

Emotion Detection from Text and Sentiment Analysis of Ukraine Russia War using Machine Learning Technique

Abdullah Al Maruf¹, Zakaria Masud Ziyad², Md. Mahmudul Haque³, Fahima Khanam⁴
Department of Computer Science and Engineering, Bangladesh University of Business and Technology
Mirpur, Dhaka, 1216, Bangladesh

Abstract—In the human body, emotion plays a critical function. Emotion is the most significant subject in human-machine interaction. In economic contexts, emotion detection is equally essential. Emotion detection is crucial in making any decision. Several approaches were explored to determine emotion in text. People increasingly use social media to share their views, and researchers strive to decipher emotions from this medium. There has been some work on emotion detection from the text and sentiment analysis. Although some work has been done in which emotion has been recognized, there are many things to improve. There is not much work to detect racism and analysis sentiment on Ukraine -Russia war. We suggested a unique technique in which emotion is identified, and the sentiment is analyzed. We utilized Twitter data to analyze the sentiment of the Ukraine-Russia war. Our system performs better than prior work. The study increases the accuracy of detecting emotion. To identify emotion and racism, we used classical machine learning and the ensemble method. An unsupervised approach and NLP modules were used to analyze sentiment. The goal of the study is to detect emotion and racism and also analyze the sentiment.

Keywords—Emotion detection; racism; sentiment analysis; social media; machine learning; ensemble; Ukraine-Russia

I. INTRODUCTION

Emotion is a strong feeling caused by one's circumstances and conditions, moods, interpersonal connections, or pleasure and dissatisfaction. The emotional experience includes perceptions of the world, cognitive capacities, behavioral reactions, metabolic anomalies, and instrumental activity. Emotions are difficult to define because they are a fleeting state of mind. Images, speech, facial expressions, textual information, emoticons, and other kinds of expression may all be used to determine emotion. Textual data is essential for research [1]. Massive amounts of text-based data have been generated regularly in recent years via social media and conversations such as messenger, Whatsapp, Twitter, and other means [2]. The progression of digital communications and its popularity, particularly virtual networking, keeps individuals interested in how they connect and communicate amongst themselves. People have become accustomed to expressing feelings lightly and intuitively through social media communication and the simplicity of responses. People use social media to keep up with what's happening in the world and to share their opinion and feedback via likes, comments, and shares, among other things [3]. Today's most popular social media platforms are Facebook, Twitter, and Instagram. People visit Facebook to keep in touch with friends, family, and loved ones, learn about what's happening worldwide, and express what's important

to them, according to Facebook's vision and mission. People incline to be more verbose on Facebook, yet posts go through more simple "likes" than lengthy comments. Since February 2017, Facebook has included additional capabilities that allow users to express their specific feelings in reaction to a post, such as the ability to mention "love" or "sadness" instead of liking a post [4]. People use Twitter to express their thoughts, feelings, and views through short messages or tweets at any time. Individuals' emotional states of mind, such as joy, worry, and hopelessness, are captured overtly or indirectly within those short messages along with bigger communities, such as the viewpoints of people in a particular country [5][6]. There are five ways to emotion identification from text, including keyword-based, lexical/corpus-based, learning-based, hybrid-based, and deep learning-based approaches, but each has its limitations[7]. Specifically, identifying object words from tweets is named object-oriented feature to perform sentiment—Bi-gram, uni-gram model with object-oriented feature effective better [8]. Emotion was discovered and recognized using machine learning and deep learning methods. There are also several classifications for identifying emotions; some of them are the k-nearest neighbor(KNN) algorithm, Support vector machine(SVM), decision tree(DT), random forest(RF), linear discriminant analysis(LDA), etc. DT is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, KNN is a method for classifying objects based on closest training examples in the feature space, LDA is a method used in statistics, pattern recognition, and machine learning to find a linear combination of features which characterizes or separates two or more classes of objects or events and SVM analyze data with recognizing patterns used for classification and regression analysis. Based on the lexical approach, Real-time emotional analysis was carried out, and the data was obtained from online social media[9]. Sentiment analysis may be performed using a deep learning model that has been pre-trained, as well as unsupervised algorithms such as Valence Aware Dictionary and sEntiment Reasoner(VADER), textblob, and k-means clustering. It's also possible to gauge sentiment using lexicons. The sentiment was analyzed using large-scale data from Twitter and a machine learning-based technique. Sentiwordnet and sentiment are sentic computing-based public lexicons [10]. Emotion is identified from audio sources following machine learning techniques and categorized into six basic emotions. Auto weka performed the best outrun compared to SVM, KNN, and multi-layer perceptron (MLP) [11]. Finding out mental instability RF shows (87%) accurate

