Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://journal.nsps.org.ng/index.php/jnsps

J. Nig. Soc. Phys. Sci. 7 (2025) 2300

Journal of the Nigerian Society of Physical Sciences

Optimized aspect level sentiment analysis of tweet data using deep learning and rule-based techniques

S. N. Enemuo[a], O. N. Akande[b,*], M. O. Lawrence[c], I. C. Saidu[d]

[a]*Computer Science Department, Landmark University, Kwara State, Nigeria*
[b]*Computer Science Department, Nile University, Abuja*
[c]*Computer Science Department, Miva Open University, Abuja*
[d]*Computer Science Department, Baze University, Abuja*

Abstract

Social media platforms are no longer just for interacting with others and having fun; they are also now a place for people to voice their opinions on topics that are important to them. It has developed into a potent platform in recent years whereby elections can be won or lost. Governments and organizations are therefore becoming more and more interested in the opinions that citizens voice on social media platforms. Towards coming up with a localized sentiment analysis system, this research explored the use of rules and optimized Convolutional Neural Network (CNN) Deep Learning (DL) technique for aspect-level sentiment analysis of users' tweets. A total of 26,401 tweets about the security situation in Nigeria were used as the test data. Most tweets in the Nigerian Twitter space are expressed in Pidgin language, which is a blend of English and local words, therefore, a rule-based technique was used to capture the local words and abbreviated character representations used in the Nigerian tweet data used for testing. Moreover, existing works have shown that CNN does not adequately capture modeling sequence information and long-distance dependency of texts. Therefore, this study employed existing word embedding to boost the performance of CNN for sentiment polarity classification. The proposed rule and optimized CNN model outperformed the traditional CNN and state-of-the-art Glove and word2Vec models with an accuracy of 89.31%, a recall value of 88.21%, a precision value of 88.15%, and an F1 score of 87.62%.

DOI:10.46481/jnsps.2025.2300

Keywords: Aspects extraction, Sentiment analysis, Deep learning, Data security

Article History :
Received: 09 August 2024
Received in revised form: 16 March 2025
Accepted for publication: 24 March 2025
Published: 09 April 2025

Communicated by: B. J. Falaye

1. Introduction

Every day, there is a greater requirement to record the opinions that customers or citizens have regarding goods or regulations. It serves as a method for obtaining opinions that can be utilized to evaluate how effectively a product is accepted by consumers or how adversely a government policy impacts the local populace. Sentiment is a general emotion, viewpoint, attitude, or perspective of a circumstance or object that can be communicated orally, in writing, or through sign language. Sentiment analysis, often known as opinion mining, is the study of how people feel about a given good, service, or set of government regulations. Its use spans a variety of academic disciplines, including computer science, politics, marketing, the health sciences, communications, etc.

*Corresponding author Tel. No: +234-703-956-37773.
Email address: `akande.oluwatobi@gmail.com` (O. N. Akande)

Additionally, the global acceptance of platforms of social media and the increasing adoption and internet utilization have accelerated the growth of this research area. Nowadays, people and organizations use information from various channels to make decisions. Also, to make a smart choice of whether or not to buy a product, consumers would rather read reviews about it and consider the opinions of other consumers, companies use consumers' opinions or feedback from existing products to improve new products. As a result of the amount of client feedback on product websites, businesses may not always need to conduct surveys or create a pool of opinions to gather them. The viewpoint could be categorized as conventional or comparative [1]. The common view gives an indirect or direct sentiment regarding an entity. A direct opinion, such as "the drug is potent", differs from an indirect, unfavorable opinion, such as "my headache got worse after taking the drug." Comparative opinion, on the other hand, conveys an opinion on one thing about another. For instance, a comparison viewpoint that favors the use of Panadol for treating headaches is "I prefer Panadol to Paracetamol for the treatment of headaches." Extracting and condensing these enormous amounts of thoughts, especially from social media networks, is still a laborious task.

Therefore, Social Media Networks (SMNs) are no longer only a place for people to communicate socially; they are now a powerful political instrument that can affect election results. They are being used as a tool for creating procedures and establishing agendas among politicians, the media, and the public [2]. They have become into effective platforms on which elections are won and lost because prominent members of the administration are avid users. Additionally, social movements can use SMNs as a method to disseminate their ideas, gather supporters, and take action [3]. This is evident in the roles that Facebook and Twitter played during significant political milestones: Barack Obama's emergence as the first Black American president in 2008 [4]; Recep Tayyip Erdoğan's leadership in Turkey in 2011 [5, 6]; Miro Cerar's rise in Slovenia in 2014 [7]; and Giorgio Napolitano's influence in Italy in 2013 [8]. Similarly, political developments in Belgium [9], China, and the United States have also been shaped by social media platforms [10]. Therefore, citizens have seen social media networks as a tool to express their views about issues bothering them as well as favorable and unfavorable Government policies. With this in view, instead of placing a ban on SMNs as some Governments' have been doing, these platforms can be maximized as tools for obtaining feedback from the masses about their policies. With an emphasis on security, citizens' opinions on SMNs could be fetched and analyzed to know the state of protection of a particular country.

However, harvesting opinions on SMNs (most importantly tweets) is more complex due to the inherent nature of tweets [11]. Since most tweets are not written in official English, it is permitted to use acronyms, symbols, emojis, and other abbreviations that vary from country to country. In the work of Ibekwe *et al.* [12], analysis of sentiment falls into document, sentence, and aspect-based levels. To determine if a document reflects a positive or negative sentiment, document-level sentiment analysis looks at the entire opinion document. This demonstrates how analysis of sentiment at the level of documents provides the overall sentiment polarity linked to a single entity. However, the opinion document is broken down into sentences at the sentence level, and the sentiments of each sentence are then extracted and classified as either positive or negative. This demonstrates that, compared to the document level, sentences provide a more precise sentiment polarity. ABSA, on the other hand, makes an effort to conduct a more thorough and in-depth study of conveyed sentiments that are concealed in each sentence. The aspect level directly analyses the opinion rather than focusing on language constructs like documents, sentences, paragraphs, phrases, or clauses. Aspect in this instance is the properties used to qualify the specific entity of interest. Using product review as a case study, the product itself is the entity while the price, quality, quantity etc. are the various properties that define the aspect of the product. Therefore, the multifaceted nature that the aspect of an entity may exhibit makes aspect-based sentiment analysis a hectic task. Different aspects may yield opposite polarities, and identifying the sentence parts that represent the aspects may be challenging. For instance, a comment that "It is not safe to live in Nigeria, but the cost of living is low" will be categorized as negative if the aspect being considered is "safe" but positive if the aspect being considered is "condition of living". Furthermore, aspects could be explicitly or implicitly expressed. A statement that it is not safe to live in Nigeria was implicitly expressed because the yardstick used to arrive at that conclusion was not explicitly stated. Though there may be unrest in some areas of some states in the country, this could be less than 5% of the state's geographical location. However, the opinion that the cost of living is low was explicitly expressed.

As a result, aspect-level sentiment analysis is a difficult undertaking because explicit aspects are easily defined whereas implicit aspects are difficult to identify [13]. Additionally, aspect detection and sentiment classification are two subtasks that are part of aspect-level sentiment analysis. While aspect detection primarily concentrates on extracting aspects, classification of sentiments focuses on classifying the extracted sentiments into appropriate polarity (i.e., positive, neutral, or negative); this can be accomplished using knowledge-based techniques (lexicon approach), machine learning techniques (statistical approach), and hybrid methods.

