach [14] was recommended to classify topics and analyze sentiments from Twitter data. This method exploited the Hadoop framework to extract and structure real-time Twitter data by applying Apache Flume and Apache Hive and classifying topic-focused tweets as positive or negative based on their polarity. Subsequently, this method implemented the Hybrid Lexicon-Naive Bayesian Classifier (HL-NBC) for sentimental analysis by integrating the Lexicon feature with the NBC, which enhances the classification's efficiency. A sentiment classification [15] method was proposed to analyze the sentiments of Twitter data in Arabic. This method adopts LSTM neural network to classify the sentimental polarity with less parameter computation, limited time consumption and high accuracy. A big data approach [16] for sentiment analysis was presented by customizing the Hadoop framework and deep learning classifier. This approach facilitates the Hadoop cluster to distribute data and extract essential features. Afterward, a deep recurrent neural network classifier is applied to assign a real-valued review, ultimately classifying the Twitter data's sentimental polarity.A deep learning-based scheme [17] was presented for sentiment mining on product reviews from Twitter. This scheme integrates Term Frequency and Inverse Document Frequency (TF-IDF) weighted Glove word embedding with CNN-LSTM architecture. This scheme produced high predictive performance and accuracy by evaluating different word embeddings and weighted functions. A transformer-based technique [18] was implemented to perform sentiment analysis on Twitter data by encoding representations from transformers and adopting deep intelligent contextual embedding. This technique exploits spell correction, sentiment-aware tokenization, word segmentation, and normalization as preprocessing, subsequently, deep intelligent contextual embedding with transformer-based word representation to contend with language ambiguity. A novel deep learning-based framework [19] was discovered for conducting Arabic language sentiment analysis from the user's information shared on social networks. It joints one-layer CNN architecture with two LSTM layers leveraged by the FastText word embedding model, which learns semantic and syntactic information. A deep learning method [20] was proposed to perform sentiment analysis on Twitter data. This method categorizes tweets as a Negative or Positive class from twitter data using efficient deep learning architecture with tuned hyperparameters on CNN layers followed by Bidirectional LSTM neural networks.

A new Attention-based Bidirectional CNN-RNN Deep Model [21] was introduced for sentiment analysis. This model used pre-trained GloVe word embeddings for initialization, then, Bi-LSTM and GRU networks on top of the embedding layer to extract past and future context for informative semantic representations. A novel deep learning model [22] was proposed to classify emotions using sentiment analysis for US airline service. This model effectively integrates different wordembeddings with deep learning to examine the Twitter data for multi-class sentiment analysis and utilizes DNN and CNN for feature extraction.The sentiment analysis approach [23] was developed for movie reviews on Twitter data. This approach introduced a new feature by combining tweet words, word2vec, and stop wordsfrom providinga high effect on computing the movie reviews and applyingCNN and LSTM models with natural language processing.A novel Multi-level Hybrid Aspect-Based technique [24] was developed to perform finer-grained sentiment analysis at the aspect level on Twitter data. This technique combines a feature ranking process with an amended feature selection method to enhance Twitter sentiment classification, which uses an Artificial Neural Network, specifically the Multi-Layer Perceptron (MLP), to achieve superior outcomes. A customized ensemble deep-learning language framework [25] was introduced to improve sentiment classification performance. This framework employs advanced word embedding and an LSTM network to comprehend contextual relationships between words and identify unfamiliar or infrequent terms related to the coronavirus pandemic by analyzing suffixes and prefixes found in the Twitter coronavirus data.

A deep learning-based framework [26] was developed to measure the airport service quality to conduct sentiment mining. This framework applies the LSTM model as a predictor of positive or negative sentiment among the passengers from the airport tweet data in two languages. Hybrid neural network architecture [27] was presented to forecast customer sentiments using sentiment mining. This architecture comprises both CNN and LSTM on top of pre-trained word embeddings to analyze the sequential arrangements in large social media reviews. This architecture is applied to perform domain-independent sentiment analysis, which is conducive to analyzing the reviews from different domains. Twitter sentiment analysis [28] approach was presented to examine the effect of Twitter data analysis on the mental health status of the public. This approach adopts a deep learning algorithm using RNN to categorize the sentiments as positive, negative and neutral opinions from the Twitter data based on hashtag keywords, such as COVID-19, coronavirus, deaths, and new cases. The topic-level sentiment analysis framework [29] used topicmodeling and deep learning. This framework detects topics using online learning from large-scale streaming data and effectively conducts sentiment mining by applying topic-level attention LSTM. A hybrid deep learning approach [30] was introduced to classify sentiments utilizing examining consumer opinions using online reviews. This approach adopts Keras embedding and the CNN-LSTM model to classify online reviews of airline data from airlinequality and Twitter.To accurately perform the sentiment analysis, the research work [31] presents a hybrid deep learning model that integrates the Robustly optimized BERT with the LSTM.Applying the transformer model addresses the limitations of the sequence model by presenting a compact potential representation of word embedding.

## 3. Problem Formulation

Various machine and deep learning approaches face difficulty capturing text semantics and cannot accurately recognize the user's sentiment polarities. Twitter users frequently utilize informal language, abbreviations, slang, and emoticons. Such linguistic deviations add noise to the text, makingextracting sentiment difficult.Misspelling, grammatical errors, ambiguities and incomplete expression further complicates opinion mining. In twitter, human generates raw natural language texts that are linguistically and grammatically incorrect, leading to inaccurate encoding or embedding for opinion classification and becoming ineffective.Ignoring word or phrase aspects leads to misinterpretation of tweet opinion polarity. The rise of transformer models in opinion mining reaches significant progress as it captures the richer representations of input text. However, these models face difficulty addressing the vanishing gradient problem arising from long-range dependencies in the

embeddings. It necessitatesovercoming challenges, such as i) linguistically and grammatically incorrect tweets, ii) ignoring the effect of emojis and emoticons, iii) not considering the aspect of the words or phrases and iv) the vanishing gradient problem. Therefore, the proposed approach focuses on developinga dedicated framework using advanced deep learning techniques on Twitter data to conquer the listed challenges.
In opinion mining, Twitter is a familiar platform witha major concern for analyzing users' opinions. Even though such opinion mining faces several issues, which leads to cause confusion and difficulty in the classification models. Thus, the proposed approach formulates an effective classification model to analyze the sentiment expressed by the user.
Let D be the entire Twitter dataset $D = \{(tw_1; l_1), (tw_2; l_2) \ldots\ldots (tw_n; l_n)\}$, the opinion classification task can be described as learning a mapping function $f: TW \rightarrow L$ where tweet $tw \in TW$. Where the Twitter text set denoted as $TW = (tw_1, tw_2 \ldots tw_n)$ and the label set denoted as $L = (l_1, l_2 \ldots l_n)$.

**Task 1:**To improve the quality of the tweets, preprocessing techniques are applied on the twitter dataset TW as,

$$TW \rightarrow \text{Preprocess}(TW) = (tw_1, tw_2 \ldots tw_n; P) \quad (1)$$

**Task 2:** By contemplating the effect of emoji and emoticons in the tweets, the emojis and emoticons from tweets are converted into text as,

$$f: (\{EM_n\}) \rightarrow \{TX_n\} \quad (2)$$

Where,$EM_n \in TW$. In equation (2), $EM_n$ represents a set of emojis and emoticons in the tweets and $TX_n$ represents the set of corresponding textual representations for each emoji and emoticon in tweets.

