JOURNAL OF NORTHEASTERN UNIVERSITY
Volume 26 Issue 02, 2023 ISSN: 1005-3026 https://dbdxxb.cn/ Original Research Paper

RUSSIA-UKRAINE CONFLICT TWEETS SENTIMENT ANALYSIS USING BI-DIRECTIONAL LSTM NETWORK FOR POST-TRAUMATIC STRESS DISORDER EARLY DETECTION

Eman Sedqy Shlkamy[1] Khaled Maher [2] and Ahmed Ahmed Hesham Sedky [3]
[1] Department of Information Systems, College of Computing and Information Technology, Arab Academy for Science Technology and Maritime Transport, Alexandria, Egypt
E.shlkamy15020@student.aast.edu

Abstract. Sentiment analysis techniques have a vital role in analyzing people's opinions. The continuous and rapid growth of data posted on social media sites drives people's opinions. However, most of the research focuses on analyzing sentiment to determine how the war will affect the global economy. As a consequence, in international conflict research, national leaders and other powerful figures are typically given more attention than public opinions and emotions. This paper aims to go over sentiment analysis, and focus on analyzing public emotions and opinions during the Russia-Ukraine Conflict to early detect Post-Traumatic Stress Disorder (PTSD) symptoms. This is the first study to provide a model that represents the intention to analyze how the Russian-Ukraine military conflict affected mental health. This can prevent people against mental illnesses and suicide, as well as point the way forward for future study in this field. The method utilized in this study is a single bidirectional LSTM network for English tweet sentiment analysis, with positive, negative, and neutral classifications as a multi-class classification strategy aimed at detecting PTSD indicators. Natural language processing (NLP) is used to extract emotional content from text data via sentiment analysis. By developing a Deep Learning (ML) model using text data, we hope to identify individuals with PTSD via sentiment analysis. Text sentiment analysis was used to learn Natural language processing (NLP) is used to extract emotional content from text data via sentiment analysis. By developing a Deep Learning (ML) model using text data, we hope to identify individuals with PTSD via sentiment analysis. We used one bidirectional long short-term memory (Bi-LSTM) layer in combination with the global Max pooling ID approach and achieved an accuracy of 91.64% based on sentiment analysis of text. In terms of accuracy, the results of the proposed framework outperform earlier state-of-the-art investigations. Creating an early detection model can aid in reducing the symptoms of post-traumatic stress disorder (PTSD).
Keywords: Sentiment Analysis · Neural network · Bidirectional LSTM · Deep neural network· Natural language processing (NLP)· Post-Traumatic Stress Disorder (PTSD)

1 Introduction

War is one of the most mentally stressful events anyone can experience. Even after a conflict has ended, the indirect effects can last for years. As per the International Society for Traumatic Stress Studies, traumatic stress disorder (PTSD) is a common mental health condition that occurs as a result of a stressful experience. A person suffering from post-traumatic stress

Submitted: 28/07/2023 **Accepted**: 18/08/2023

disorder (PTSD) may experience the following symptoms: Night-mares, flashbacks, and returning symptoms that affect feelings and thoughts, such as depression, anger, worry, shame, guilt, and a loss of interest in enjoyable activities and difficulties remembering key things from the past. The symptoms of PTSD and Complex Posttraumatic Stress Disorder (C-PTSD) are very similar. However, (C-PTSD) may also be associated with: Emotional issues, such as difficulty controlling one's emotions, as well as self-esteem issues. PTSD recovery requires an understanding of emotions. The role played by social support in healing and regaining well-being has been demonstrated time and again [16].

Vietnam veterans developed PTSD after the war ended as a result of the psychological effects of war. It was also officially acknowledged that the diagnosis had been made. Psychiatrists were able to recognize victims of war and violence as patients as a result of this construction. As a result of the addition of diagnostic categories, compensation claims were also simplified [17].

The American Psychological Association defines cognitive processing therapy The American Psychological Association defines cognitive processing therapy (CPT) as a cognitive-behavioral therapy designed to treat PTSD and comorbid symptoms. Trauma-related therapy aims to change negative emotions and beliefs. Supporting patients in processing such painful memories and emotions is part of the therapist's job [18]. Therefore, understanding emotions is critical to supporting mental health. In this work, sentiment analysis is used to identify post-traumatic stress disorder (PTSD).Humans have emotions as part of their nature. As a result, it facilitates a better understanding of what humans are experiencing and facilitates a more appropriate response. Understanding emotions is the key to understanding human behavior. When one is aware of emotions, it is easier to decide what one need and want. People can better communicate feelings when they are emotionally aware, prevent or resolve conflicts, and move through difficult emotions more easily [1].

Neurosurgeon Antonio R. and colleagues present conclusive evidence of a fundamental connection between emotion and decision-making. However, Damasio showed that without emotions, we may not be able to make good decisions [2].

During conflict, a lack of social connection can lead to negative emotions such as anxiety and depression. However, when negative emotions are removed, depression, loneliness, social anxiety, and other mental health effects are significantly reduced [3]. Human behavior is largely determined by opinions, as these opinions have a significant impact on how people behave when making decisions [4]. The general definition of "sentiment analysis" is the process of extracting and interpreting subjective and human-relevant information using meaningful techniques such as natural language processing, text analysis, and mining. Various sources of information can be used for sentiment analysis, such as written text or speech, while entities can be events, topics, individuals, etc. Many other tasks can be categorized as sentiment analysis, such as B. Opinion Mining, Sentiment Mining, Sentiment Analysis and Mining [34]

Sentiment and emotions are strongly intertwined. Emotions are a complex group of neurological expressions comprising three components: sentimental occurrences, psychological reactions, and socio-logical reactions [5]. Early recognition of the problem is of utmost importance.