Knowledge-based techniques use explicitly defined emotional words (sentiment lexicon) to classify a text. A catalog of sentiment words, multi-word phrases, idioms or expressions, rules of opinion, and parse trees are all included in a sentiment lexicon. These are used to identify each part of a sentence's sentimental orientation. Sentiment words could be compiled using a corpus-based approach, dictionary-based approach, or manually (which is labor-intensive thereby rarely used). Available sentiment lexicon includes WordNet-Affect, SentiWordNet, WordStat Sentiment Dictionary, SentiStrength lexicon, EffectWordNet, linguistic annotation scheme, affective lexicon, SenticNet, etc. Consequently, the sources of the sentiment words used determine how accurate this technique is [14]. Additionally, sentiments can be retrieved using Machine Learning (ML) approach, the unsupervised and supervised. The labeling and utilization of aspects and non-aspects for training are necessary for supervised machine learning approaches. They construct a mathematical model of both the inputs and expected outputs on a set of observations or data.

The supervised learning algorithms learn through iterative optimization of a goal function - a function that can be used to predict new inputs [15]. Generally, supervised learning algorithms can be classified into probabilistic, decision-tree, linear and rule-based classifiers. The hybrid approach employs both techniques for sentiment extraction. Even though CNN classifier has reportedly been shown to miss certain legitimate aspect phrases, a rule-based strategy might improve its effectiveness. Although several works on sentiment analysis of tweets in English and other languages have been reported in the literature, existing models have not been trained to handle a special variant of English language called "Pidgin language"; this is a blend of English and local words that is common among African countries. Since most tweets are not written in official English, it is permitted to use acronyms, symbols, emojis', and other abbreviations that vary from country to country. As a result, this work investigates the application of rule based and deep machine learning algorithms for sentiment analysis and categorization based on aspect. The rule-based technique was utilized to find and extract aspect-based sentiments from preprocessed tweets with a focus on Twitter data available on the Nigerian Twitter Space, while the CNN deep learning technique was used to classify sentiment polarity.

This research article is an extended version of the preliminary study presented at an international conference [2]. However, this extended version provides more detailed background information about aspect sentiment analysis, an overview of techniques employed in existing works was added, and a total of 26401 tweets were used as against 26378 previously used in the conference article, therefore, an improved training and testing results were obtained. Also, a comparison of the results obtained from CNN model and the proposed CNN with rule-based technique was provided. A comparison of results obtained from the proposed technique and results available in existing works was also added in this extended version. Most importantly, the contributions of the study presented here are the following:

(a) For the aspect-based sentiment extraction task, this study further affirmed the prowess of the rule-based technique in identifying both local and sequential features that may be hidden in a text. Due to the uniqueness of the tweet data available on the Nigerian twitter space, this work demonstrated how rule-based techniques can be used to extract aspect-based sentiments by studying the unique tweet composition, structure, grammatical, lexical and semantics of tweets.

(b) Regarding the task of classifying sentiment polarity, a CNN deep learning technique that was optimized with existing Word2Vec and Glove embedding models was used for sentiment polarity classification. This involved normalizing the polarity values obtained from CNN, Word2Vec and Glove embedding models by computing the mean and median polarity values generated by the three models. Afterwards, the mean of the two resulting polarity values was computed. This yielded an improved result compared to when only the CNN model was used.

(c) A web platform that Governments and interested individuals can use to examine expressed sentiments in tweets was developed.

The remainder of this article is structured as follows: an overview of related works is provided in Section 2. The methodology behind the proposed model is presented in Section 3, and the results are given in Section 4. The study's conclusion and future works are covered in Section 5.

2. Related works

The literature reviewed in this section is in four categories: The first category summarizes works that have focused on aspect level using one technique or another. The second category focused on works that employed deep learning on sentiment analysis task. The third subsection focused on related works that employed rule-based techniques for their sentiment analysis task, while the last subsection reviewed articles that extracted sentiments from tweet data.

2.1. Aspect level sentiment analysis

Several researches have been carried out towards developing an efficient Aspect Based Sentiment Analysis (ABSA) system. An overview of techniques employed in ABSA and the results obtained are provided in this section. The research in Abdelgwad *et al.* [1] introduced a DL model for the analysis of sentiment in the Arabic language. The suggested model combined CNN, a bidirectional gated recurrent unit, and a conditional random field for the aspect extraction task while an interactive attention network was used for the polarity extraction. The word embedding vectors and character level word encoding from CNN were combined to create the context information for each word under consideration. When the Arabic hotel reviews dataset was used to assess the suggested models, an F1 score of 72.83% was achieved compared to 69.44% obtained from CNN-BGRU-CRF that used word2vec and 70.67% obtained from CNNBGRU-CRF that used fastText word embedding. Also, the aspect extraction model yielded an accuracy of 75.8% while the polarity extraction model yielded an accuracy of 83.98%.

Al-Dabet *et al.* [16] also proposed two DL models for ABSA prediction. CNN and LSTM were used to extract sequential and local features used for the aspect identification while a combination of bidirectional independent LSTM, GRU's and multiple attention layers were used for the polarity classification. These algorithms were used to overcome the limitations of RNN. SemEval 2016 Arabic dataset was used to evaluate the proposed models. Accuracy of 58.08% was recorded for the sentiment identification model while 87.31% was recorded for the polarity classification model. To employ contextual affective knowledge present in words for aspect sentiment extraction, a SenticNet graph convolutional network was introduced by Liang *et al.* [17]. The graph-based model was used to extract the affective dependencies of sentiment aspects. When the suggested model was tested using SemEval 2016 datasets from the restaurant domain, accuracy and an F1score of 91.97% and 79.56%, respectively, were attained. To improve the challenges of Latent Dirichlet Allocation (LDA), which is ineffective for short sentences due to data sparsity issues and a lack of co-occurrence patterns, a sentence segment LDA technique that is appropriate for extracting features from both short and lengthy sentences was proposed by Ozyurt *et al.* [18]. The topic modeling method divides a sentence into sections, each of which addresses a distinct element. Afterwards, words that are not associated with aspects are removed and those related are grouped under the same category. F1-Score of 82.39% and precision of 81.36% were achieved when the proposed model was evaluated with the product reviews dataset. Additionally, Yadav *et al.* [19] developed a method to overcome this restriction of the ABSA system, which uses position index sequence to calculate the association between aspect words and other contextual words. The need for positional embedding was eliminated while a lexicon replacement and masking preprocessing techniques were employed to retrieve information about the aspects in the sentence. The masking preprocessing technique yielded a better accuracy of 88.65% when compared to an accuracy of 87.38% obtained by the lexicon replacement technique. The overall system yielded an accuracy of 89.30% which was better than existing techniques.