**Task 3:**From the Twitter set, unwanted entities and errors are eliminated for de-nosing from tweets and textual representations,

$$\text{Denoise}\{(tw_1, tw_2 \ldots tw_n; P), \{TX_n\}\} \rightarrow \{w_1, w_2, .. w_n\} \quad (3)$$

**Task 4:**People express their opinion in various aspects in each entity. Aspects Aare extracted from denoisedtweets as,

$$\{w_1, w_2, .. w_n\} \rightarrow [A] \quad (4)$$

Where $A = \begin{bmatrix} (a_{11}, a_{12} \ldots a_{1n}) \\ (a_{12}, a_{22} \ldots a_{2n}) \\ . \\ . \\ (a_{m1}, a_{m2} \ldots a_{mn}) \end{bmatrix}$ And $A \in E, TW$.

To map the polarity of the aspects, the polarity score obtained by,

$$P_s(A) = \sum_{i=1}^{n} (\sum_{j=1}^{m} A_{ij}) = (P_{sp}, P_{sn}) \quad (5)$$

**Task 5:**Based on the aspect $A$ and polarity $P_s$,opinion classification $OC$ is carried out to categorize the sentiment classes of the users based on the equation (5).

$$[A \vee P_s] \Rrightarrow f(OC) \rightarrow (c_1, c_2) \quad (6)$$

Where, $c_1$indicates positive sentiment and $c_2$ indicates negative sentiment.

## 4. Proposed Methodology

The proposed system focuses on implementing big data decision-making for opinion recognition from tweets by denoising, RoBERTa-based aspect-level feature representation, and classification using a self-attention-based Bi-LSTM model, illustrated in Figure 1.Initially, this work applies preprocessing methodson the Twitter dataset to eliminate unnecessary data and symbols from the tweet content, involving the transformation of emojis and emoticons into text. Subsequently, a novel denoising strategyis carried out in two ways. One iseliminatingunwanted entities with the help of key entity recognition.Another denoising strategy is conducted by incorporating semantic knowledge from three resources, WordNet, and ConceptNet, to deal with ambiguities in the preprocessed text at the semantic level. In a big data environment, the denoising step is succeeded by adoptingthe RoBERTa model for the aspect-level embedding. The aspect-level tweet embedding captures the specific features ofthe entities in the tweets and produces insights into the opinion orientation of different aspects given in the tweets. Finally, this work employs a BiLSTMmodel with a design of a self-attention mechanism to accurately categorize the opinion labels for a huge collection of tweets generated in the Twitter network.

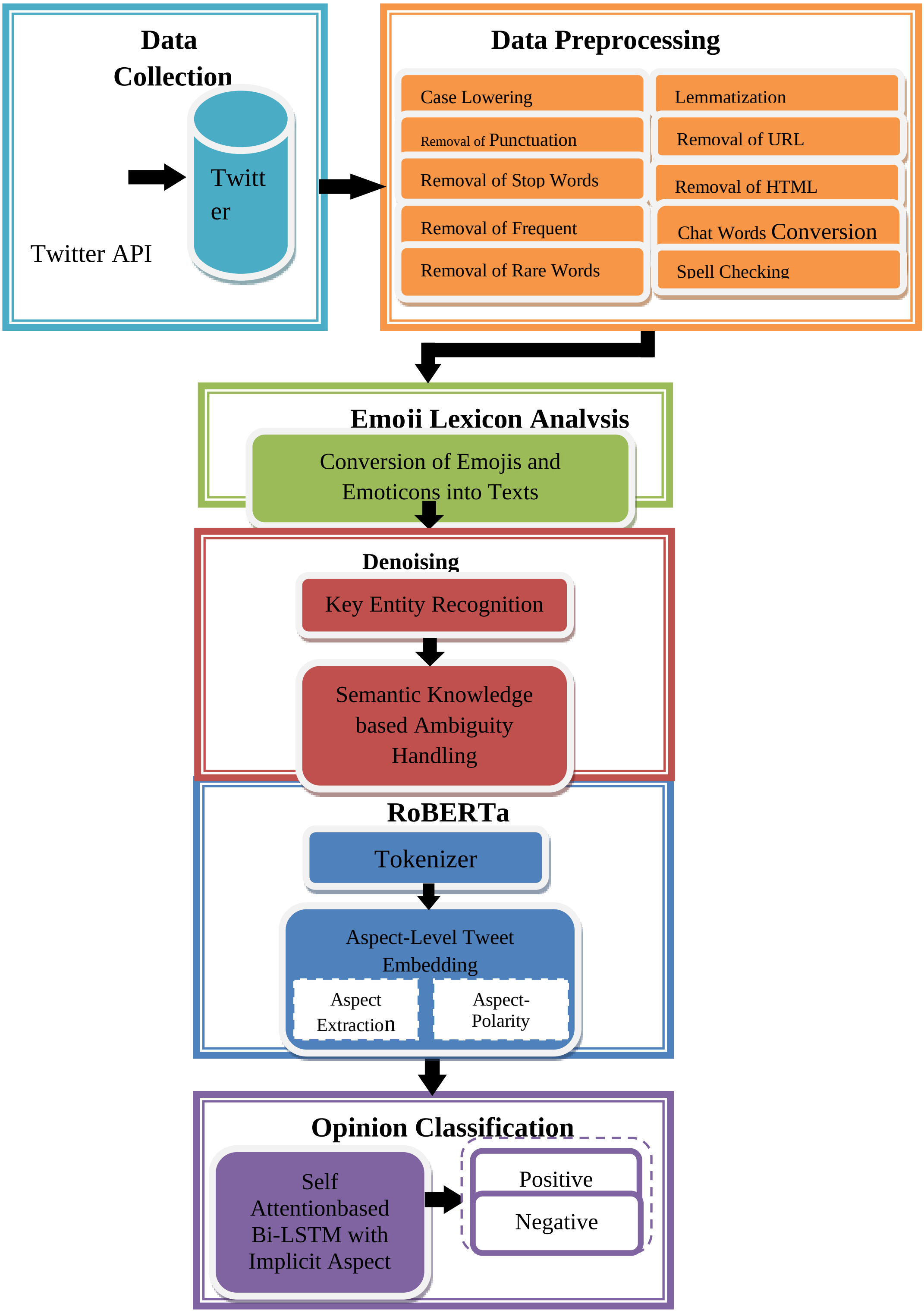

Data Collection
Twitter API
Twitt er
Data Preprocessing
Case Lowering
Lemmatization
Removal of Punctuation
Removal of URL
Removal of Stop Words
Removal of HTML
Removal of Frequent
Chat Words Conversion
Removal of Rare Words
Spell Checking
Emoji Lexicon Analysis
Conversion of Emojis and Emoticons into Texts
Denoising
Key Entity Recognition
Semantic Knowledge based Ambiguity Handling
RoBERTa
Tokenizer
Aspect-Level Tweet Embedding
Aspect Extraction
Aspect-Polarity
Opinion Classification
Self Attentionbased Bi-LSTM with Implicit Aspect
Positive
Negative

**Figure1: Proposed Methodology**

## 4.1. Preprocessing

Text data obtained from different sources, especially social media, are typically unstructured and may contain a sustainable amount of noise, spelling mistakes, and grammatical errors in their raw form. In such a way, real-time Twitter data usually comprises noise and no sentiment. Therefore, preprocessing the Twitter data before analysis is important. The purpose of this preprocessing stage is not only to improve the analysis but also to reduce the dimensionality of the input data. Many words are irrelevant and do not affect the text's sentiment, so they should be removed. In this proposed scheme, preprocessingis conducted to reduce the risk of incorrect analysis, raise the information gain and provide better performance. This proposed scheme adopts a set of preprocessing steps to transform tweets, which comprises special characters such as emojis, emoticons, hashtags, and user mentions, as well as web constructs including email addresses and URLs, and other sources of noise like phone numbers, percentages, money amounts, time, date, and generic numbers, into sentences that are constructed in a more standard form. Preprocessing of tweets is conducted using the following steps:

- Case lowering for transforming all tweets in the Twitter dataset to lowercase to assure case stability.
- Removal of punctuation marks from the tweets to concentrate on important words.
- Elimination of common stop words such as "the", "a", "an", "that", and "those" from tweets to compress the Twitter dataset, as such words do not convey valuable meaning.
- Elimination words frequently occurring in the Twitter dataset may not provide significant information.
- Removal of rare words in the Twitter dataset as they may introduce noise or do not create high impact and helps to reduce the dimensionality of Twitter data.
- Lemmatization reduces words to their base or dictionary form to normalize various word forms and minimize vocabulary size.
- Elimination of URLs or hyperlinks from the tweets as they are inappropriate textual content for opinion classification.
- Removal of any HTML tags or entities from the tweets to obtain clean and cohesive text helps to focus solely on the tweets that carry opinions.
- Conversion of chat words or abbreviations to their full forms for better clarity and to enhance the lexicon coverage.
- The spell checker is responsible for correcting any spelling mistakes in the tweets and providing improved text quality.

As a result, the data preprocessing stage implemented in the proposed approach enhances the proposed model's accuracyand speed-up the opinion mining. These steps involve removing irrelevant information from the raw text and transforming the tweets into a suitable format that can efficiently help to be processed by deep learning algorithms.

### 4.1.1. Emoji Lexicon Handling

As emojis provide valuable cues for opinion classification, this proposed approachanalyzes emoji lexicons better to understand the polarity behind the user's opinion. With the growing usage of emoji and emoticons on Twitter, such symbols have a high influence on predicting the sentiment of tweets. Each emoji and emoticon denotes a certain emotion relevant o particular sentiment. Thus, the first part of the proposed approach involves transforming emojis and emoticons into their respective text form. After preprocessing, the twitter data comprises text and emojis, reflecting the extensive utilization of emojis in tweeting conversation. This proposed approach interprets user opinions by extracting and mapping the emojis and emoticons to their corresponding text. The preprocessed tweets containtext, emojis and emoticons, which are extracted. Then, its corresponding texts are generated by matching the emojis and emoticons with emoji lexicons. This conversion permits polarity score computation of opinion classification to treat emojis and emoticons as additional textual features, ensuring a more holistic analysis of opinions.These generated texts are next fed to the further processing of opinion classification.

### 4.2. Tweet Denoising

In the proposed system, the denoising processaims to remove the entities that are irrelevant to the contribution of the opinion through the key entity recognition and semantic knowledge source-assisted ambiguity handling processes for the preprocessed tweets.

#### 4.2.1. Key Entity Recognition

In the tweet denoising phase, the proposed approachinitially focuses on utilizing key entity recognition designed to identify and filter out unwanted entities in Twitter datasets, aiming to improve opinion classification accuracy. Key entity recognition identifies and extracts irrelevant entities with no sentiment information, such as number, date and time, location, and organization names. Such entities distract the sentiment-bearing content of the tweets while analyzing the opinions in the tweets. By applying equation (7), the denoising strategy of the proposed approach removes irrelevant with the help of key entity recognition from preprocessed tweets.

$$KER\ score = \alpha * \left(log\left(\frac{N}{Df+1}\right)\right) + \beta * \left(Cos_{sim(T,KE)}\right) \quad (7)$$

In the equation, the coefficients $\alpha$ and $\beta$ are weighting parameters, and the first term measures the relevance score of the key entity within the tweets to identify the essential entities. N: The ratio between the total number of tweets$N$ and the document frequency $Df$ of the number of tweets in which the key entity appears. The second term denotes the lexicon similarity measure between key entities $KE$ of relevance and tweets $T$. By combining these two terms, this equation assistsin generating key entity recognition scoresforeach entity. Higher scores signify a higher probability of the key entity being pertinent within the tweet dataset, while lower scores signify that the entity is more prone to be irrelevant and potentially be removed.

#### 4.2.2. Semantic Knowledge-basedAmbiguity Handling

Semantic knowledge sources include lexical databases, ontologies, and knowledge graphs, furnishing structured language representations with proper meanings. Such resources can capture the relationships between words, their senses, and contextual information, facilitating a deeper comprehension of textual content. In this proposed approach, semantic knowledge from WordNet [32]and ConceptNet [33]are collectively utilized as a part of denoising to handle ambiguity among words in the tweets. WordNet is the lexicon database that impartsa comprehensive hierarchy of semantic relations for assessing rich lexical information. ConceptNet is a knowledge graph that provides a wide range of background concepts for exploring the connection between concepts and improving semantic enhancement. Exploiting the semantic knowledge from such resources helps to improve text understanding and quality via word sense disambiguation.

WordNet assembles words into synonyms and definesthe relationship among them. Such synonyms associate the word with the same meaning to disambiguate its sense. WordNet also organizes relationships of hyponymy and hypernymy to produce extra context to disambiguate word senses. To conduct word sense disambiguation, WordNet is employed to recognize and disambiguate ambiguous words in tweets. The most appropriate interpretation in the given context is determined by representing ambiguous words within synonyms to their respective senses. The disambiguation reduces noise and empowers the text's accurate analysis by solving potential ambiguities. As the semantic network, ConceptNet captures common-sense relationships between concepts in the tweets to ensure the representation of wider contextual understanding. Integrating ConceptNet into text analysis providesa better understanding of tweet texts by deploying the semantic relationships between words. The valuable contextual knowledge from semantic relationships of ConceptNet aids in disambiguating word senses. The acquisition of semantic knowledge from WordNet and ConceptNet resources achieved better denoising via word sense disambiguation and enhanced the understanding of texts in tweets.

### 4.3. RoBERTa Model-assisted Aspect-Level Feature Extraction

RoBERTa[34] is an enhanced version of the pretrained BERT model that belongs to the Transformers family of models implemented for sequence-to-sequence modeling to tackle the challenge of long-range dependencies and large-scale data handling. While the architecture of RoBERTa is closely similar to BERT, its main purpose is to improve upon BERT's performance. As a result, RoBERTa is furnished with a larger number of parameters compared to BERT models. RoBERTa basic has 123 million features, while RoBERTa wide exhibits 354 million features. Generally, the Transformer model has an encoder with a self-attention layer, a feed-forward network, a decoder with a self-attention layer, an encoder-decoder attention layer, and a feed-forward network.

In this proposed model, only the encoder part of RoBERTahas been utilized as the text encoder. Transformer models comprise three main components: tokenizers, transformers, and heads. The tokenizer converts denoised texts into sparse index encodings. The transformers component transforms the sparse content into contextual embeddings, enabling deeper training. The heads wrap the transformers model and allow the contextual embeddings to be utilized for downstream tasks. While BERT and RoBERT share similarities, BERT varies from other language models as it can learn contextual representations from both ends of sentences. BERT utilizes a vocabulary of 30,000 character-level Byte-Pair Encoding tokens for tokenization. In contrast, RoBERTa employs byte-level Byte-Pair Encoding with a huger vocabulary of 50,000 subword units. In addition, RoBERT refines the BERT model by training on more large-scale data with larger batch sizes, longer sequences, and for an extended period. RoBERT also deploys dynamic masking while training, permitting it to learn from different input sequences. The input data is prepared using the RoBERTa tokenizer, which breaks down the text into subword tokens, preserving semantic meaning while minimizing the impact of out-of-vocabulary words. Each token is assigned a unique input ID and an attention mask to indicate its relevance. The RoBERT model comprises 12 layers, each with 768 hidden states, and processes the input at many levels of abstraction. The purpose of the RoBERTa's base layers is to generate a useful word embedding, serving as a feature representation. It further enables the subsequent layers to conveniently capture the meaningful information contained within the word embedding.