Understanding emotions is one of the most important components of personal growth and development. It is a necessary talent for the imitation of human intelligence. Emotion processing is critical for AI development and the closely linked challenge of polarity detection. Almost all researchers in Natural Language Processing and Computational Linguistics presume that speakers have an emotional value or attitude toward specific aspects of a topic. As a result, the growing fields of affective computing and sentiment analysis, which make use of information retrieval and human-computer interaction [6], have emerged.

The growing need to process opinion-based web and social network content has made sentiment analysis (SA) one of the fastest growing research fields. Due to their applicability and effectiveness, machine learning techniques are commonly used to classify emotions. Furthermore, machine learning models are widely used for sentiment classification due to their high classification accuracy. The use of deep learning in sentiment analysis as a subfield of machine learning has also grown in popularity in recent years. The structure and function of the self-learning human brain provide inspiration for deep learning algorithms. Machine learning and lexicon-based systems are the two main paradigms in the development of sentiment analysis. Although dictionary-based machine learning techniques and traditional machine learning techniques have produced highly accurate results, the feature engineering required to use them is difficult and time-consuming. A rapidly growing subfield of machine learning is deep learning, which is based on artificial neural networks (ANNs). Deep learning algorithms have recently been applied to sentiment analysis and achieved good results. Their ability to automatically learn and build parametric models using datasets makes it easier to create computational models without manual selection of attributes [10]. Surveys of deep learning algorithms for sentiment classification reveal that deep networks perform better. Recurrent neural networks (RNNs) are deep learning neural networks that are specifically built to learn data sequences and are primarily used to classify text [11]. However, it was discovered that RNNs suffer from vanishing gradients as well when dealing with large data sequences [12]. This problem was proposed to be solved by LSTM neural networks, which have proven to be effective in many real-world applications [13].

Bidirectional LSTM and LSTM models can obtain more semantic features, which are helpful for sentiment classification [14], and also show the effectiveness of bidirectional LSTM in performing sequential data models [12]. The proposed architecture based on a single-layer bidirectional LSTM has the advantages of being computationally efficient and suitable for real-time, real-time sentiment analysis applications [15]. The disadvantage of recurrent neural networks is that they have short-term memory, which allows previous information to be stored in the current unit. However, this ability drops off rapidly with longer sequences. To solve this problem, LSTM models with longer memory retention time have been developed. To achieve more meaningful results, bidirectional long-term memory (Bi-LSTM) is introduced, which combines LSTM layers in both directions. As a contribution to this study, this article seeks to gain insight into the early detection of posttraumatic stress disorder (PTSD) in the Russia-Ukraine conflict. They are efficient, highly accurate, and can be used in the classification process for sentiment analysis based on Twitter tweets, using a layer of bidirectional long short-term memory (Bi-LSTM) for early detection of post-traumatic stress disorder (PTSD). We were able to overcome the limitations of sentiment analysis of English-language tweets in support of mental health.The paper is structured into several sections as follows. Section 2 provides an overview of related works in the field, while Section 3 elaborates on the methodology and Proposed Model. The experimental procedures and evaluation are discussed in Section 4, followed by the presentation of results and analysis in Section 5. Section 6 focuses on the discussion and evaluation of the proposed model. Finally, Section 7 concludes the paper by summarizing the main findings and discussing avenues for future research.

2 Related works

This section introduces the sentiment analysis literature by focusing on deep learning methods for detecting user sentiment or opinions from user posts. Sentiment analysis, as an NLP task, has been extensively studied over the past few decades. They use various deep learning architectures such as convolutions. The authors of this paper refer to [19] for sentiment analysis of drug ratings using deep learning architectures. According to the article, the best performing deep learning architecture for sentiment analysis of drug ratings is BERT (Bidirectional Encoder Representation for Transformers) with LSTM (Long Short-Term Memory) model with a micro F1 score of 0.

9046. However, its cultivation cycle is very long. Convolutional Neural Networks (CNNs) produce acceptable results and require less training time. Khan Hasib and Ahsan Habib [20] proposed a new form of deep learning with different emotion classification capabilities. Using four-layer DNN and CNN for labeling, the combined accuracy rate is 91%. Zahra Rajabi and Ozlem Uzuner [21] study different approaches to sentiment classification in short, informal Twitter text. The multi-filter CNN-Bi-LSTM achieves 85.1% accuracy, outperforming existing models. In [22], Arnab Roy and Muneendra Ojha compared tweet sentiment classification using Google BERT, an attention-based bidirectional LSTM, and a Convolutional Neural Network (CNN). Compared with models, machine learning techniques are less efficient and less accurate in studying emotions [22].According to [23], the Bi-LSTM-based social media sentiment analysis training model outperforms the traditional LSTM model in terms of accuracy and F1 measure. In [24] LSTM and Bi-LSTM, as they are two central architectures of the recurrent neural networks (RNN), were used for evaluating the performance of the predictive models, whereby the LSTM achieved an overall accuracy of 90.59%, while the Bi-LSTM achieved an accuracy of 90.83% [24]. The proposed model in [25] uses a Bi-LSTM self-attention-based CNN model-for classifying user reviews. Based on the experimental results, the model achieved a high level of classification accuracy and F1 measure value is achieved. On the other hand, Sharat Sachin and Abha Tripathi [26] have implemented baseline models for LSTM, GRU, Bi-LSTM, and Bi-GRU on an Amazon review dataset. With an accuracy of 71.19% the bi-GRU model outperformed the other model as it scored higher on all performance measures. The work of Sakirin Tam[27], which proposed the combination of CNNs and Bi-LSTMs for sentiment classification on Twitter, the ConvBiLSTM model with Word2Vec had a 91.13% accuracy rate and outperformed the other models .

The reviewed studies indicate that mental illness can now be detected based on a variety of data types. There was higher prediction accuracy with DL models compared to traditional ML techniques. According to the findings, it is likely that DL models will assist clinicians in the diagnosis of mental health conditions in a more accurate and efficient manner. In order for DL techniques to become a useful tool for treating mental illnesses, more efforts will have to be made to link them with therapeutic treatment [28].