In the work of Ray and Chakrabarti [13] suggested a roughly related project that used rule- and CNN-based deep learning algorithms for aspect-based sentiment analysis of reviews of products, movies, and restaurants. Skip-Gram Model was used for the word embedding, while SenticNet was used as the concept-level opinion lexicon. Almost Positive, Positive, Very Positive, Almost Negative, Negative, Very Negative, and Neutral are the seven-point sentiment score categories that were also suggested. With CNN, the suggested CNN + Rule-based technique, and CoreNLP + Rule-based, overall accuracy of 0.75, 0.80, and 0.87, respectively, were attained. This demonstrated how the classification accuracy of CNN might be improved by combining it with a rule-based approach. Wu *et al.* [20] reported that existing neural network models do not perform well when there are limited training data. For ABSA, they subsequently suggested the Multisource Textual Knowledge Fusing Network (MTKFN). A dataset with conjunction relations at the sentence level and document level sentiment labels were utilized to train the proposed model, which combined sentiment and structural information for sentiment analysis. Bidirectional Long Short-Term Memory (BiLSTM) networks were utilized to create the sentiment labels at the document level, while Rhetorical Structure Theory was used to segment the clauses at the sentence level. In order to better forecast the sentiment polarities of aspect phrases, this was done. The proposed model yielded a better accuracy and F1 score when compared to other techniques.

Aspect position and Entity-oriented Knowledge Convolutional Graph (APEKCG) were suggested by Ahmed *et al.* [21] to address limitations identified in current studies which overlook the non-explicit arrangement of aspect-category information inside the context. The Entity Oriented Knowledge Dependency Convolutional graph (EKDCG) and the Aspect Position-Aware (APA) make up its two modules. The APA selects the most relevant features and the EKDCG is built in three phases and the result of the modules is fed into the classifier. The efficacy of the suggested models was tested on SemEval-15&16, MAMS, Restaurant, Laptop and AWARE datasets. The accuracy of the suggested method exceeded that of the existing models with 90.22%, 89.13%,89.02%, 84.32% and 79.64% respectively. Similarly, Bharathi *et al.* [23] employed deep learning ideas to construct ABSA models for the identification of sentiments in experience, service and product aspects. Word2Vec was utilized for the extraction of the most informative features after the preprocessing of the datasets. BiLSTM, CNN and Hopfield Network were employed for classification of the extracted features. The efficacy of the DL models was tested on AWARE datasets and CNN outmatched the other models with 92.08% accuracy, 88.44% precision and 86.20% recall.

Gu *et al.* [22] suggested Syntax-Aware and Graph Convolutional Networks (SAGCN) which is a solution to the drawbacks experienced by existing models such as graph convolutional networks which do not fully utilize the information contained in a sentence's unique aspects, and they do not take into account the improvement of the model through the addition of external general sentiment knowledge. To improve the model's ability to recognize sentiment information, aspect-specific elements were first integrated into contextual information and then external sentiment knowledge is applied. To extract the sentence's semantic information, a multi-head self-attention mechanism and a Point-wise Convolutional Transformer (PCT) were used. This model incorporates semantic and syntactic aspects of sentence-specific parts of sentiment analysis. The suggested model significantly outperforms the benchmark models examined by testing it on three (3) ABSA benchmark datasets.

2.2. *Sentiment analysis using rule based techniques*

There is no general rule that specifies how tweets should be written; in fact, informal languages, symbols, and emoticons can be used in their constructs. This makes rule-based techniques that use grammatical rules to model inherent features that may be present in tweets suitable for sentiment analysis. They are majorly used to capture features that other machine learning techniques fail to capture. For instance, Siddiqua *et al.* [24] presented a rule-based classifier to improve the Naive Bayes classifier's capacity for sentiment categorization. The emoticon and sentiment word occurrences in the text data were extracted using the rules that were created. The Nave Bayes classifier was then trained using the sentiment lexicons that were created. The suggested model was trained using the 1.6 million tweet Stanford Twitter Sentiment140 dataset, and it was tested with 359 tweets. Recall, F1, and accuracy were all attained at 91.76%, 86.08%, and 84.96%, respectively. Sentiment analysis was conducted on an annotated corpus of review texts gathered from online hotel reservation sites to investigate constraints between efficiency and accuracy in the work of Braoudaki *et al.* [25]. The Deep learning models were trained using preprocessed reviews that have sentiment tag added to words and expressions in the text based on linguistic rules. The tags were either negative or positive indicating the sentiment polarity. The suggested approach utilized four deep learning architectures that take as training input either the annotation of the text alone or the review tests in addition to certain details on the tag annotation.

2.3. Sentiment analysis of tweet data

SMNs have been continuously embraced as a convenient platform for individuals to express their opinions about any subject of discussion. Therefore, they have become a rich source of opinion mining that has attracted the attention of researchers and has been used for several predictions. For instance, Birjali *et al.* [26] employed a machine learning technique for suicide related sentiment prediction from Twitter data. An English vocabulary of words that points at suicide was manually built, and the WordNet embedding model was used for the semantic analysis of the prediction system. Afterwards, Sequential Minimal Optimization (SMO), Naive Bayes, and CART classifiers were used for the classification task. When 892 tweets collected with keywords available in the manually compiled vocabulary were used to evaluate the system's performance, the highest precisions of 89.5%, 87.50%, and 83.1% were achieved with SMO, CART, and Naïve Bayes, respectively. A FI score of 85.7%, 89.3%, and 82.9% was also achieved with SMO, CART, and Naïve Bayes, respectively. A real-time sentiment prediction system was proposed by Yoo *et al.* [27]. CNN was used to build the sentiment analysis model, while LSTM was employed for the prediction model. Word2vec algorithm was used for the word embedding. A better accuracy (84.56%), precision, recall, and F1 score was recorded when compared to Naïve Bayes, SVM, and random forest classifiers. The authors in Sailunaz and Alhajj [28] proposed a recommender system that used sentiments and emotions expressed in tweets and replies to tweets. An influence score that uses agreement, sentiment and emotion scores computed from the tweets and their replies was used to generate the final recommendation. With the Naïve Bayes algorithm as the classifier, an accuracy of 66.86% was recorded when compared to 23.32% and 55.23% achieve with SVM and random forest.

3. Materials and methods

Twitter data is sensitive, therefore rule-based and DL methods were used for aspect detection, extraction, and classification to guarantee that no aspect of a tweet is overlooked. For aspect sentiment extraction and detection, the rule-based method was used, and for sentiment polarity classification, the deep learning method. The processes used include model development, model evaluation, and model data display. Tweet data collection and preprocessing are also included.