#### 4.3.1. Aspect-Level Tweet Embedding

The aspect-level tweet embedding stage of the proposed approach involves two sub-processes: aspect extraction and aspect polarity mapping, described as follows.

**Aspect Extraction:** Aspect extraction is an indispensable task in opinion mining, intended to recognize and extract aspects from the opinionated text. Aspects can be both explicit and implicit, representing the particular topics that opinions are expressed. This extraction process includes identifying instances of modifiers that denotes opinions related to a certain aspect. This proposed approach extracts pairs of words based on syntactic dependency paths using a dependency parser tree. This step generates a collection of nouns, adjectives, verbs, adverbs, and word pairs that describe the opinion associated with specific aspects and serve as the outcome. Nouns are utilized to recognize groups of places, people, or things that help differentiate subjects and competitors, whereas adjectives name attributes of nouns permitting more in-depth descriptions. Verbs signify occurrences, actions, or states of being, denoting several events linked with aspects. Adverbs alter adjectives, verbs, prepositions, clauses, or sentences, ensuring further intensity. Extracting noun aspects facilitates differentiating subjects and competitors to attain opinion classification. With the help of dependency parsing, syntactic grammatical dependency relations between words in the tweets are analyzed, and aspect sentiment words, potential noun phrase aspects, and aspect-sentiment word pairs are extracted.

**Aspect-Polarity Mapping:** Aspect polarity mapping involves mapping the extracted aspects tweets to their respective polarity scores, which signifies the opinion expressed towards each aspect. The polarity score can range from positive or negative. The aspect polarity pairs provide a more precise understanding of opinions towards different aspects of the entities. The mapping of aspect polarity is accomplished by computing polarity scores.The equations to calculate the polarity scores of positive and negative aspects are given as follows;

$$Ps^{pos} = (\sum(w_i * s_{ip}) + c) / (\sum(w_i) + n) \quad (8)$$

$$Ps^{neg} = (\sum(w_i * s_{in}) + c) / (\sum(w_i) + n) \quad (9)$$

In equations (8) and (9), $Ps^{pos}$ and $Ps^{neg}$ are the positive and negative polarity scores, respectively. $w_i$ signifies the weights allotted to each opinion-bearing word based on its relevance to the aspect. $s_{ip}$ denotes the positive polarity score assigned to the opinion-bearing word, whereas $s_{in}$ denotes the negative polarity score assigned to the opinion-bearing word. $c$ is a constant parameter to adjust the overall opinion intensity, and $n$ is the number of opinion-bearing words related to the aspect.

### 4.4. Bi-LSTM-based Opinion Classification

This section explains the opinion classification process in the proposed approach, depicted in Figure 2, which comprisesself-attention-based Bi-LSTM with implicit aspect term informationthat takes input as aspect-based embeddings from the RoBERTa model.

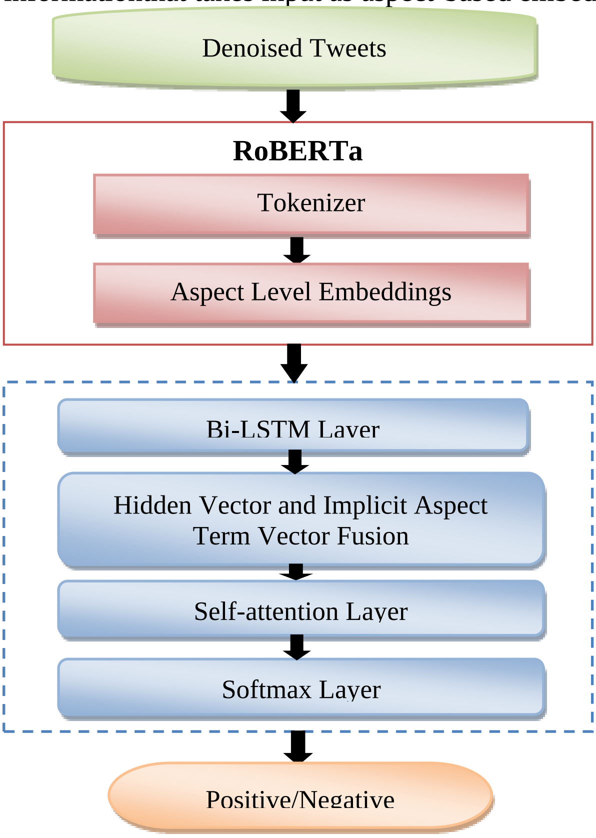



**Figure 2: Architecture of the Proposed RoBERTa-Self Attention-based Bi-LSTM model**

### 4.4.1. Self-Attention-based Bi-LSTM

Theproposed approachdesignsa self-attention mechanism in the Bi-LSTM model to classify the opinion expressed by the users on the tweets. Owing to the limitation of LSTM in capturing only the forward parts of a sentence, a bidirectional LSTM comprises two independent LSTMs for both forward pass and backward pass at every time step. Bi-LSTM facilitates the network to use contextual information from both directions of the sentence, yielding a more accurate representation of the current word's semantics. The BiLSTMcomprises forward LSTM to process the input sequence in the actual order and backward LSTM to process the sequence in reverse order to capture both past and future context simultaneously. The proposed model comprisesa Bi-LSTM, attention, and softmax layers.

**Bi-LSTM Layer:** LSTM contains a cell memory stateand three gates: the forget gate, the input gate, and the output gate. From the RoBERTamodel, word vectors in sentences are captured. Based on the computation of the single LSTM unit, the hidden vectors of the Bi-LSTM layer are produced. In the Bi-LSTM, the forward hidden vector$fh_t$ and the backward hidden vector $bh_t$ are then merged to produce the final hidden vector of the BiLSTM model. The parameters of the two opposite directions of Bi-LSTM are independent but share similar word embedding for a sentence. $\{fh_1, fh_2 \ldots. fh_n\}$ and $\{bh_1, bh_2 \ldots. bh_n\}$ indicate the forward hidden vector and the backward hidden vector, respectively. $h_n$represents the vector formed by combining $fh_n$ and $bh_n$ .The final hidden vector $h_t$ of the Bi-LSTM is, $h_t = [fh_n, bh_n]$.