results from the manually surveyed datasets [12]. The basic facts are that domain adaptability and accuracy are the main constraints for emotion recognition from text. The advantages of the suggested approach are an improvement in emotion detection accuracy. The dataset was compiled from a number of sources. The dataset will aid researchers in overcoming this industry's problem of domain adaptation. The research project also analyzes the war between Russia and Ukraine's tweets for mood. The topic has fewer datasets that are readily available. Therefore, gathering a dataset is the main contribution to this field. The desired model to address accuracy issues is ensemble model learning. The main goal of the research was to increase accuracy, thus we utilized a variety of ensemble techniques. We proposed a unique technique where ensemble and classical machine learning are applied. We also used an unsupervised approach to analyze the sentiment. In our work, we will show an approach that detects emotion, analyzes sentiment, and detects racism. We used various methods and techniques to identify emotion, racism, and sentiment analysis. To increase the speed and quality of the learning process, we applied machine learning and ensemble method using parameters. The following is the structure of this paper: Some recent studies in this field are addressed in Section 2. Section 3 details our study technique. Section 4 contains the results, and We'll talk about our strengths and weaknesses and provide suggestions for the future. Then, we eventually finished the paper.

II. RELATED WORK

According to a case study [12], a 3-turn dialogue detects and distinguishes emotions. In three turns, it contains a collection of human emotions on Twitter. They also talk about the Amiens system, which detects emotions in text messages. The algorithm applied the Long short-term memory (LSTM) model to recognize human emotions based on the in-depth reading. The system's primary input is a mix of word2vec and doc2vec embedding. Then, for that issue, utilize the most current Bi-Long Memory Short Term category, which used word embedding as input and predicted human emotion. When the Amiens score is 0.7185, the output results reveal considerable variations in f points above the model's base. As a potential future project, they plan to expand hybrid approaches through emotion handling and emotion lexicons management. A case study showed the interaction of emotion by applying hybrid and machine learning methods for six basic emotion categories. A comparison study of speaker-independent and dependent recognition was represented [13]. As different formats of emotion detection systems exist, the text identification system is one of them. A machine learning-based automated system was implemented to understand a textual form of data to analyze its class. SVM was performed with a 63.5% accuracy rate in the following system of automatic text identification [14]. Linguistic-related emotion identification has some complexity, and little research has been done on it. For every individual language, the data preprocessing step becomes the most challenging part. Preprocessing steps followed tokenization, segmentation, and other extraction methods to filter raw data for better applicability. A system was implemented to perform a hybrid approach to classify six basic emotions for Punjabi words [15]. Urdu language-based work focused on commercial-related emotion detection [16]. The algorithm support vector classifier(SVC), KNN, RF,