3.1. Tweet data collection and preprocessing

The Kaggle Sentiment140 event made 1.6 million annotated tweets about security were used to train the proposed sentiment analysis system while 26401 tweets about the security situation of Nigeria were used to assess the efficacy of the suggested sentiment analysis model. The tweets were collected between 6th of December, 2019 and 11th of April, 2020 due to the lingering security situation in the country and is publicly available online [29]. During the data collection, it was observed that Twitter API cannot be used to collect tweets that are more than a week; therefore, a site scraping script was used to complement the tweets obtained via the Twitter API. Furthermore, there were redundant and duplicated tweets among the tweets retrieved. Also, some tweets had several URL links, @mentions, usernames, hashtags, numbers, acronyms, superfluous symbols, tags, tabs, tweet id, user information, language attributes, emoticons, whitespace, and media assets like photos and videos. These were all undesirable traits that were eliminated during the preprocessing phase. This was accomplished using a script that could recognize and pinpoint regular expressions and/or patterns that matched elements that were considered to be superfluous and required removal. The following pseudocode was used to carry out the preprocessing task with the aid of a builtin Python RegEx library:

Input: Tweet data
Output: Preprocessed Tweets
for each statement in tweet do

1. Refined_tweet = RegEx.remove(r'[∧a-zA-Z]', statement)
2. Refined_tweet = RegEx.remove(r'\s+[a-zA-z]\s+', statement)
3. Refined_tweet = RegEx.remove(r'\s+', statement)
4. Refined_tweet = RegEx.remove(r '< [∧>]+>', statement)
5. Refined_tweet = RegEx.remove(r 'http\S+', statement)
6. Refined_tweet = RegEx.remove(r 'pic.\S+', statement)
7. Refined_tweet = RegEx.remove(r'@\s\w+', statement)
8. Refined_tweet = statement.trim (),

where + denotes strings with a pattern that appears one or more times, S denotes strings with no whitespace characters, w denotes matches with any word characters (letters from A to Z, numbers from 0 to 9, and the underscore character), and s denotes matches with whitespace characters in strings. URLs, or strings that start with image, are also included in step B. Twitter was removed, @mentions in the tweets were removed, HTML tag contents were removed, symbols, punctuation, and other non-alphabetical elements were removed, single characters were removed, multiple concurrent spaces were removed, and spaces and white spaces were removed from the input tweet in steps *c* through *h*.

3.2. Model development

The framework of the proposed aspect-based sentiment extraction and classification model using rule-based and DL model is shown in Figure 1. It has two modules:

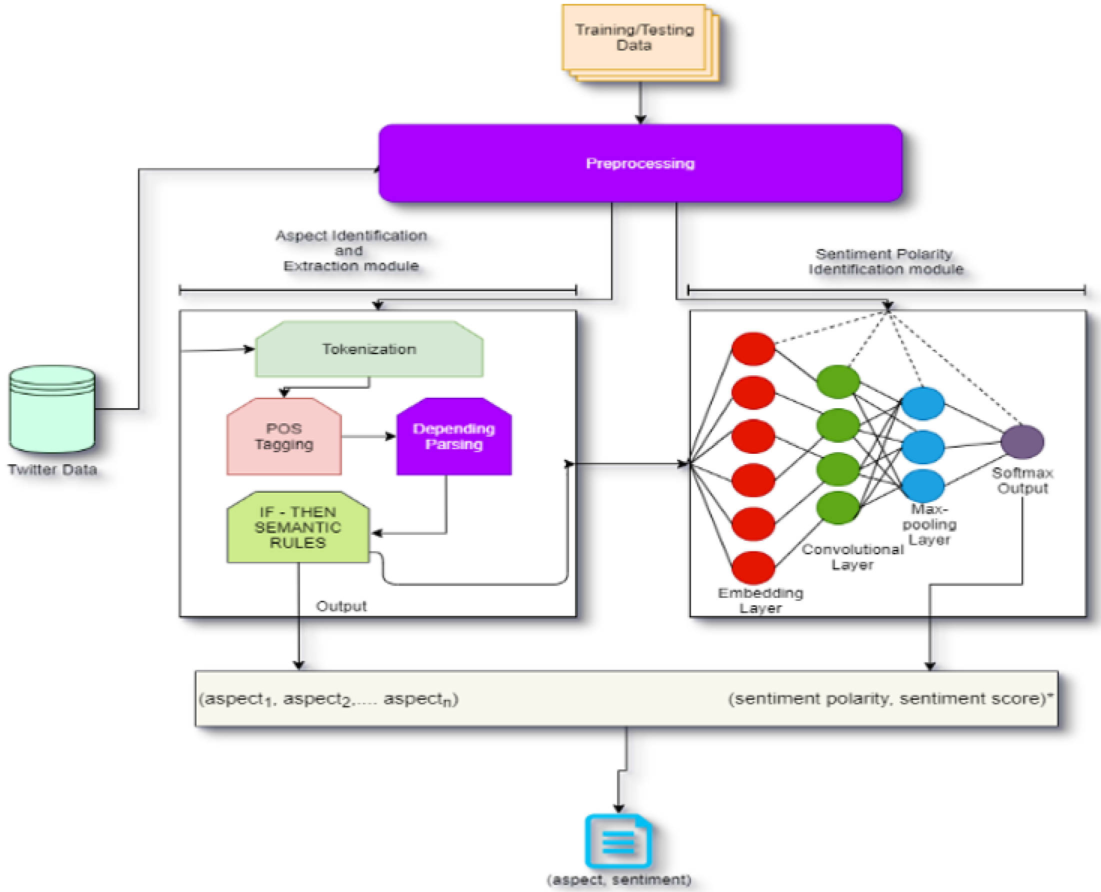

Figure 1. Framework of the research.

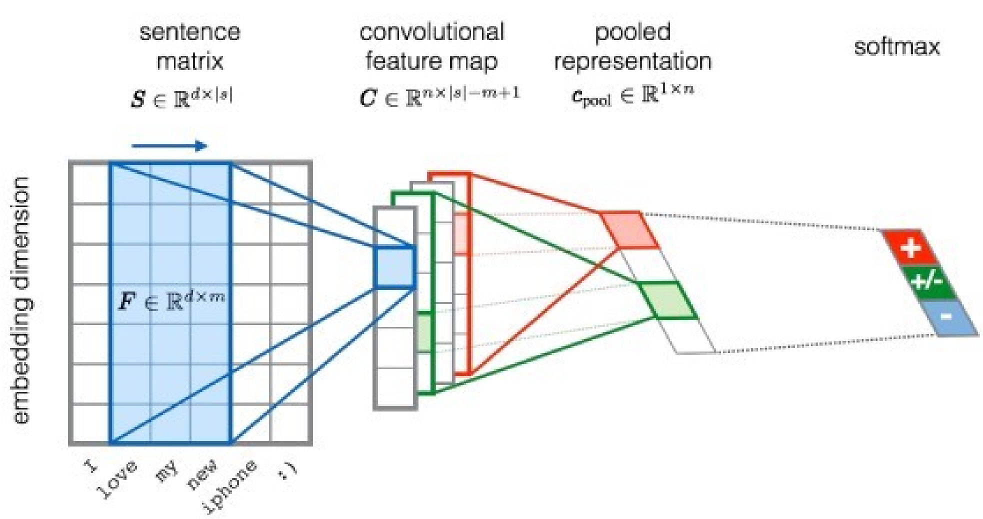

Figure 2. The CNN based sentiment polarity classification module.

1. Aspect identification and extraction module
2. Sentiment polarity classification module

3.2.1. *Aspect identification and extraction module*

The components for detecting and extracting aspect terms from sentences are part of the aspect identification and extraction module. The relevant elements were extracted from the

Table 1. Aspect characteristics of extracted tweets.