The work presented in this paper has successfully analyzed the urgent need for methods to help detect potential depression at an early stage from textual data [29].

In a study entitled Early Post Traumatic Stress Disorder Diagnosis after Cesarean Delivery using Artificial Neural Networks (ANN), [30]. The derived decision model demonstrates high levels of accuracy based on the results. In spite of the fact that there may be only partial information or readily available information, this is the case. Early in the process, ANNs have proven to be accurate. Among patients diagnosed with PTSD, 94% are symptom-free and 66% are recognized to be suffering from the condition.

The development of a neural network model combined with a support vector machine is presented in [31]. It is able to accurately predict psychological disorders, including anxiety, behavioral problems, depression, and PTSD, in mild traumatic brain injury patients with an accuracy of 73% to 95% for each condition, as well as an overall accuracy of 82.35%.

The authors analyzed semi-structured interviews conducted with PTSD sufferers through sentiment analysis to elucidate how the disorder affects them. A super learner (SL) combines different sentiment analyzers, features to create a model. By using the VADER analyzer 80.4% accuracy was achieved. Based on the results of this study, sentiment extraction from natural language is sufficient to detect post-traumatic stress disorder (PTSD) [32].

An accurate prediction of mental health conditions is crucial for effective intervention. This study [33] aims to evaluate the performance between Decision Tree and Random Forest, in predicting mental health outcomes. The chosen approach involves implementing the Decision

Tree algorithm and assessing its precision rate, as a measure of accuracy, in predicting mental health conditions. Which yielded an 82% precision rate, highlighting its accuracy

The purpose of this study was to develop models for diagnosing post-traumatic stress
This study [34] proposes a predictive model that utilizes structured data and unstructured narrative notes from Electronic Medical Records to accurately identify patients diagnosed with Post-Traumatic Stress Disorder (PTSD) and demonstrates the highest sensitivity (0.77) and F-measure when both structured and text data are incorporated as input.

The authors of [35] showcase the potential of leveraging psychological and societal factors to mitigate the effects of ongoing economic and social crises amid the Ukraine-Russia conflict, utilizing a methodology that involves employing two Bidirectional Recurrent Neural Networks (Bi-RNN) along with the VADER sentiment analyzer, trained on a substantial dataset comprising approximately 11.15 million tweets, with 93% accuracy. Table 1 below summarizes the research papers discussed in this section.

Table 1. Literature Review Summary

Author	Objective	Algorithm	Results
[19]	Benchmark comparison, hybrid modeling	CNN, LSTM, RNN	BERT followed by Bi-LSTM 0.9046% Bi-LSTM followed by a CNN 0.8944%
[20]	Classifying emotions	Four-layer DNN & CNN	91% accuracy
[21]	Classifying emotions	Multi-filter CNN-Bi-LSTM	85.1% accuracy
[22]	Sentiment classification	Google BERT Attention-based Bidirectional LSTM, and (CNN)	BERT 64.1% BI-ATTENTIVE LSTM 60.2% CNN 59.2%
[23]	Sentiment analysis	Deep learning LSTM- Bi-LSTM	Bi-LSTM-based training model was found to be better than the traditional LSTM
[24]	Sentiment analysis	RNN-based LSTM Bi-LSTM	LSTM 90.59% Bi-LSTM 90.83%
[25]	Sentiment	Bi-LSTM self-	high classification

	analysis	attention-based CNN	accuracy and F1 measure value
[26]	Sentiment analysis	LSTM, GRU, and Bi-LSTM, and Bi-GRU	Bi-GRU model outperforms with an accuracy of 71.19%
[27]	Sentiment Classification	Integrating structure of CNN and Bi-LSTM	91.13% accuracy
[28]	Review DL and ML techniques for mental illness detection	ML DL	Higher prediction accuracy with DL models compared to traditional ML techniques data
[29]	Sentiment analysis Depression Detection	Knowledge base depressive vocab	Successfully methods for detecting depression early
[30]	Sentiment classification Cesarean PTSD in women diagnosed early	ANN	Free PTSD 94% PTSD 66%
[31]	Psychological condition prediction	Neural network + SVM	Over all 82.35%
[32]	Sentiment analysis PTSD Detection	ML	80.4%
[33]	Mental health prediction	Decision tree algorithm and trained a model	82%
[34]	PTSD prediction model	MLNN RF	Provides insight that may serve as a potential PTSD indicator
[35]	Sentiment analysis data	BI-RNN based VADER a pre-trained sentiment analyzer	BI-RNN 93%

3 Methodology and Proposed Model

In this section, we present a comprehensive description of the approach employed to investigate the multi-class classification of individuals' emotions during conflict

3.1 Data acquisition, cleaning, and preprocessing

In this study, Twitter tweets from the period of the Ukraine-Russia conflict were analyzed, and a dataset comprising 200,567 tweets was extracted from Kaggle.com [36]. Subsequently, the collected data was subjected to preprocessing in order to prepare it for training and validating the developed model. The cleaning process involved several steps to ensure data quality, including the identification and removal of duplicates, as well as the elimination of irrelevant and null values. Specifically, NAN values were eliminated in accordance with the research goal and problem domain. Examples of the resulting clean data are presented in Table 2.

Table 2. Original vs cleaned Tweets

Tweet	Clean text
#Odessa #Odesa #ðŸ•¯ï¸ ðŸ˜¢ðŸ’”ðŸ˜¢ðŸ•¯ï¸ May the stars carry your sad-ness away. May the flowers fill your heart with beauty. May hope forever wipe away your tears, and above all , may silence make you strong... http	ii May star carry sadness away May flower fill heart beauty May hope forever wipe away tears may silence make strong
With an aching heart I add my voice to that of the com-mon people, who implore the end of the war. In the name of God, listen to the cry of those who suffer, and put an end to the bombings and the attacks! #LetsPrayTogether #Ukraine #Peace	With aching heart I add voice common people implore end war In name God listen cry suffer put end bombing at-tacks

The fundamental component of a sentiment analysis system lies in the NLP (Natural Language Processing) phase. Pre-processing plays a crucial role in enhancing the accuracy of the model by facilitating NLP and text analysis techniques such as lemmatization and tokenization. Moreover, considering the unique characteristics of social media data, additional data preprocessing techniques specific to social media data are essential. Figure 1 illustrates the steps involved in the subsequent section.