and Naive Bayes(NB) was applied to datasets (Smartphone and Sports), and SVC performed best among them with accuracy above 80%. NB achieved the goal of detecting Bangla text-related emotion. As a complex linguistic analysis, three emotions category presented in this process [17]. Facial expression means a lot of things about the mentality of a person. The machine learning-based method SVM algorithm was used to detect the actual emotion of a person from an image. Framework and UI environment constructed in their system [18]. In the present world, the most used handheld electronic device is a smartphone. A self-automated system named "iself" is represented in their work where it can detect the user's emotions from their smartphone data. The system's structure describes how emotion detection happened by analyzing smartphone users' data [19]. Social media is a great source for mass data collection. Five preprocessing techniques were used for Arabic language emotion detection contextual research. The minimal sequential optimization(SMO) classifier performed better than NB and SF in classifying six basic emotions distinguished from the text [20]. Emoticons can be an expression of emotion. The study reviewed the Arabic tweets and labeled them into the four emotion categories. Testing was performed using SVM and MNB algorithms for developing the system[21]. Deep learning techniques were used for detecting emotion from the Persian text document. Preprocessing technique word2vec was used for normalizing the dataset. NB, DT, and SVM algorithms were executed to classify emotion, and 10-fold cross-validation was performed to evaluate the performance. The SVM algorithm scored the highest accuracy in their system [22]. A system overviewed in the research is SEDAT, which can detect emotion from Arabic social media tweets in real-time. Convolutional Neural Network(CNN)-LSTM neural network technique was applied to construct the system. Their work shows that the system is better accurate than the TeamUNCC system [23]. Arabic language-based sentiment analysis work combined three Arabic Steemers (ISRI, Light Steemer, snowball). The research outcome exhibits a comparison between deep learning and machine learning where CNN, SVM, NB, and MLP are exerted in implementation [24]. Emotion classification and social media are almost connected in parallel cause it is a rich data sources. In this case, comparative model analysis on different datasets is highlighted. TME model acquires maximum accuracy than other models, and WME remains close to TME [25]. From youtube users' text comments, emotion identification was conducted in this research. Unsupervised machine learning technique executed to determine emotion category from YouTube comment data corpus. The system achieved an average of 92.75% precision, close to another existing system that applied the SVM algorithm[26]. The microblog-based emotion classification task is represented by using deep learning techniques. CNN model was selected for this purpose, though the experiment was held in four phases. Sina Weibo was the data source for the experiment. This study exposed CNN as the maximum accurate (97.60%) for detecting emotion from microblogs [27]. Another research is a deep learning-based emotion recognition system. It also demonstrates how excellent deep learning is at detecting text emotion. Data preparation includes sentence segmentation and word embedding. SVM and LSTM are employed to categorize seven different emotion groups. Among these, LSTM has a higher efficiency (94.7%) than SVM [28]. Most of the existing

sentiment analysis and emotion detection research is performed on deep learning or machine learning-based unique techniques. This study presents a system that converts multi-label classification problems into a single binary classification and then solves the problem by applying the deep learning method. The proposed system model Binary Neural Network (BNET) shows the best accuracy in evaluation measured in (multi-label accuracy) Jaccard [29]. Although several textual-based emotion recognition systems exist, the researcher concentrated on discovering emotional states from poetry text in this study. The suggested method is capable of detecting 13 different emotion classes. The system recognizes the context and value of tokens. Comparative investigation demonstrates that the suggested model outperforms the CNN-BiLSTM model in this system [30]. Emotion-based research is becoming more popular since computer machines can predict likely outcomes. Several research and records used a variety of procedures and strategies to identify particular human emotions. Previous work had various drawbacks; thus, the suggested system employed the layered LSTM approach to identify emotion. The suggested system outperformed the LSTM and SVM approaches, reaching a 99.2% accuracy rate [31]. Another research represents EmoDet2, a system that innovates deep learning architecture for identifying emotion from the text. The suggested system used two different algorithms: BERT and BiLSTM. EmoDet2 has been created by applying the ensemble method. The system F1 score (0.75) outperforms the baseline model, which was encouraging. Semeval-2019 dataset had the best accuracy in recognizing emotion from the text, allowing the EmoDet2 system to outperform the baseline model [32]. A suggested method was given in another study that used a machine-learning approach to identify emotions. The dataset is divided into two types, one binary, and the other multi-class. The system was programmed with seven machine-learning models for the experiment. The sentiment is divided into two categories: joyful and unpleasant. The LR-SGD and the tf-idf classifier produced ideal results in a comparison study. The voting classifier beat all others, obtaining 84 percent accuracy using both TF and tf-idf [33]. According to a study by [34], social media is getting a lot of attention in today's world. Public and private opinions on various topics are constantly expressed and distributed via social media platforms. One of the most prominent social media platforms is Twitter. Twitter is a social media platform that allows buddies, relatives, and colleagues to interact and keep in touch by exchanging short, frequent messages. Twitter is the main microblogging website that allows users to post status updates (known as "tweets"). These tweets occasionally reflect thoughts on a variety of issues [35]. Twitter is a real-time microblogging platform that allows individuals or groups to express their opinions on a topic and have them appear on a timeline. Web search apps and real-world applications such as current global trends and world events, and extracting the most up-to-date information regarding occurrences use microblog data for analysis and conclusion-making. Sentiment Analysis and Opinion Mining are two types of text mining that involve the analysis of sentiments, opinions, and emotions and the assessment of the text's content. When evaluating people's thoughts, feelings, assessments, attitudes, and responses to services, goods, organizations, personalities, events, themes, and issues, and their qualities, sentiment analysis is another name for opinion mining. Sentiment and subjectivity are strongly influenced by the context and domain in which they occur.