Item	Value
Distinctive Appearing Aspects	79,318
Total number of aspects that occur more than once	17,147
Most Prominent Aspect	Nigeria
Number of times an aspect appeared	15,702
Highest degree of polarity connected to an aspect	Negative
Polarity breakdown for aspect	Negative:7547, Neutral:4566, Positive: 1549
Total number of distinct aspects	62,171
Average aspect frequency per tweet	7.2092
Tweets having aspect appearances that are below average	11,156
Number of tweets containing a single aspect	385
Breakdown of the polarity of tweets with a single aspect	Negative:167, Neutral:138, Positive: 80

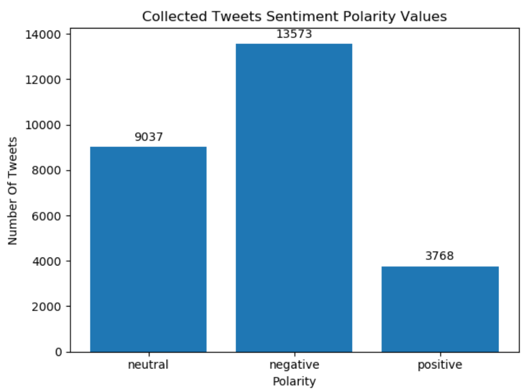

Figure 3. Value distribution of sentiment polarity over the collected tweets.

tweets using the rule-based approach, which applies natural language processing rules. The opinion word lexicon was created using SentiWordNet, a sentiment lexicon in which words are given numerical sentiment values. After removing terms with no sentimental significance from the 26,401 tweets that were taken into consideration, 51,949 words remained. The aspect extraction was implemented using the Spacy Python Library. This involves:

1. Tokenization: This method divides sentences into smaller, analyzed units called tokens.
2. Part-Of-Speech (POS) Tagging: In this stage, tokens are identified and categorized according to the part of speech to which they belong.
3. Dependency Parsing: Following partof-speech tagging, a variety of aspects-dependencies were gathered and specified using syntactic tokens.
4. Putting IF-THEN rules into action: Aspects are extracted based on recognizable semantic links between tokens using rules in the "IF-THEN" format. Studying sentence structure, grammatical and lexical syntax, and sentence semantics allows one to deduce these rules.

3.2.2. *Sentiment polarity classification module*

Based on the Keras deep-learning package, the sentiment polarity identification and classification module used a six-layer CNN. The network comprises of a fully linked layer with Softmax output, an embedding layer, two convolutional layers, and two max-pooling layers. Figure 2 illustrates this.

The Embedding Layer. There is a need to encode every word or phrase from the input sentence as a word vector; this vector representation of terms in a sentence is word embedding. It is an input layer to the network that provides a means of mapping real-number vector representation to terms or phrases from a vocabulary. The technique often calls for the transformation of a high-dimensional, sparse vector space into a smaller, more compact vector space.

The embedding vector's many dimensions each correspond to a word's latent feature. The resulting word vector matrix for a sentence S of length I_R, derived from a corpus vocabulary of size V_R, is provided as:

$$W_R \in \mathbb{R}^{d\times|V|}.$$

The ith word in the phrase was then transformed using the matrix-vector (dot) product into an n-dimensional vector w_i, so that:

$$w_i = W x_i, \tag{1}$$

where x_i is the ith word's encoded description. The deep learning network can use this embedding matrix as a feature vector of learnable parameters.

Given a total of 42,000 trainable parameters in just the embedding layer, the input tensor for the model's embedding layer has a shape of $(300, 140)$, where 300 represents the word embedding dimension and 140 represents the selected sentence length. The first convolutional layer has 128 filters with a kernel size of 2, while the second convolutional layer has a filter size of 256 and a kernel size of 3. There are 128 units in the fully connected layer. With the exception of the softmax neuron, which has a sigmoid activation function, all layers employ the ReLU activation function. The binary cross-entropy function was used as the loss function.

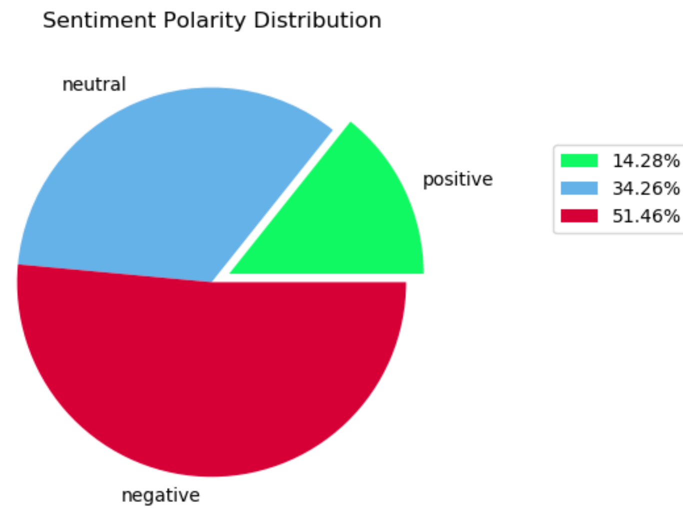

Figure 4. Percentage distribution of sentiment polarity in collected tweets.

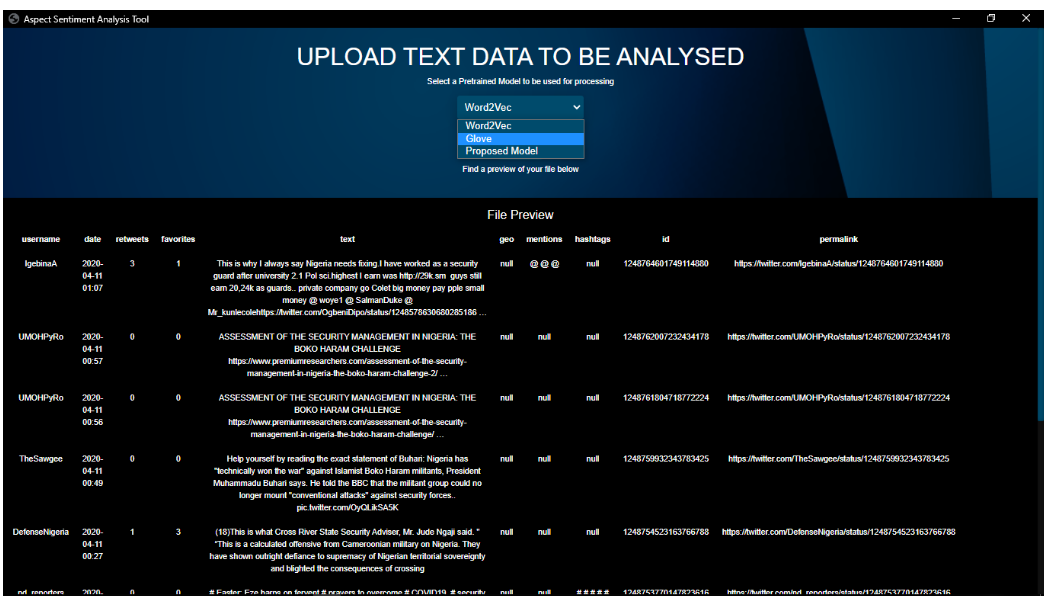

Figure 5. Tweet sentiment prediction interface.