1. Eliminate emojis, URLs, mentions, Latin characters, hashtags, mentions, digits, and tags.

2.NLTK library used to remove stop words, add custom stop words and Lemmatizer (by using WordNetLemmatizer.) Keras used for tokenizing words and pad sequences. The tokenization method focuses on converting text into tokens until it becomes vectors. This step is essential to remove unwanted words, as shown in the next steps.

- Replace all NaN values to empty strings using replace() method in python
- Use NLTK library to load stop words for English
- Remove stop words using the NLTK python library
- Add a Custom stop words list
- Python Regex Compile to remove emoji
- RegEx python Functions using sub-function replace one or many matches with a string
- Remove (URLs, Latin letters, characters, hashtags, digits, HTML tags, and punctua-tions)
- Using NLTK library "WordNetLemmatizer" Lemmatizer
- Keras used for Tokenized words and pad sequences

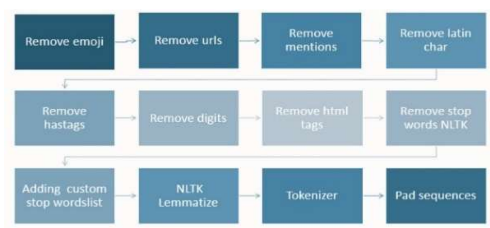

Fig. 1. Pre-processing pipeline (NLP)

3.2 Data Visualization Tools

A comprehensive analysis was conducted to explore the relationships within the data. Visualizations were created using various pre-built Python libraries to provide insightful representations. Bar graphs, line graphs, and word clouds were employed to visualize and interpret the data in a visually informative manner

3.3 Classification

In our study, we proposed a streamlined and effective Deep Learning-based classification model utilizing a single bidirectional LSTM network for English sentiment analysis. By employing a multi-class classification approach to classify people's emotions during the war, we found that English sentiment analysis achieved optimal results when utilizing a single bidirectional LSTM. The architecture of our proposed model, depicted in Figure 2, consists of several key components.

Firstly, an embedding layer where every index, corresponding to a unique word in the data set, is transformed into a real-valued feature vector. The Bi-LSTM layer is capable of reading input reviews in both directions, forward and backward Thus, the model could capture more sentiment information. The output of Bi-LSTM can be summarized by concatenating the forward and backward states, then performing pooling to reduce the number of parameters and computations. Dense Layer will describe how neurons are connected to the next layer of neurons. This works for changing the dimension of the output by performing matrix vector multiplication. Lastly, the Dropout layer will remove noise from neurons' inputs. This prevents overfitting the model.

Fig. 2. The architecture of the proposed Model

4 Experimental and Evaluation

This section encompasses several key aspects, including system configuration, data analysis, performance metrics, results, and comparative analysis. It provides detailed information on the setup and configuration of the system used in the research. Furthermore, it outlines the process of analyzing the data. Performance metrics are utilized to evaluate the effectiveness of the proposed model. The section also presents the results obtained from the experiments conducted and conducts a comparative analysis to assess the performance of different approaches or models.

4.1 Model Configuration

The training and testing datasets were divided into an 80:20 ratio. During the training phase, classification is performed based on the network's performance, and the degree of classification error is calculated. This error is then minimized by adjusting the weights or parameters of the network. The iterative process continues until the model reaches a state of convergence, where further adaptation is no longer necessary. To evaluate the performance of the model, two important parameters were chosen:

- Metric: Accuracy report and confusion matrix were utilized to measure the performance of the model in the experiment.

- Loss function: The loss function calculates the loss value, which is subsequently minimized by adjusting the weights of the network during the training phase. For the multiclass classification models used in this study, the categorical_crossentropy loss function was employed.

- Optimizer: An optimizer function was employed to update the network weights based on the output of the loss function. The Adam optimizer was utilized in our experiments to optimize the performance of the model and enhance convergence during the training process. Tune Hyper parameters

Neural networks possess the ability to capture intricate relationships between inputs and outputs. However, it is crucial to address the potential issue of overfitting, where the model learns to fit the noise in the data rather than the underlying patterns. Overfitting can result in a decrease in the model's classification capacity. To mitigate overfitting, dropout layers were employed in our model.In order to optimize the performance of the model and achieve a better representation of the problem at hand, it is essential to adapt the hyperparameters to the specific characteristics of the dataset. By fine-tuning these hyperparameters, the model can be refined to enhance its performance. Table 3 presents the values of the hyperparameters utilized in our proposed model. These hyperparameters were carefully selected to ensure optimal performance and mitigate the risk of overfitting, allowing the model to effectively capture the underlying patterns in the data.

- A network's complexity: is determined by the number of layers it has. A careful consideration must be given to this value. An over fit model will be able to learn too much information about the training data by using too many layers. If the model has too few layers, under fitting can occur. For text classification datasets, we experimented with a single BI-LSTM layer.

- Number of units per layer: is an important consideration in neural network design as it impacts the network's ability to perform effective transformations. Each layer should have sufficient capacity to capture and process relevant information. In our study, we experimented with different numbers of units per layer and found that using 32 / 64 units yielded favorable performance outcomes. These unit configurations were identified as effective performers in our model, allowing for effective information processing and transformation within each layer.

- Dropout rate: In order to implement regularization techniques, dropout layers were incorporated into the model architecture. These layers are designed to mitigate overfitting by randomly dropping a fraction of the input units during training. Specifically, a dropout rate of 0.25 was employed in this study as a precautionary measure against overfitting.