It is not only due to language changes but also to the dual meaning of feelings of the same term in various domains. The processes of extracting nontrivial patterns and intriguing information from unstructured script texts are called opinion mining and sentiment analysis, respectively [36]. Before the internet, it took a long time for information about a company's stock price, direction, and general attitudes to spread among individuals. Web technology has ushered in a new era of rapid information transmission and retrieval. Applying positive or negative information about a company, product, person, or other entity may be as simple as clicking a mouse or utilizing microblogging services like Twitter [37]. Social media platforms have grown in importance as a forum for political debate worldwide. Users may use Twitter to send tweets, which are short communications of up to 140 characters. The number of people using Twitter is steadily increasing. Over 100 million active users throughout the world send over 250 million tweets every day, according to the business. In several fields, including the financial market, politics, and social movements, sentiment in Twitter data has been utilized for prediction or assessment. Emerging events or news are frequently followed by a surge in Twitter activity, offering a unique chance to assess the relationship between stated public mood and political outcomes. Furthermore, sentiment analysis may be used to investigate how these occurrences impact public perception. It provides a fresh and contemporary perspective on the dynamics of the election process and public opinion to the general public, the media, lawmakers, and academics. They present a method for real-time analysis of popular sentiment toward presidential candidates, as stated on Twitter during the 2012 U.S. election [38]. Many readers value social media platforms such as Twitter and Facebook because they enable users to easily discuss and express their thoughts on various issues and send messages worldwide. Every day, millions of quick messages (tweets) are shipped on Twitter, making it one of the most prominent microblogging and social networking services. The challenge of assessing users' tweets in terms of emotions, ideas, and viewpoints in a wide range of activities and domains is addressed by Twitter sentiment analysis [39]. Twitter sentiment analysis is a tricky issue, even though sentiment analysis has recently garnered significant popularity in various fields. Businesses use sentiment analysis to look at customer reviews of their services, while the government and other organizations use it to observe public health and forecast political trends, among other things. Manual procedures were frequently used for this before the advent of social networks. Manual techniques were frequently used for this before the beginning of social networks [40]. Customers who want to buy a quality product based on user reviews at reliable online retailers can take advantage of the service offered by the Twitter Sentiment Analysis for Business Project. Additionally, it benefits commercial enterprises that wish to accelerate their firm's growth by revamping goods or services following customer demands and preferences [41]. One of the most crucial tasks in text mining is automatic sentiment analysis of texts, which aims to establish if a given document has a positive, negative, or neutral attitude. Nowadays, It has gotten a lot of attention because opinion mining on microblogging sites is widespread. Both public and commercial companies increasingly rely on the ability to determine whether a written document has a favorable or unfavorable impression [42]. SVM, LR, NB,