Table 2. Performance of the proposed model during training

Models	Precision (%)	Accuracy (%)	F1-Score (%)	Recall (%)
Word2vec Model	85.70	854.21	84.75	84.46
Glove Model	86.25	84.70	85.00	80.35
Proposed CNN + Rule based Model	88.45	88.00	87.95	87.15

The Convolutional Layer. This layer accepts the embedding matrix as input. The embedding matrix vectors are then subjected to the convolution procedure to produce additional features. This is illustrated in Eq. (2) as:

$$(f * g)(t) \triangleq \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau, \quad (2)$$

where τ is the difference between the point for which the convolution is being calculated and the t at a specific point in the

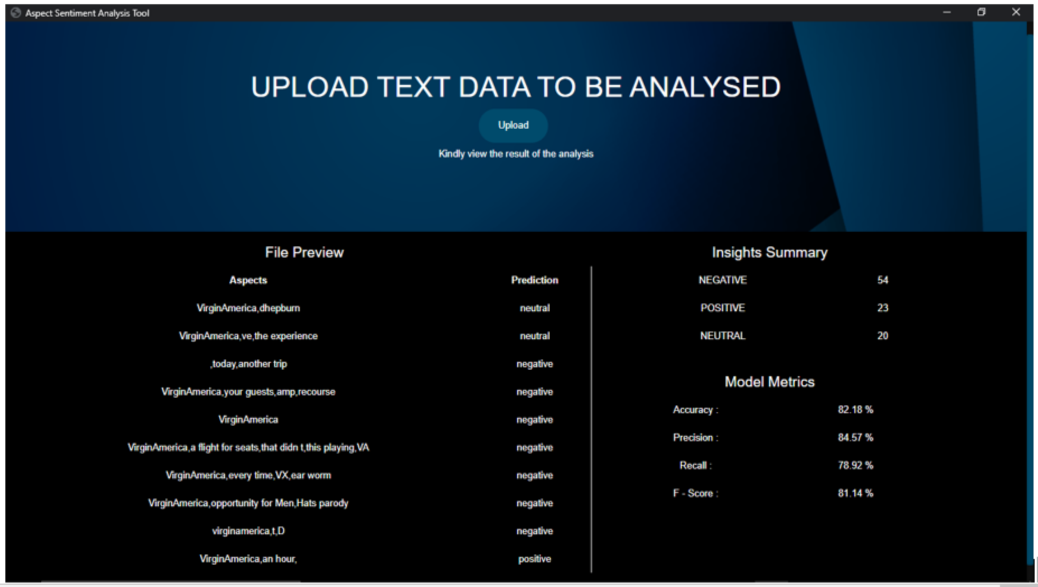

Figure 6. Tweet sentiment analyis interface.

Table 3. Performance of the proposed model during testing.

Models	Precision (%)	Accuracy (%)	F1-Score (%)	Recall (%)
Word2vec Model	85.15	83.25	82.10	82.15
Glove Model	85.50	82.68	81.86	79.25
CNN Model	80.75	80.26	80.89	81.15
CNN+ Rule based Model	88.15	89.31	87.62	88.21

Table 4. Comparism with results from exisiting studies.

Authors	Methodology	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)
Hu *et al.* [14]	CNN + REL + LEX	88.0	-	-	-
Hu *et al.* [15]	CNN + First-Order Logic Rules	89.3	-	-	-
Poria *et al.* [30]	CNN LP + CNN	- -	85.01 86.10	87.42 88.27	86.20 87.17
Yin *et al.* [31]	CNN + WLE + SLE	87.1	-	-	-
Li *et al.* [32]	CNN + N Gram filter	89.0	-	-	-
Zhao *et al.* [33]	CNN + Capsule Model	87.3	-	-	-
Zhang *et al.* [34]	CNN+Knowledge Rules	89.5	-	-	-
Ray & Chakrabarti [13]	CoreNLP + Rule-based CNN	0.75 0.80	- -	- -	- -
Abdelgwad *et al.* [1]	CNN-BGRU-CRF (word2vec)CNN-BGRU-CRF (fastText)	- -	- -	- -	69.44 70.67
Al-dabet *et al.* [16]	CNN-LSTM	87.31			58.05
Proposed CNN + Rule based Technique		89.31	88.21	88.15	87.62

Where BGRU-CRF: Bidirectional Gated Recurrent Units-Conditional Random Field, LP: Linguistic Patterns WLE: Word Level Embedding, SLE: Sentence Level Embedding, LEX: Lexicon.

period integral, and f and g are arbitrary denotations of the feature vectors at point t. A new feature f is the result of the convolution process; it is derived from a window of words wi-r:i+r. After applying the filter p Rh * d, the window of words was recovered from a window slice of the input sentence h = 2r + 1. The new feature was computed using Eq. (3):

$$f_i = g(wi - r : i + r .p), \tag{3}$$

where the activation function g is non-linear. Additionally, feature maps were created using the filter for all potential windows of the encoded maps input phrase illustrated in Eq. (4):

$$f = [f_1 \ , \ f_2, \ f_3, \ldots . f_n]. \tag{4}$$

The Pooling Layer. Getting the most significant feature with the highest value from each feature map matrix is the core function of this layer. As a result, various filters with various window sizes are employed, but only one feature is taken from each filter. This is accomplished by combining data from the complex filters and producing their representation in such a way that:

$$\hat{f} = \max([f_1 \ , \ f_2, \ f_3, \ldots . f_n]), \tag{5}$$

where $\hat{f}$ is the feature that has the highest value. This is accomplished by applying max-over-time pooling to each fixed-size feature vector $v_i \in \mathbb{R}^d$ for every i-th feature map matrix.

Dropout Regularization. The enormous number of parameters that must be learned is one of the causes of deep neural network overfitting (Ray & Chakrabarti [13]). In order to prevent overfitting in the model, dropout regularization is typically suggested. This method randomly modifies a portion of a layer's neurons, which prevents co-adaptation and forces each neuron to learn useful characteristics on its own.

Fully-connected layer. This layer receives the features produced by the dropout regularization layer using Eq. (6) such that:

$$G = \sigma(w * \hat{f} + b). \tag{6}$$

The weight of the matrix is denoted by w, the rectified linear activation function is denoted by, and b denotes the bias, which is applied to a neural network as a threshold for learning patterns.

The SoftMax layer. This layer uses the output from the fully connected layer to create the probability distribution of the class labels, which results in:

$$y = argmaxP\,(y = j\,|\,x,\ w,\ a). \tag{7}$$

$$\therefore y = argmax(\frac{e^{x.w_j+\ a_j}}{\sum_k x.e^{x.w_k+a_j}}). \tag{8}$$

While the bias for the class is represented by a j, the weights of the entities that are members of class j are represented by w j. Using the probability function (y=argmaxP()), a discrete probability distribution of the networks' classification results was created. The model parameters were then updated or changed in accordance with the training data's actual classification label using the back propagation procedure. Before the model training procedure is finished, these forward and backward propagation steps are iterated a predetermined number of times. The trained model will be able to forecast the sentiment polarity of the tweets after the model training process is over. These forecasts produce a probability distribution with a range of 0 to 1. These forecasts produce a probability distribution with a range of 0 to 1. Sentiment polarity levels tending towards zero are more negative, and values tending towards one are the opposite. As a sentiment score, this probability distribution would be used. Since the output ranges from 0 to 1, 0.5 would be a logical positive polarity threshold. However, a threshold of 0.18 was imposed on values for neutral tweets due to the complexity of the tweets, as seen by their constructions and semantic components.