- Embedding dimensions: For the purpose of word embedding, the dimensionality of the vector representation plays a crucial role. It determines the number of dimensions that each vector possesses. In this study, an embedding dimension of 100 was chosen for the word embedding process.

Table 3. Optimal hyper parameters of the proposed model

Hyper parameters	Value

Optimizer	Adam
Loss function	categorical_crossentropy
Batch size	64
Epochs	4
Dropout rate	0.25
Bi-LSTM Nodes	64
Max length	300
Learning rate	0.001
Regularizes	L2

4.2 Evaluation

A performance metric includes accuracy, precision, recall, and F1-Score. The next subsection briefly explains the state of the art evaluation techniques.

4.3 Accuracy Report

In order to assess the effectiveness of our model, several performance metrics including accuracy, precision, recall, and F1-score were employed. The accuracy report provided an evaluation of these metrics. Precision indicates the percentage of correct positive predictions, recall represents the proportion of positive cases correctly identified, and the F1-score indicates the overall correctness percentage of positive predictions. Additionally, the support value represents the frequency of occurrence for each specific class within the dataset. The obtained results are presented in Table 4.

To further evaluate the performance of our proposed architecture layers, a comparison was made with existing state-of-the-art models in terms of accuracy [38].

Table 4. Accuracy report for each class

Class	Precision	Recall	F1-score	Support
Neutral	0.89	0.90	0.89	14,617
Positive	0.93	0.86	0.89	6,097
Negative	0.94	0.94	0.94	19,398

Table 5. Comparison of the dataset with other approaches

Method	Accuracy
LSTM+CBA+LA	90.10 %

CNN+LSTM	88.90%
WALE+LSTM	89.50%
FARNN-ATT	89.22%
Single BI-LSTM	90.585%
Proposed Model	**91.64%**

4.4 Training and validation accuracy

The number of epochs for training was set to 20 however, to ensure efficient use of time and mitigate overfitting, an early stop callback with a patience value of 4 was implemented based on validation loss. In the context of training, epochs refer to the iterations of the entire dataset. Patience represents the number of consecutive epochs in which no further improvement is observed, leading to the termination of the training process. The quality of the model is assessed by evaluating the associated errors or losses. For instance, a higher validation loss suggests that the trained model may not generalize well to unseen validation data, indicating its inadequacy for the testing phase. The outcomes of the training and validation, along with the model loss, are depicted in Figure 3 and Figure 4, respectively.

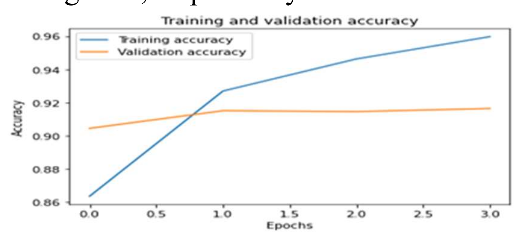

Fig. 3. Training and validation accuracy

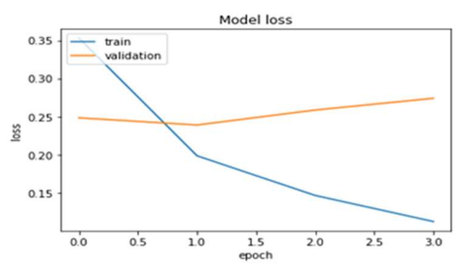

Fig. 3 Model Loss

4.5 Confusion Matrix

The evaluation of classification accuracy often involves the use of a confusion matrix. In this matrix, the predicted labels are represented along the x-axis, while the true labels of the samples in the test set are shown along the y-axis. Based on the predictions made on the test set, the confusion matrix provides insights into the classification performance.

Figure 5 illustrates the confusion matrix, where the true label indicates the actual sentiment of the tweets, while the predicted label represents the sentiment predicted by the model. By analyzing the confusion matrix, it can be deduced that a total of 36,769 tweets were correctly predicted with the same sentiment as their true labels (summing the values in the diagonal elements of the matrix: 13,254+5,284+18,231). On the other hand, a total of 3,353 tweets were predicted with an incorrect sentiment (summing the values in the off-diagonal elements of the matrix: 355+1,018+674+139+1,076+91).

From the perspective of the true labels, it can be concluded that 91.64% of the tweets in the dataset were correctly classified with the corresponding sentiment.

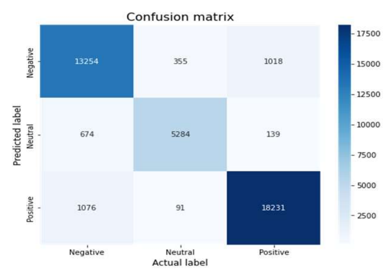

Fig. 5. Confusion Matrix

5 ~~Data Analysis~~ Results and Analysis

5.1 Tweets Sentiment Analysis

Sentiment analysis serves as a valuable tool for researchers to gain insights into the objectives, communication content, and effects of social media posts. Through the utilization of word clouds, visual representations of sentiment analysis are employed to classify tweets based on the presence of negative words and emotions. The objective of this study is to investigate the association between the occurrence of negative words and emotions and the presence of specific mental disorders. By analyzing the patterns and prevalence of negative sentiments expressed in social media content.

5.2 Sentiment Words according to Polarities

Figure 6 presents the distribution of tweets based on their sentiment classification, revealing the proportions of positive, neutral, and negative sentiments. The findings indicate that 15.27% of the tweets exhibited positive sentiments, while 36.96% were classified as neutral, and a substantial majority of 47.77% displayed negative sentiments. The analysis indicates a higher prevalence of negative sentiments compared to neutral and positive sentiments. Negative tweets constituted the largest share, suggesting a prevalent expression of negativity among individuals. Conversely, positive reactions were relatively less common. The predominance of negative sentiments suggests a significant shift in the prevailing sentiment among the population, indicating the seriousness of the situation and the emotional state of individuals, as evident from the content of negative tweets.