**Self-Attention Layer:**Exploiting an attention layer after Bi-LSTM helps the network to gather essential information for each tweet by its hidden vector $h_t$from the Bi-LSTM layer as an input. The transformation of hidden representation permits the network to capture and handle the hidden states efficiently. The proposed approach concentrates on implicit aspect terms to capture the hidden aspects that affect the opinions expressed by users. An effective self-attention mechanism with implicit aspect term information is introduced.An embedding vector is trained for each given implicit aspect-term to leverage the implicit aspect-term information. If the aspectterm contains multiple words, such terms are represented by the average of its constituent word vectors. Assume the implicit aspect-term is comprised of $m$ words $\{ia_1, ia_2, \ldots, ia_m\}$, and its vector representation is given by equation (10). Each hidden vector $h_i$ is combined with the implicit aspect-term vector $v^{iat}$of the respective aspect-term, acting as input to the attention layer for learning an attention weight. The fusion vector $f_i$is calculated using equation (11) in which aspect-level vector representation ($\overrightarrow{v^{iat}{}_{KE}}$) obtained from the RoBERTa model for the extracted key entities in the denoising phase. Moreover, the aspect-level feature embeddings arecontextually learned in the Bi-LSTM model from the perspective of the Key Entities (KE); hence, the key entity-based weight is integrated as an element-wise multiplication in equation (11). By equation (11), the hidden representation $h_t$ is transformed into another hidden representation $u_t^A$by utilizing a fusion vector, a weight matrix $W_w$ and a bias vector $b_w$ as formulated in equation (12).

$$v^{iat} = \frac{1}{m}\sum_{i=1}^{m} e_i \qquad (10)$$

$$f_i = h_i + \omega_{KE} \odot \overrightarrow{v^{iat}{}_{KE}} \qquad (11)$$

$$u_t^A = \tanh(W_w f_i + b_w) \qquad (12)$$

Subsequently, the attention value $\partial_t$is calculated by contemplating $u_t$and a word-level context vector $u_w$using equation (13). This context vector supports discriminating the significance of different words within the sentence. The attention values are calculated so that the sum of the attention values for a sequence equals 1. The higher the attention value, the more relevant the corresponding word is in the context of sentiment polarity. Lastly, sentence vector s is computed using equation (14) and considered as the opinion features weighted by attention.

$$\partial_t = \frac{exp\ (u_t^T u_w)}{\sum_t exp\ (u_t^T u_w)} \qquad (13)$$

$$s = \sum_t \partial_t h_t \qquad (14)$$

**Softmax Layer:** The softmax layer is utilized as the classifier in the proposed approach for opinion classification. It produces a high-level sentence representation by multiplying each word's hidden vector with its respective weight. This multiplication results in a sentence vector $s$, which serves as the opinion feature for opinion polarity classification.

$$\tilde{y} = softmax(W_s s + b_s) \tag{15}$$

In equation (15), $\tilde{y}$ is the classified result through the model, the weighted matrix is represented as $W_s$, and the bias is denoted as $b_s$. Thus, the proposed approach effectively categorizes the opinion of the tweets with the help of the hybrid deep learning models of RoBERTa and Bi-LSTM associated with the denoising process.

## 5. Experimental Evaluation

This section presents the experimental settings and compares the performance of the proposed RoBERT-Bi-LSTM model with state-of-the-art methods.

### 5.1.1. Experimental Setup

The experimental setting adopts RoBERTa and Self-attention-based Bi-LSTM to classify opinions by training on two sentiment analysis datasets. The proposed model's experiments were conducted utilizing a machine running Ubuntu 16.04, a 64-bit operating system with a 3GHz Intel CPU and 16GB of memory. The experimental framework used Python machine learning libraries to implement the RoBERTa and Self-attention-based Bi-LSTM model for opinion classification. For optimal performance, the setting involves configuring preprocessing, emoticons conversion, de-nosing strategy, and aspect-level embedding generation. The trained model is then utilized to categorize the opinion of users from tweets as positive and negative.

### 5.1.2. Dataset Description

This section provides information about the dataset used in the proposed model. The datasets adopted in this proposed approacharethe sentiment140 andCoronavirus tweets NLP-text classification.Their detailed descriptions are given below;

**Sentiment140 Dataset:** The Sentiment140 dataset was obtained from Kaggle, a public repository for benchmark datasets [35]. This dataset comprised 1.6 million tweets and was collected using the Twitter search API by Stanford University. This dataset has well-balanced class distribution samples with 0.8 million positive and 0.8 million negative tweets. The tweets in the sentiment140 dataset are labeled as 0 for negative sentiment and 4 for positive sentiment. To annotate the tweets, the dataset authors examined the presence of emoticons as an indicator of positive or negative sentiment. The dataset involves six features: target, id, date, flag, user, text, and 1,600,000 annotated tweets.

**Coronavirus Tweets Dataset:** The Second Twitter dataset utilized in this proposed approach is Coronavirus Tweets NLP-Text Classification Dataset [36]. This dataset comprises 44,957 COVID-relevant tweets gathered from users who used hashtags such as #coronavirus, #coronavirusoutbreak, and #covid19. These tweets were acquired between March and April 2020 and were initially manually categorized into five polarities: Extremely Negative, Negative, Neutral, Positive, and Extremely Positive. To enable a uniform polarity distribution, this work merges the extremely negative category with the negative category and the extremely positive category with the positive category. This reconstructed dataset comprises 17,031 negative tweets, 8,334 neutral tweets, and 19,592 positive tweets.

### 5.1.3. Evaluation Metrics

To establish the dependability of the proposed model, the experiment employs the following evaluation metrics:precision, recall, Specificity, Accuracy, and Area Under the Curve (AUC). These evaluation metrics assess the efficacy of deep learning models for opining mining on Twitter data and demonstrate the strengths and weaknesses of a proposed model.

**Precision:**The ratio between the number of correctly classified positive tweets and the total number of tweets that the proposed model classified as positive to reduce the number of false positives.

**Recall:**The ratio between the number of correctly classified positive tweets and the total number of positive tweets in the Twitter dataset is favorable to reducing the number of false negatives, where the proposedmodel incorrectly classifies a negative sentiment for a positive tweet.

**F-score:** It is a harmonic mean of precision and recall. It is a balanced measure that merges both precision and recall to provide an overall evaluation of the proposed model's performance.

**Accuracy:** The accuracy is the evaluation ratio between the total number of correctly classified instances and the total number of instances in the Twitter dataset, which is beneficial to evaluate the proposed model's ability to classify tweets into positive or negative sentiments accurately.

**AUC:** The area under the ROC curve measures the overall performance and efficiency of the proposed opinion classification model.

## 5.2. Experimental Results

The experimental scheme exemplifies the performance deviation of the proposed deep learning and transformer-based opinion classification model using the following classification metrics, precision, recall, specificity, accuracy, and AUC. For experimental analysis, the experiment setting compares the performance of the proposed model in two ways on the sentiment140 dataset and Coronavirus Tweets dataset, respectively. As provided in Table 1 and Table 2, at first, the performance of the proposed model assesses across the single baseline models such as Logistic Regression, SVM, Random Forest, MLP, DNN and LSTM. Secondly, the performance of the proposed model emulates against hybrid models,including RoBERTa+LSTM, RoBERTa+GRU andRoBERTa+BiLSTMto assess the performance of the proposed approach against the different combinations of RoBERTa pretrained model and text classifiers. Moreover, the comparative evaluations incorporate the existing sentiment analysis model, a Hybrid RoBERTa and LSTM model with a data Augmentation (HRoLA) approach [34].