3.3. Performance evaluation

Accuracy, precision, recall, and F1 score were used to assess the performance of the proposed rule and deep learning-based model for sentiment analysis in the following ways:

1. Accuracy represents the percentage of properly classified tweets relative to the total number of tweets utilized for testing. When the distribution of the class is nearly equal, accuracy is more effective. For the accuracy, Eq. (9) was utilized.

$$Accuracy = \frac{TN + TP}{TP + TN + FP + FN}. \tag{9}$$

2. Precision: this is also called positive predictive value. It calculates the proportion of successfully categorized tweets that were actually categorized. Eq. (10) was used to compute this:

$$Precision = \frac{TP}{TP + FP}. \tag{10}$$

3. Recall also known as sensitivity assesses how thorough the methods were in classifying a tweet's polarity. It was computed using Eq. (11):

$$Recall = \frac{TP}{TP + FN}. \tag{11}$$

4. F1-Score: This extrapolation metric combines recall and precision by taking the harmonic mean of the two. It was computed using Eq. (12):

$$F1\,Score = \frac{2 * Precision * Recall}{Precision + Recall}, \tag{12}$$

where TP connotes True Positive and it was used to evaluate positive tweets that are rightly classified as positive, TN means True Negative which estimates negative tweets that are rightly classified to be negative, FN means False Negative which measures positive tweets that are wrongly classified to be negative while FP known as False Positive measures negative tweets that are wrongly classified as positive.

4. Results and discussion

As described in the preceding section, a web scraping script and the Twitter API were utilized to collect the 26,401 tweets that were used to test the proposed model. Table 1 provides a summary of the elements that were taken from the tweets that were recorded.

4.1. Sentiment polarity module

The retrieved sentiments were divided into positive, neutral, and negative polarity using the CNN model. However, several disparities were noted as a result of the compositional changes between the training and testing sets of twitter data. To support the suggested model, pre-trained word embedding that already exist were used. The final polarity score of a tweet was calculated as a normalized value of all three model's values after applying all three models. The mean of the two polarity values that arise from this normalization was calculated after computing the mean and median polarity values produced by the three models. Assume that the ultimate polarity score of the model is represented by the letter f, and that the scores of the models are p, w, and g, respectively. In that situation, Eq. (13) was used to compute the final polarity:

$$f = \text{mean}(p, w, g) + \text{median}(p, w, g). \tag{13}$$

This addresses outliers by taking into account the scores of all three models and provides a number that is relatively close to a range (0.046) of values shared by two models. For instance, the word2vec model and glove both receive scores of 0.264 for a tweet with a predicted model score of 0.565. The polarity of the first model is neutral, and the majority polarity of the categorized tweet would be negative (using the previously specified scale for classifying polarity), as negative is the value that the two models share. Given the first model's score, the normalized value of 0.3115 still falls on the negative side but is closer to neutral. Figures 3 and 4 depict the distribution of polarity across the gathered tweets using this method. As anticipated, the number of tweets labeled negatively exceeds those of other categories. It was unexpected how many neutral tweets there were in total. The neutral tweets, on the other hand, were found to be statements, suggestions, and questions when the testing tweets were evaluated.

4.2. Web-based graphical user interface

A web-based Graphic User Interface (GUI), depicted in Figure 5 was created for simple tweet sentiment prediction. This makes interacting with the sentiment prediction system simpler for even a beginner. The spreadsheet file may be readily uploaded via the GUI once the tweets to be studied have been extracted into it. This shows a summary of the file that was submitted, including the date, username, mentions, hashtags, tweet IDs, and permalinks, as well as the number of retweets and the text that was actually tweeted. The area of this interface where the user can select the preferred pre-trained model to be processed is an important feature. The Glovec, word2vec, or the suggested model might be this. The outcomes will be displayed in the interface depicted in Figure 6 once the preferred pre-trained model has been chosen. The aspects taken from each tweet and their corresponding polarities, insights summaries, and performance evaluation results in terms of accuracy, precision, recall, and F-score are presented as the categorization results.

4.3. Model evaluation

Modern word2vec and Glove word embedding models were contrasted with the suggested rule-based and machine learning model. In all performance criteria, the proposed model outperforms the baseline pre-trained embedding models thanks to its locally trained word embedding during the training stage. Table 2 displays the training results. As shown in Table 3, the suggested model performed better than the current models during testing in terms of accuracy, recall, and F-score.

Several authors have employed different techniques to improve the perforomance of CNN for sentiment analysis and polarity prediction. The results obtained from these studies are presented in Table 4. It will be obsered that the results obtained from our proposed technique is at par with what was reported in existing studies. This showed that combining CNN model with rule based technique yielded a more accurate polarity prediction. This is simply because the rule based technique could capture non-English words present in the tweets.

5. Conclusion

The rule- and CNN-based deep learning technique has been fully described in this study as a novel method for gathering and assessing thoughts expressed in tweets. While the CNN-based deep learning technique was used to identify and categorize the sentiment polarity of the tweets, the rule-based technique was utilized to discover and extract sentiments from the processed tweets. The 26,401 security-related tweets that were retrieved on twitter using security and Nigeria as keywords were used to evaluate the performance of the proposed model. 7547 negative elements in all, 4566 neutral, and 1549 good aspects were all that could be gleaned from the tweets. In terms of accuracy, recall, precision, and F-score, the suggested model outperformed the most recent word2vec and Glove models. A web-based graphic user interface was additionally created to simplify the sentiment analysis process. Users can import tweets that have been saved in spreadsheet forms. Following the selection of the preferred word embedding model, the statistics of the positive, neutral, and negative polarity, accuracy, recall, precision, and F1 score values for the uploaded tweets will be provided.

Data Availability

We do not have any research data outside the submitted manuscript file.

References

[1] M. Abdelgwad, T. Hassan, A. Soliman, A. I. Taloba & M. Fawzy, "Arabic aspect-based sentiment analysis using bidirectional GRU-based models", Journal of King Saud University – Computer Information Science **34** (2021) 6652. https://doi.org/10.1016/j.jksuci.2021.08.030.

[2] O. N. Akande, E. S. Nnaemeka, O. C. Abikoye, H. B. Akande, A. Balogun & J. Ayoola, *TWEERIFY: A web-based sentiment analysis system using rule and deep learning techniques*, In Proceedings of International Conference on Computational Intelligence and Data Engineering, ICCIDE, Singapore: Springer Nature Singapore, 2021, pp. 75–87. https://doi.org/10.1007/978-981-16-7182-1_7.

[3] O. Edeama, "The Role of Social Media and Social Networking in Information Service Provision: A Practical View", in *Social Media Strategies for Dynamic Library Service Development*, 2015, pp. 124–139. https://doi.org/10.4018/978-1-4666-7415-8.ch007.

[4] P. Borah, "Facebook Use in the 2012 USA Presidential Campaign", in *Social Media in Politics: Case Studies on the Political Power of Social Media*, 2014, pp. 201–211. https://doi.org/10.1007/978-3-319-04666-2.

[5] S. Bute, "The role of social media in mobilizing people for riots and revolutions", in *Social Media in Politics: Case Studies on the Political Power of Social Media*, 2014, pp. 213–223. https://doi.org/10.1007/978-3-319-04666-2.