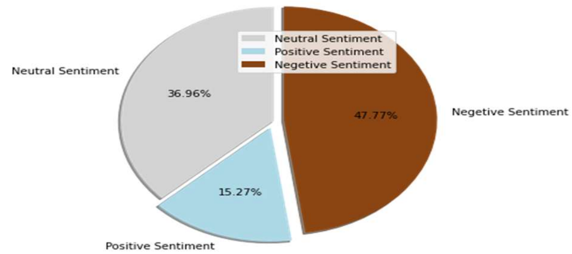

Fig. 6. Tweets Sentiment percentage

5.3 Sentiment Words according to Feelings

Based on the analysis of emotions expressed in the tweets depicted in Figure 7, it was observed that anger emerged as the most prevalent and prominent emotion. Following anger, the subsequent emotions in terms of frequency and prominence were optimism, sadness, and happiness.

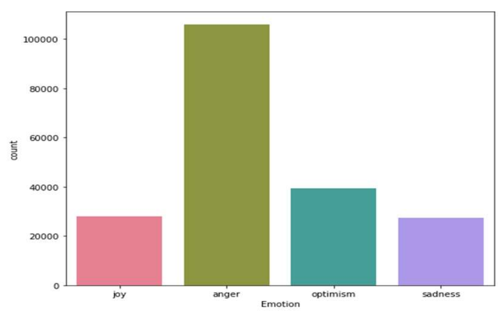

Fig. 7. Emotion Analysis of Tweets

5.4 Word Cloud

A word cloud is a visual representation that showcases the most commonly used words within a dataset. The size of each word in the cloud corresponds to its frequency, with larger fonts indicating more frequent usage. By analyzing the word cloud, we can gain insights into people's emotions during times of conflict. Figure 8 represents the word cloud for negative sentiment words, while Figure 6 illustrates the word cloud for frequently occurring negative emotions.

Upon examining the emotion word cloud in Figure 9, it is apparent that the symptoms associated with Post-traumatic stress disorder (PTSD), as outlined in the DSM-IV-TR criteria, are prominently mentioned. In the DSM-IV-TR, PTSD is classified into three symptom categories, comprising a total of 17 specific symptoms. The diagnosis of PTSD requires meeting the criteria for all three categories. For instance, an individual must exhibit at least three of the seven symptoms related to persistent avoidance, which can include consciously avoiding thoughts, feelings, or discussions linked to the traumatic event, as well as experiencing memory loss.

It is noteworthy that the inclusion of DSM-IV-TR criteria in the word cloud suggests a broad association between the expressed emotions and potential indicators of PTSD within the analyzed dataset.

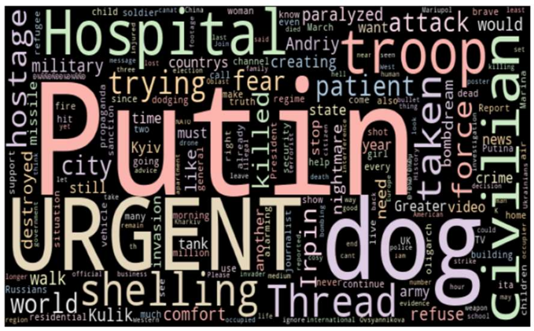

Fig. 8 Frequent words visualization for negative words

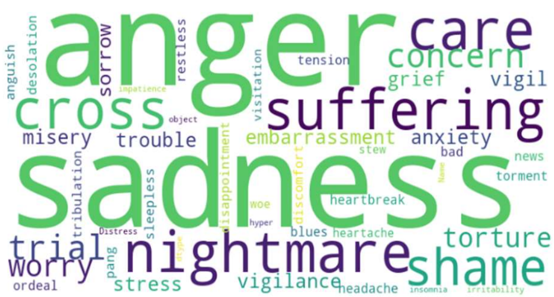

Fig. 9. Frequent visualization of negative emotions

5.5 Most Frequently negative emotions words

Figure 10 presents a visual representation depicting the prevalence of negative emotions that occur with the highest frequency.

	Distress	Count
0	anger	105984
1	sadness	27506
2	nightmare	3496
3	care	1099
4	suffering	541
5	cross	507
6	shame	449
7	concern	272
8	trial	198
9	worry	160
10	torture	153
11	vigilance	105
12	sorrow	101
13	stress	91
14	trouble	86
15	anxiety	78
16	embarrassment	55
17	vigil	46
18	misery	44
19	grief	35
20	heartbreak	28
21	bad news	21
22	tension	14
23	blues	7
24	disappointment	7
25	headache	6
26	ordeal	6
27	torment	6
28	anguish	5
29	tribulation	5

Fig. 10. Most Frequently Occurring Negative Emotions

5.6 Most Frequently occurring Words

Figure 11 exhibits the top 20 most frequently occurring words within the dataset Which are Support,President,Kyiv,today,Putin,dog,invasion,europe,need,us,military,word,putines,stand, force,thread,minister,urgent,help and like.

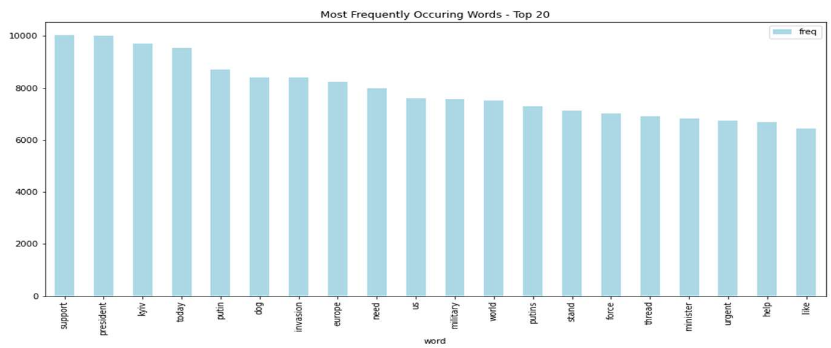

Fig. 11. Most Frequently Occurring 20 Words - Top

6 Discussion

In this research study, a single Bi-LSTM network was developed as an efficient model to enhance the identification of word indicators related to Post-traumatic stress disorder (PTSD) in individuals who may be at risk of developing the disorder. The analysis conducted revealed significant impacts on the lives and mental health of individuals due to the ongoing conflict between Ukraine and Russia. The war has had a profound effect on people, but it is noteworthy that many individuals are still demonstrating resilience and maintaining a positive outlook amidst the adversity.