**Table 1: Comparative Results of the Sentiment Classification Model on Two Different Tweet Datasets**

| Models | Sentiment140 Dataset | | | | | Coronavirus Tweets Dataset | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Precision (%) | Recall (%) | F-Score (%) | Accuracy (%) | AUC (%) | Precision (%) | Recall (%) | F-Score (%) | Accuracy (%) | AUC (%) |
| SVM | 66.62 | 83.61 | 74.15 | 70.92 | 70.94 | 76.45 | 69.12 | 66.00 | 69.72 | 67.86 |
| RF | 69.39 | 78.94 | 73.29 | 70.34 | 71.37 | 77.07 | 72.57 | 70.20 | 72.57 | 69.51 |
| LR | 69.72 | 70.97 | 70.34 | 70.13 | 70.13 | 66.87 | 96.16 | 78.89 | 71.24 | 67.92 |
| MLP | 70.94 | 68.93 | 71.92 | 71.93 | 73.27 | 81.31 | 81.36 | 81.33 | 81.47 | 80.60 |
| DNN | 72.33 | 71.6 | 71.34 | 72.6 | 71.59 | 78.47 | 78.12 | 78.29 | 78.34 | 78.34 |
| LSTM | 75.12 | 72.95 | 74.02 | 74.52 | 74.51 | 82.79 | 83.25 | 83.02 | 83.31 | 83.31 |
| RoBERTa + BiLSTM (Proposed) | 85.86 | 97.52 | 91.32 | 94.05 | 93.89 | 96.37 | 94.65 | 95.51 | 95.37 | 95.18 |

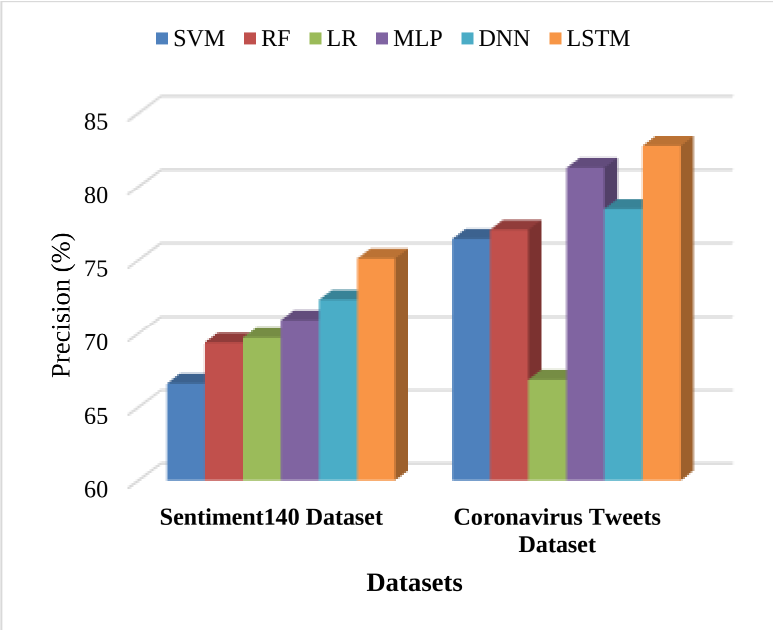
From the analysis of Table 1, it is determined that the accuracy comparisons of baseline models on sentiment140 and Coronavirus Tweets datasets. Among all the baseline models, LSTM performs best on both datasets by achieving the highest accuracies,83.31% on the Coronavirus Tweets dataset and 74.52% on the sentiment140 dataset. Contrarily, SVM records lower accuracy of69.72% on the Coronavirus Tweets dataset and moderate accuracy of 70.92% on the sentiment140 dataset. From the performance of baseline models, it is determined that the deep neural network of the LSTM model is the most suitable technique for the opinion classification on both the Sentiment140 and Coronavirus tweet datasets. Hence, the proposed approach employs the Bi-LSTM model to categorize the opinion labels of the tweets and accomplished higher classification accuracy at 94.05% and 95.37% on Sentiment140 and Coronavirus tweet datasets, respectively.

**Table 2:Performance of the Proposed Model against Comparative Models**

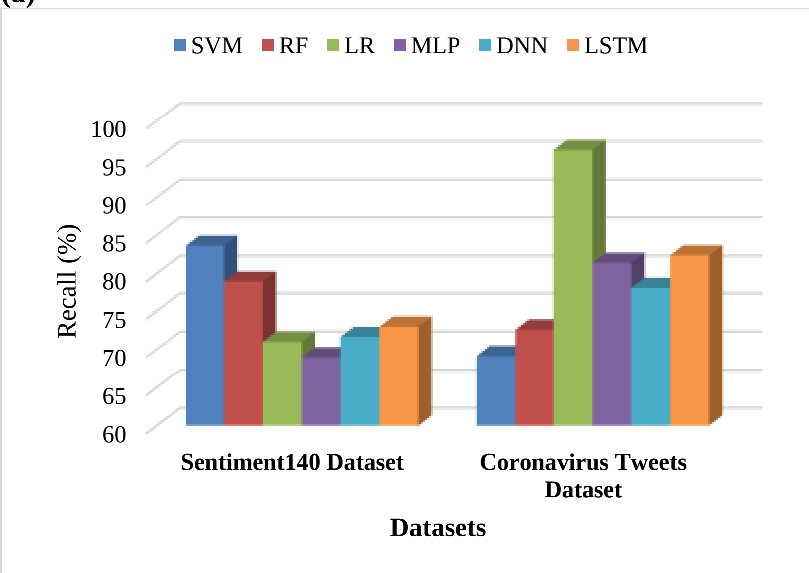
| Input | Models | Sentiment140 Dataset | Coronavirus Tweets Dataset |
|---|---|---|---|

| | | **Precision (%)** | **Recall (%)** | **F-Score (%)** | **Accuracy (%)** | **Precision (%)** | **Recall (%)** | **F-Score (%)** | **Accuracy (%)** |
|---|---|---|---|---|---|---|---|---|---|
| Raw Tweet (Without Feature Extraction) | RoBERTa + GRU | 98.23 | 80.45 | 82.8 | 84.88 | 85.99 | 83.24 | 84.59 | 84.59 |
| | RoBERTa + LSTM | 89.95 | 78.28 | 83.71 | 84.52 | 81.43 | 88.19 | 84.67 | 83.78 |
| | RoBERTa + Bi-LSTM | 90.78 | 86.51 | 88.36 | 89.12 | 93.1 | 87.88 | 90.41 | 90.32 |
| With Feature Extraction | RoBERTa + LSTM [34] | 90 | 90 | 90 | 89.70 | 91.25 | 90.58 | 90.91 | 89.91 |
| | RoBERTa + Bi-LSTM (Proposed) | 85.86 | 97.52 | 91.32 | 94.05 | 96.37 | 94.65 | 95.51 | 95.37 |

The comparative performance of accuracy of the proposed RoBERTa-BiLSTM model across the existing hybrid models such as RoBERTa+GRU, RoBERTa+LSTM and RoBERTa+BiLSTM are presented in Table 2for the Sentiment140 and Coronavirus Tweets datasets. Compared to the existing hybrid models, the proposed RoBERTa+BiLSTM model achieved higher accuracy of 95.37% on the Coronavirus Tweets dataset and 94.05% on the Sentiment140 dataset. It indicates that the proposed combination of RoBERTa and BiLSTM model surpasses other methods in accurately classifying a given tweet's opinion. The comparative sentiment analysis research work, HRoLA [34] obtained the second-highest accuracy of 89.91% on the Coronavirus Tweets dataset and 89.70% on the Sentiment140 dataset by applying the RoBERTa and LSTM models associated with the data augmentation process. Moreover, RoBERTa+GRU, RoBERTa+LSTM, and RoBERTa+Bi-LSTM models relatively achieved a moderate accuracy score of 84.88%, 84.52%, and 89.12% on the Sentiment140 dataset, respectively.The main reason behind the performance improvement of the proposed RoBERTa-BiLSTM is the denoising and aspect-level embedding in the RoBERTa model for contextual understanding to precisely understand the opinionated tweet patterns.