[6] D. Irak, "A close-knit bunch: Political concentration in Turkey's Anadolu Agency through Twitter interactions", Turkish Studies **17** (2016) 336. https://doi.org/10.1080/14683849.2016.1138287.

[7] B. Bansal & S. Srivastava, "On predicting elections with hybrid topic-based sentiment analysis of tweets", Procedia Computer Science **135** (2018) 346. https://doi.org/10.1016/j.procs.2018.08.183.

[8] T. R. White & T. Anderson, "The use of Twitter in the 2013 Italian political election", in *Social Media in Politics: Case Studies on the Political Power of Social Media*, 2014, pp. 63–77. https://doi.org/10.1007/978-3-319-04666-2.

[9] P. Borah, "Facebook use in the 2012 USA Presidential Campaign", in *Social Media in Politics: Case Studies on the Political Power of Social Media*, 2014, pp. 213–223. https://doi.org/10.1007/978-3-319-04666-2.

[10] E. Bulut & E. Yörük, "Digital populism: Trolls and Political Polarization of Twitter in Turkey", International Journal of Communication **11** (2017) 4093. https://ijoc.org/index.php/ijoc/article/view/6702.

[11] P. Chikersal, S. Poria & E. Cambria, *SeNTU: Sentiment analysis of tweets by combining a rule-based classifier with supervised learning*, Proceedings of the 9th international workshop on semantic evaluation, 2015, pp. 647–651. https://aclanthology.org/S15-2108/.

[12] U. U. Ibekwe, U. M. Mbanaso, N. A. Nnanna & U. A. Ibrahim, "A machine learning sentiment classification of factors that shape trust in smart contracts", Journal of the Nigerian Society of Physical Sciences **7** (2025) 2177. https://doi.org/10.46481/jnsps.2025.2177.

[13] P. Ray & A. Chakrabarti, "A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis", *Applied Computing and Informatics* **18** (2020) 163. https://doi.org/10.1016/j.aci.2019.02.002.

[14] Z. Hu, Z. Yang, R. Salakhutdinov & E. Xing, "Deep neural networks with massive learned knowledge", in *International Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 2016, pp. 1670–1679. https://doi.org/10.18653/v1/D16-1173.

[15] Z. Hu, X. Ma, Z. Liu, E. Hovy & E. Xing, "Harnessing deep neural networks with logic rules", in *Fifty-Fourth Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, 2016, pp. 2410–2420. https://doi.org/10.18653/v1/P16-1228.

[16] S. Al-dabet, S. Tedmori & M. Al-smadi, "Enhancing Arabic aspect-based sentiment analysis using deep learning models", Computer Speech & Language **69** (2021) 1. https://doi.org/10.1016/j.csl.2021.101224.

[17] B. Liang, H. Su, L. Gui, E. Cambria & R. Xu, "Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks", Knowledge-Based Systems **235** (2022) 107711. https://doi.org/10.1016/j.knosys.2021.107711.

[18] B. Ozyurt & M. A. Akcayol, "A new topic modeling-based approach for aspect extraction in aspect-based sentiment analysis: SS-LDA", Expert Systems with Applications **168** (2021) 114231. https://doi.org/10.1016/j.eswa.2020.114231.

[19] R. K. Yadav, L. Jiao, M. Goodwin & O. Granmo, "Positionless aspect-based sentiment analysis using attention mechanism", Knowledge-Based Systems **226** (2021) 107136. https://doi.org/10.1016/j.knosys.2021.107136.

[20] K. Ahmed, M. I. Nadeem, Z. Zheng, D. Li, I. Ullah, M. Assam, Y. Y. Ghadi & H. G. Mohamed, "Breaking down linguistic complexities: A structured approach to aspect-based sentiment analysis", Journal of King Saud University - Computer and Information Sciences **35** (2023) 101651. https://doi.org/10.1016/j.jksuci.2023.101651.

[21] J. Chen, H. Fan & W. Wang, "Syntactic and semantic aware graph convolutional network for aspect-based sentiment analysis", IEEE Access **12** (2024) 22500. https://doi.org/10.1109/ACCESS.2024.3364353.

[22] Q. Gu, Z. Wang, H. Zhang, S. Sui & R. Wang, "Aspect-Level Sentiment Analysis Based on Syntax-Aware and Graph Convolutional Networks", Applied Sciences 14 (2024) 729. https://doi.org/10.3390/app14020729.

[23] O. N. Akande, "Tweet data about security situation in Nigeria", Mendeley Data **1** (2022). https://doi.org/10.17632/cvtznrzmx2.1.

[24] S. Al-dabet, S. Tedmori & M. Al-smadi, "Enhancing Arabic aspect-based sentiment analysis using deep learning models",*Computer Speech & Language* **69** (2021) 101224. https://doi.org/10.1016/j.csl.2021.101224.

[25] U. A. Siddiqua, T. Ahsan & A. N. Chy, "Combining a rule-based classifier with weakly supervised learning for Twitter sentiment analysis", in *International Conference on Innovation, Science and Engineering Technology (ICISET)*, 2016, pp. 1–4. https://doi.org/10.1109/ICISET.2016.7856499.

[26] A. Braoudaki, E. Kanellou, C. Kozanitis & P. Fatourou, "Hybrid data-driven and rule-based sentiment analysis on Greek text", Procedia Computer Science **178** (2020) 234. https://doi.org/10.1016/j.procs.2020.11.025.

[27] S. Yoo, J. Song & O. Jeong, "Social media contents-based sentiment analysis and prediction system", Expert Systems with Applications **105** (2018) 102. https://doi.org/10.1016/j.eswa.2018.03.055.

[28] M. Birjali, A. Beni-hssane & M. Erritali, "Machine learning and semantic sentiment analysis-based algorithms for suicide sentiment prediction in social networks", Procedia Computer Science **113** (2017) 65. https://doi.org/10.1016/j.procs.2017.08.290.

[29] K. Sailunaz & R. Alhajj, "Emotion and sentiment analysis from Twitter text", Journal of Computer Science **36** (2019). https://doi.org/10.1016/j.jocs.2019.05.009.

[30] S. Poria, E. Cambria & A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network", Knowledge-Based Systems **108** (2016) 42. https://doi.org/10.1016/j.knosys.2016.06.009.

[31] R. Yin, P. Li & B. Wang, "Sentiment lexical-augmented convolutional neural networks for sentiment analysis", in *International Conference on Data Science Cyberspace (DSC)*, 2017, pp. 630–635.

[32] R. Bharathi, R. Bhavani & R. Priya, "Leveraging deep learning models for automated aspect-based sentiment analysis and classification", International Journal of Electrical and Electronics Engineering **10** (2023) 120. https://doi.org/10.14445/23488379/ijeee-v10i5p111.

[33] M. Yang, W. Zhao, J. Ye, Z. Lei, Z. Zhao & S. Zhang, "Investigating Capsule Networks with Dynamic Routing for Text Classification", in *International Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 3110–3119. https://doi.org/10.18653/v1/d18-1350.

[34] B. Zhang, X. Xu, X. Li, X. Chen, Y. Ye & Z. Wang, "Sentiment analysis through critic learning for optimizing convolutional neural networks with rules", Neurocomputing **356** (2019) 21. https://doi.org/10.1016/j.neucom.2019.04.038.