The identification of individuals expressing negative sentiments through tweets can serve as an important marker for vulnerability, and appropriate authorities should be notified to provide necessary support. The findings also indicate that emotions expressed in tweets can serve as valuable early detection evidence for individuals experiencing PTSD. The proposed approach achieved competitive results, particularly in terms of accuracy, outperforming the novel methods compared in this study.

The overall accuracy of 91.64% indicates that the sentiment classifier performed well in categorizing sentiment polarities. The model exhibited favorable scores in precision, recall, F1-scores, and the calculation of the confusion matrix. The confusion matrix provided detailed insights into the classification performance, particularly regarding positive, negative, and neutral classifications. This information supported the percentages of tweets classified as positive, negative, and neutral, with 47.77% classified as negative, 36.96% as neutral, and 15.27% as positive.

7 Conclusion

This paper introduces a single bidirectional LSTM network for the analysis of public emotions and opinions during the Russia-Ukraine conflict, with the aim of early detection of Post-Traumatic Stress Disorder (PTSD) symptoms. Detecting mental health issues at their early stages is crucial for providing appropriate interventions to potential sufferers. The study conducts a comprehensive analysis of tweet sentiments to identify early signs of PTSD.

Recognizing and understanding emotions is essential for addressing and preventing negative emotions, such as anger and sadness. These emotions, when triggered during conflicts, can lead to various mental health outcomes, including depression, loneliness, and social anxiety. By eliminating negative emotions, these adverse outcomes can be significantly reduced [3]. Hence, careful understanding and analysis are required for effective elimination. The results of the study confirm that the percentage of negative emotions expressed in tweets related to the conflict is 47.77%.

The paper presents an English tweet sentiment analysis model based on deep learning, employing multiclass classification to detect PTSD symptoms. Preprocessing steps are applied to prepare the collected data for training and validating the developed model. The proposed model utilizes Bi-LSTM, which enhances contextual understanding through a two-way network, providing better information extraction compared to other approaches.

The analysis reveals that the symptoms of PTSD align well with the criteria defined in the DSM-IV-TR. Due to the lack of support for grid automatic search in Keras libraries for LSTM, various experiments are conducted to optimize hyper parameters and refine the model.

By employing a single Bi-LSTM memory layer and global Max pooling 1D, an accuracy of 91.64% is achieved. Compared to previous studies, the proposed framework exhibits improved performance in terms of accuracy.

The paper highlights the importance of raising awareness among individuals about their mental health during times of war. It suggests the development of technologies, such as online clinics, to support people during conflicts when mobility is restricted. Furthermore, understanding the impact of emotions on mental health can contribute to early detection of mental illnesses. By emphasizing emotions during crises, efforts can be made to promote and maintain mental health and well-being.

The presented framework offers a robust explanation and classification approach.
Among the future prospects of this study, this study will consider a multilingual with more languages including the Arabic text in order to evaluate the shifting emotions and feelings of individuals over time, and also to ascertain if there are any noticeable changes in their mental health over time.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Funding No funding was received for conducting this study.

References

1. KidsHealth Behavioral Health Experts. Understanding Your Emotions (for Teens). Nemours KidsHealth. https://kidshealth.org/en/teens/understand-emotions.html. (accessed 2023-03-06).
2. Grewal, D.; Salovey, P. Feeling Smart: The Science of Emotional Intelligence. American Scientist 2005, 93 (4), 330. https://doi.org/10.1511/2005.54.969.
3. Pettinelli, M. The Psychology Of Emotions, Feelings and Thoughts. Connexions 2007.
4. Liu, B.; Zhang, L. A Survey of Opinion Mining and Sentiment Analysis. In Mining Text Data; Springer US: Boston, MA, 2012; pp 415–463.
5. Zucco, C.; Calabrese, B.; Cannataro, M. Sentiment Analysis and Affective Computing for Depression Monitoring. In 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); IEEE, 2017.
6. Hovy, E. H. What Are Sentiment, Affect, and Emotion? Applying the Methodology of Michael Zock to Sentiment Analysis. In Language Production, Cognition, and the Lexicon; Springer International Publishing: Cham, 2014; pp 13–24.
7. Kadhim, A. I.; Cheah, Y.-N.; Ahamed, N. H. Text Document Preprocessing and Dimension Reduction Techniques for Text Document Clustering. In 2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology; IEEE, 2014.
8. Toutanova, J. D. ; M.-W. C. ; K. L. ; K. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding.
9. Hourrane, O.; Idrissi, N.; Benlahmar, E. H. An Empirical Study of Deep Neural Networks Models for Sentiment Classification on Movie Reviews. In 2019 1st International Conference on Smart Systems and Data Science (ICSSD); IEEE, 2019.
10. Sukheja, S.; Chopra, S.; Vijayalakshmi, M. Sentiment Analysis Using Deep Learning – A Survey. In 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA); IEEE, 2020.
11. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. Neural Computation 1997, 9 (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.
12. Yash Indulkar; Patil, A. Sentiment Analysis of Uber & Ola Using Deep Learning. In 2020 International Conference on Smart Electronics and Communication (ICOSEC); IEEE, 2020.
13. Pasupa, K.; Seneewong Na Ayutthaya, T. Thai Sentiment Analysis with Deep Learning Techniques: A Comparative Study Based on Word Embedding, POS-Tag, and Sentic Features. Sustainable Cities and Society 2019, 50, 101615. https://doi.org/10.1016/j.scs.2019.101615.
14. Zielinski , Melissa . ISTSS. What Is Traumatic Stress? https://istss.org/public-resources/trauma-basics/trauma-during-adulthood (accessed 2023-03-07).