**(a)**




**(b)**

**Figure 3(a-b): Performance of Proposed Model Against Baseline Models, 4(a): Precision and 4(b): Recall**

Figure 3 (a) illustrates the precision comparisons of baseline models on sentiment140 and Coronavirus Tweets datasets. In the case of the Coronavirus Tweets dataset, the proposed RoBERTa+BiLSTM outperforms other models by achieving the highest precision with the value of 96.37%, denoting that it accurately classified positive opinions in the aspect of the coronavirus disease. On the contrary, RoBERTa+GRU attained the highest precision on the Sentiment140 dataset at 98.23% than the proposed RoBERTa+BiLSTM.However, the true positive rate of the RoBERTa+GRU is only 80.45%, which is a 17.5% decrease from the proposed approach due to the key entity-based aspect-level tweet embedding in the RoBERTa and self-attention mechanism in the BiLSTM model. The comparative hybrid deep models, such as RoBERTa+LSTMand RoBERTa+BiLSTM, reached the recall values of 78.28% and 86.51% on the Sentiment140 dataset and 88.19% and 87.88% on the Coronavirus Tweets dataset respectively. Even though the hybrid deep models yield fluctuations in their precision and recall performance, the accuracy of the BiLSTM model with the combination of RoBERTa only outperforms the GRU and LSTM models, depicted in Figure 4.

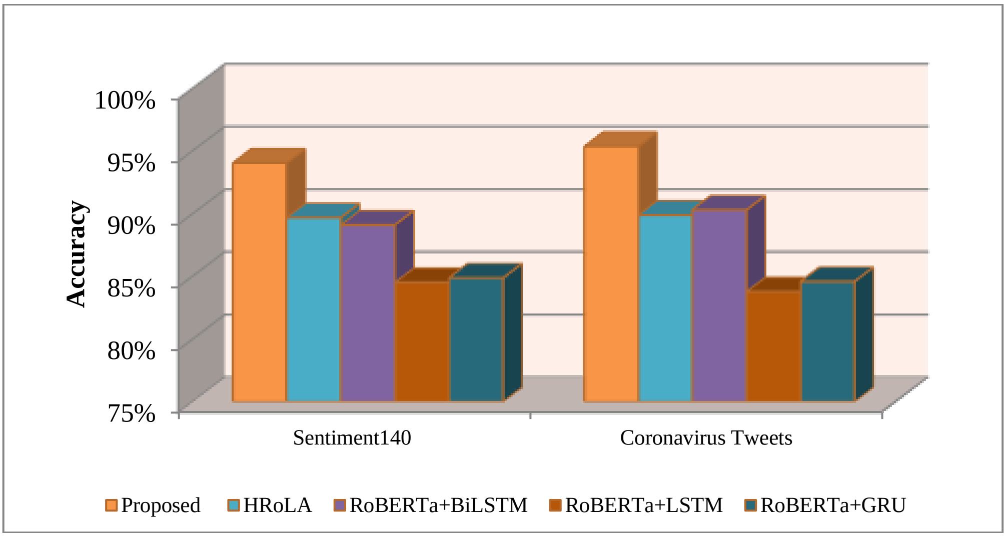
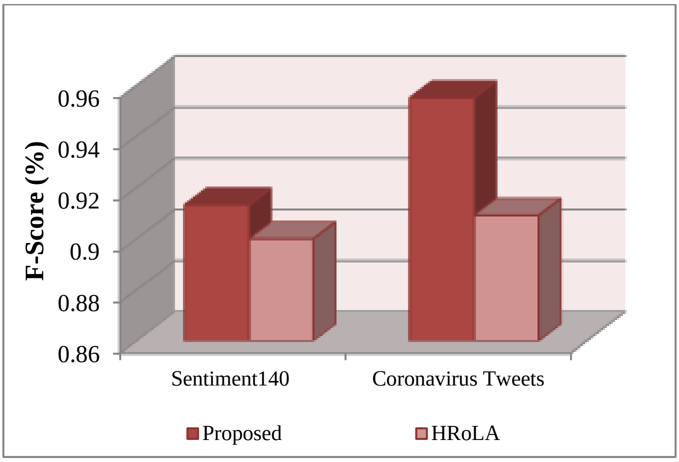
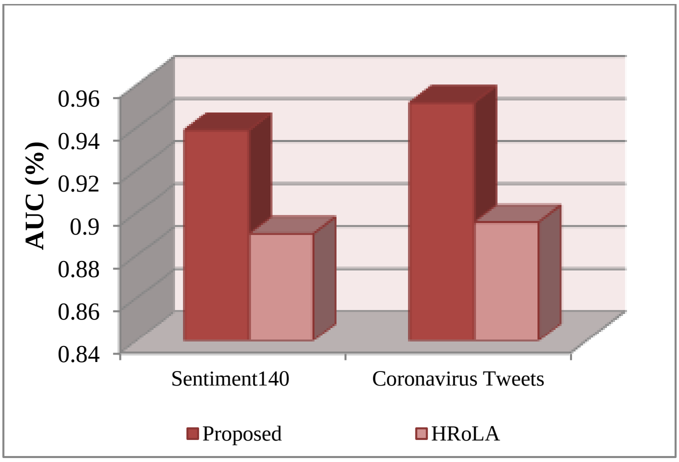



**Figure 4: Sentiment Classification Performance of Hybrid Models**

Among all the existing hybrid models, the proposed approach obtains comparatively higher accuracy at 94.05% and 95.51% on both tweet datasets due to the lack of interpreting the opinion from the perspective of aspect-level tweets in the existing HRoLA approach and comparative hybrid models. It is accomplished by combining the denoising and pre-trained model with the self-attention-based Bi-LSTM in the proposed approach. As shown in Figure 5(a), the proposed approachperforms better on both datasets by accomplishing the best F-Score of 95.51% on Coronavirus Tweets dataset and 90.91% on the sentiment140 dataset. The results indicate that the proposed RoBERTa+BiLSTM model outperforms the existing HRoLA approachand illustrates the potential influence of the denoising and aspect-level feature extraction process on the opinion classification outcome. The proposed RoBERTa+BiLSTM model uses the RoBERTa Transformer with aspect-level embeddings to capture contextual information from the tweets andthe self-attention-based Bi-LSTM layer to inherently learn the sequential information with respect to aspect-level key entities. Moreover, the proposed approach conducts Emojis lexicon analysis and semantic knowledge-based ambiguity handling, enhancing text quality for better model learning. Itmanifests that the proposed model efficaciously captures opinions expressed in the tweets from both the Coronavirus and Sentiment140 datasets.




**(a)**




**(b)**

**Figure 5(a-b): Performance of AUC on Baseline Models**

Figure5 (b) portrays the AUC comparisons of proposed and existing opinion classification approach on sentiment140 and Coronavirus datasets. For the Sentiment140 and Coronavirus datasets, the proposed approachattained an AUC of 93.89% and 95.18%, while the HRoLAapproachgets performance degradation of AUC at89.02% and 89.56%, respectively. Even though the existing HRoLA approach employs the RoBERTa transformer and LSTM model to analyze the sentiment of the tweets, the generalized examination of the raw tweets misleads the discrimination of the positive and negative classes.Hence, theproposed approachoutperformed the existing HRoLA approach on the Sentiment140 and Coronavirus datasets.As a result, integrating theRoBERTa and BiLSTMmodel with self-attention in the proposed approach, in addition to the denoising and aspect-level feature extraction,persistently assists in achieving the highest performance, signifying its superiority in capturing opinion information.