15. Gersons, B. P. R.; Carlier, I. V. E. Post-Traumatic Stress Disorder: The History of a Recent Concept. British Journal of Psychiatry 1992, 161 (6), 742–748. https://doi.org/10.1192/bjp.161.6.742.
16. Cognitive Processing Therapy (CPT). https://www.apa.org. May 25, 2017. https://www.apa.org/ptsd-guideline/treatments/cognitive-processing-therapy (accessed 2023-03-07).
17. Colón-Ruiz, C.; Segura-Bedmar, I. Comparing Deep Learning Architectures for Sentiment Analysis on Drug Reviews. Journal of Biomedical Informatics 2020, 110, 103539. https://doi.org/10.1016/j.jbi.2020.103539.
18. Hasib, K. Md.; Habib, Md. A.; Towhid, N. A.; Showrov, Md. I. H. 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); IEEE, 2021.
19. Rajabi, Z.; Shehu, A.; Uzuner, O. A Multi-Channel BiLSTM-CNN Model for Multilabel Emotion Classification of Informal Text. In 2020 IEEE 14th International Conference on Semantic Computing (ICSC); IEEE, 2020.
20. Roy, A.; Ojha, M. Twitter Sentiment Analysis Using Deep Learning Models. In 2020 IEEE 17th India Council International Conference (INDICON); IEEE, 2020.
21. Chen, L.-C.; Lee, C.-M.; Chen, M.-Y. Exploration of Social Media for Sentiment Analysis Using Deep Learning. Soft Computing 2019, 24 (11), 8187–8197. https://doi.org/10.1007/s00500-019-04402-8.
22. Alam, K. N.; Khan, M. S.; Dhruba, A. R.; Khan, M. M.; Al-Amri, J. F.; Masud, M.; Rawashdeh, M. Deep Learning-Based Sentiment Analysis of COVID-19 Vaccination Responses from Twitter Data. Computational and Mathematical Methods in Medicine 2021, 2021, 1–15. https://doi.org/10.1155/2021/4321131.
23. Bhuvaneshwari, P.; Rao, A. N.; Robinson, Y. H.; Thippeswamy, M. N. Sentiment Analysis for User Reviews Using Bi-LSTM Self-Attention Based CNN Model. Multimedia Tools and Applications 2022, 81 (9), 12405–12419. https://doi.org/10.1007/s11042-022-12410-4.
24. Sachin, S.; Tripathi, A.; Mahajan, N.; Aggarwal, S.; Nagrath, P. Sentiment Analysis Using Gated Recurrent Neural Networks. SN Computer Science 2020, 1 (2). https://doi.org/10.1007/s42979-020-0076-y.
25. Tam, S.; Said, R. B.; Tanriover, O. O. A ConvBiLSTM Deep Learning Model-Based Approach for Twitter Sentiment Classification. IEEE Access 2021, 9, 41283–41293. https://doi.org/10.1109/access.2021.3064830.
26. Su, C.; Xu, Z.; Pathak, J.; Wang, F. Deep Learning in Mental Health Outcome Research: A Scoping Review. Translational Psychiatry 2020, 10 (1). https://doi.org/10.1038/s41398-020-0780-3.
27. Tao, X.; Zhou, X.; Zhang, J.; Yong, J. Sentiment Analysis for Depression Detection on Social Networks. In Advanced Data Mining and Applications; Springer International Publishing: Cham, 2016; pp 807–810.

28. Orovas, C.; Orovou, E.; Dagla, M.; Daponte, A.; Rigas, N.; Ougiaroglou, S.; Iatrakis, G.; Antoniou, E. Neural Networks for Early Diagnosis of Postpartum PTSD in Women after Cesarean Section. Applied Sciences 2022, 12 (15), 7492. https://doi.org/10.3390/app12157492.
29. Dabek, F.; Caban, J. J. A Neural Network Based Model for Predicting Psychological Conditions. In Brain Informatics and Health; Springer International Publishing: Cham, 2015; pp 252–261.
30. Sawalha, J.; Yousefnezhad, M.; Shah, Z.; Brown, M. R. G.; Greenshaw, A. J.; Greiner, R. Detecting Presence of PTSD Using Sentiment Analysis From Text Data. Frontiers in Psychiatry 2022, 12. https://doi.org/10.3389/fpsyt.2021.811392.
31. Laijawala, V.; Aachaliya, A.; Jatta, H.; Pinjarkar, V. Classification Algorithms Based Mental Health Prediction Using Data Mining. In 2020 5th International Conference on Communication and Electronics Systems (ICCES); IEEE, 2020.
32. https://www.frontiersin.org/articles/10.3389/frai.2023.1163577/full
33. Zafari, H.; Kosowan, L.; Zulkernine, F.; Signer, A. Diagnosing Post-Traumatic Stress Disorder Using Electronic Medical Record Data. Health Informatics Journal 2021, 27 (4), 146045822110532.
34. Ramírez, I.; Vargas, J. A Sentiment Analysis of the Ukraine-Russia Conflict Tweets Using Recurrent Neural Networks; https://www.researchgate.net/publication/361275253: Mexico, 2022; pp 1–6.
35. ssaisuryateja. EDA_and_Sentiment_Analysis. Kaggle. May 4, 2022. https://kaggle.com/code/ssaisuryateja/eda-and-sentiment-analysis (accessed 2023-03-07).
36. Kagan, V.; Rossini, E.; Sapounas, D. Sentiment Analysis for PTSD Signals; Springer Science & Business Media, 2013.
37. Hameed, Z.; Garcia-Zapirain, B. Sentiment Classification Using a Single-Layered BiLSTM Model. IEEE Access 2020, 8, 73992–74001. https://doi.org/10.1109/access.2020.2988550.