## 6. Conclusion

Owing to social media's explosive growth with big data characteristics, deep neural networks play a significant role in handling massive tweetdata in Twitter in the field of opinion mining. Opinion mining on Twitter offers valuable insights into public sentiment, making it a remarkable tool for various fields. Even though the success of deep learning in opinion mining for analyzing opinions is high, such a model faces some difficulties inproviding effective opinion classification. This research develops the RoBERTa-Self attention-based Bi-LSTM model with emojis lexicon analysis and denoising strategy. In the proposed approach, preprocessing steps help removeirrelevant information and symbols from Twitter data. Next, the two-stage denoising strategy of the proposed approach was performed, composed of key entity recognition for unwanted entity removal and utilization of semantic knowledge resources to deal with ambiguities. The RoBERTa model in the proposed approach encompasses tokenization and aspect-level tweet embedding for efficient word embedding representation. The proposed RoBERTa-self attention-based Bi-LSTM model integrates the strengths of RoBERTa and Bi-LSTM with implicit aspect information. Subsequently, this model classified the opinion from the word embedding into positive and negative opinions. The experimental results indicate that the proposed RoBERTa-self-attention-based Bi-LSTM model surpasses the state-of-the-art methods in opinion mining. This superiority is observed over two datasets, including Coronavirus tweets datasets and Sentiment140 datasets. The performance of the RoBERTa-self-attention-based Bi-LSTM model prevailsover comparative models in accurately categorizing opinions in Twitter data and yielded 95.51% accuracy in the task of opinion-mining.

## References

[1] Cortis, K., & Davis, B. (2021). Over a decade of social opinion mining: a systematic review. Artificial intelligence review, 54(7), 4873-4965.

[2] Yue, L., Chen, W., Li, X., Zuo, W., & Yin, M. (2019). A survey of sentiment analysis in social media. Knowledge and Information Systems, 60, 617-663.

[3] Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. Procedia Computer Science, 161, 707-714.

[4] Soong, H. C., Jalil, N. B. A., Ayyasamy, R. K., & Akbar, R. (2019, April). The essential of sentiment analysis and opinion mining in social media: Introduction and survey of the recent approaches and techniques. In 2019 IEEE 9th symposium on computer applications & industrial electronics (ISCAIE) (pp. 272-277). IEEE.

[5] Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. Knowledge-Based Systems, 226, 107134.

[6] Li, Z., Fan, Y., Jiang, B., Lei, T., & Liu, W. (2019). A survey on sentiment analysis and opinion mining for social multimedia. Multimedia Tools and Applications, 78, 6939-6967.

[7] Pathak, A. R., Pandey, M., &Rautaray, S. (2019). Empirical evaluation of deep learning models for sentiment analysis. Journal of Statistics and Management Systems, 22(4), 741-752.

[8] Alsaeedi, A., & Khan, M. Z. (2019). A study on sentiment analysis techniques of Twitter data. International Journal of Advanced Computer Science and Applications, 10(2).

[9] Etaiwi, W., Suleiman, D., &Awajan, A. (2021). Deep learning based techniques for sentiment analysis: A survey. Informatica, 45(7).
[10] Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. Electronics, 9(3), 483.
[11] Alharbi, A. S. M., & de Doncker, E. (2019). Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information. Cognitive Systems Research, 54, 50-61.
[12] Abid, F., Alam, M., Yasir, M., & Li, C. (2019). Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter. Future Generation Computer Systems, 95, 292-308.
[13] Yang, C., Zhang, H., Jiang, B., & Li, K. (2019). Aspect-based sentiment analysis with alternating coattention networks. Information Processing & Management, 56(3), 463-478.
[14] Rodrigues, A. P., &Chiplunkar, N. N. (2019). A new big data approach for topic classification and sentiment analysis of Twitter data. Evolutionary Intelligence, 1-11.
[15] Gwad, W. H. G., Ismael, I. M. I., &Gültepe, Y. (2020). Twitter Sentiment Analysis Classification in the Arabic Language using Long Short-Term Memory Neural Networks. Int. J. Eng. Adv. Technol.(IJEAT), 9, 2249-8958.
[16] Khan, M., & Malviya, A. (2020, February). Big data approach for sentiment analysis of twitter data using Hadoop framework and deep learning. In 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (pp. 1-5). IEEE.
[17] Onan, A. (2021). Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. Concurrency and Computation: Practice and Experience, 33(23), e5909.
[18] Naseem, U., Razzak, I., Musial, K., & Imran, M. (2020). Transformer based deep intelligent contextual embedding for twitter sentiment analysis. Future Generation Computer Systems, 113, 58-69.
[19] Ombabi, A. H., Ouarda, W., &Alimi, A. M. (2020). Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks. Social Network Analysis and Mining, 10, 1-13.
[20] Tyagi, V., Kumar, A., & Das, S. (2020, December). Sentiment analysis on twitter data using deep learning approach. In 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN) (pp. 187-190). IEEE.
[21] Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis. Future Generation Computer Systems, 115, 279-294.
[22] Hasib, K. M., Habib, M. A., Towhid, N. A., &Showrov, M. I. H. (2021, February). A novel deep learning based sentiment analysis of twitter data for us airline service. In 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD) (pp. 450-455). IEEE.
[23] Gandhi, U. D., Malarvizhi Kumar, P., Chandra Babu, G., & Karthick, G. (2021). Sentiment analysis on twitter data by using convolutional neural network (CNN) and long short term memory (LSTM). Wireless Personal Communications, 1-10.
[24] Janjua, S. H., Siddiqui, G. F., Sindhu, M. A., & Rashid, U. (2021). Multi-level aspect based sentiment classification of Twitter data: using hybrid approach in deep learning. PeerJ Computer Science, 7, e433.
[25] Alsayat, A. (2022). Improving sentiment analysis for social media applications using an ensemble deep learning language model. Arabian Journal for Science and Engineering, 47(2), 2499-2511.
[26] Barakat, H., Yeniterzi, R., & Martin-Domingo, L. (2021). Applying deep learning models to twitter data to detect airport service quality. Journal of air transport management, 91, 102003.
[27] Behera, R. K., Jena, M., Rath, S. K., & Misra, S. (2021). Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data. Information Processing & Management, 58(1), 102435.

[28] Kaur, H., Ahsaan, S. U., Alankar, B., & Chang, V. (2021). A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets. Information Systems Frontiers, 1-13.
[29] Pathak, A. R., Pandey, M., &Rautaray, S. (2021). Topic-level sentiment analysis of social media data using deep learning. Applied Soft Computing, 108, 107440.
[30] Jain, P. K., Saravanan, V., & Pamula, R. (2021). A hybrid CNN-LSTM: A deep learning approach for consumer sentiment analysis using qualitative user-generated contents. Transactions on Asian and Low-Resource Language Information Processing, 20(5), 1-15.\
[31] Tan, K. L., Lee, C. P., Anbananthen, K. S. M., & Lim, K. M. (2022). RoBERTa-LSTM: A hybrid model for sentiment analysis with transformer and recurrent neural network. IEEE Access, 10, 21517-21525.
[32] Miller, G. A. (1995). WordNet: a lexical database for English. Communications of the ACM, 38(11), 39-41.
[33] Speer, R., Chin, J., &Havasi, C. (2017, February). Conceptnet 5.5: An open multilingual graph of general knowledge. In Proceedings of the AAAI conference on artificial intelligence (Vol. 31, No. 1).
[34] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... &Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.
[35] Sentiment 140. Available online: https://www.kaggle.com/datasets/kazanova/sentiment140 (accessed May 2023).
[36] Corona virus tweets NLP – Text Classification 2020 Available online: https://www .kaggle.com/datasets /datatattle/covid-19-nlp-text-classification (accessed May 2023).