and KNN are the techniques employed for sentiment analysis of how people feel about moving the nation's capital. As measured by the performance evaluation algorithm's accuracy, precision, recall, and F-measure findings, the support vector machine fared better than the other three algorithms. The administration will receive an overview of public opinion from the perspective of data from social media through sentiment analysis of the dialogue surrounding the relocation of the nation's capital [43]. With the development of Web 2.0 and the rising popularity of social media platforms like Face book, Twitter, and Google+, users can now exchange information and, as a result, have a say in the material published on these platforms[44]. Many analysts, businesspeople, and politicians use blogs, microblogs, social networks, and other types of websites as a massive source of information to grow their businesses by taking advantage of the copious amounts of text produced by users who provide ongoing feedback on the prominence of a particular subject through emotions, viewpoints, and feedback. The last ten years have seen a growth in sentiment analysis methods, mainly when applied to tweets, in both the academic and industrial areas [45]. Twitter sentiment analysis is a useful tool for various activities using Twitter-based analysis. Twitter sentiment analysis using attentional-graph neural networks is done by the Attentional-graph Neural Network-based Twitter Sentiment Analyzer (AGN-TSA). A three-layered neural network used by AGN-TSA combines information from the tweet's text and its user connections [46]. Recurrent neural networks (RNNs) are also used to examine sentiment in tweets. The technique can categorize tweets with an accuracy rate of 80.74% while considering a binary task, following testing 20 various design strategies [47]. Furthermore, A brand-new unsupervised learning framework built on Concept-based and Hierarchical Clustering is suggested for Twitter Sentiment Analysis. Serial ensembles combine common hierarchical clustering techniques, including single linkage, complete linkage, and average linkage algorithms. Additionally, TF-IDF performs better than the Boolean method compared to other feature representation techniques also examined [48]. One of the most intriguing study areas is sentiment analysis from Twitter [49]offered a comparative emotive analysis from a different linguistic standpoint. Several models, such as PLMs, RoBERTa, and BERT, were used to evaluate each language separately. For fine-tuning, four Nigerian languages were pooled in the training dataset. Among the models tested in the proposed system, the AfriBERTa model fared the best. At present, politics and virality in social media is a common phenomenon. This study [50] overviewed sentimental analysis from Twitter in Greece, Spain, and the United Kingdom. The dataset originated from tweets of parliament members and politicians. The experimental models are distinguished into two types multilingual and Monolingual. A comparative analysis was conducted between these regions. Another study [51] used NB Classifier to determine the tweet's word semantic analysis. The proposed system collects data from live-fetched twitter according to user input. After analyzing the stored user data, the visual heat is represented through Google Maps. There is a variety of methods that may be used to analyze sentiment. Both automatic and hybrid approaches were used for the experiment. Using Deep learning [52] performed, the analysis of sentiment on stemmed Turkish Twitter user data. Twitter API was used to create the dataset. The data was preprocessed by eliminating tokens, id numbers, and punctuation marks, among other

things. The training dataset was followed by three distinct techniques: shift, shuffle, and hybrid. As DL algorithms, CNN, RNN, and Han architecture were used. Compared to previous systems, the suggested system can acquire a modest advantage. The research [53]presented a one-of-a-kind effort on sentiment inspection based on online meal delivery. Food reviews are well-known for their commercial worth. It isn't just about Twitter users; it's also about evaluating business competitors. A comparative analysis was done on Twitter user data connected to food delivery to categorize users' attitudes and customer reviews regarding the organization. Lexicon-based categorization was used to differentiate sentiment polarity.

The literature focuses on the research's shortcomings, gaps, and advancements. A lot of studies have been done on the broad topic of emotion detection from text. But something needs to be improved. Accuracy is one among them. Therefore, the suggested model is created to increase accuracy. There isn't much work left over from the Russia-Ukraine war. This study looks into how people feel about going to war. The research also examined racist behavior. Since we lacked a labeled dataset, we used the Python library to evaluate the data and used unsupervised learning to determine sentiment. Racism originating from the war is detected via semi-supervised learning. The literature demonstrates that sentiment varies according to the domain.

III. METHODOLOGY

Social media sites like Facebook detect emotion. Another dataset was obtained from Kaggle and was used to identify racism. Finally, we collected a dataset from Twitter. The Twitter dataset is about the conflict between Russia and Ukraine. This dataset was used to identify racism and assess public opinion on the Russia-Ukraine conflict. The datasets' source and application are given in Table I. Overview of the entire system is shown in Fig. 1. As previously stated, dataset 1 was acquired via social media and manually classified. After that the dataset is divided into two parts. During the training phase, 18000 data were trained, whereas 3000 data were tested. Fig. 2 depicts the number of classes in dataset 1. Fig. 3 illustrates dataset 2, where 0 denotes a non-racist speech, and one represents a racist utterance. Despite the recent and rising interest in utilizing Twitter to study human behavior and attitudes, the capacity to use Twitter data for social science research and scientific research still has a long way to go[54]. Twitter provides an API, and by using the API, data can be collected for research [55]. We used Twitter API to collect tweets about the Ukraine-Russia crisis. Our dataset 3 had a total of 16,208 records. We used some keywords to find the exact data we needed. Fig. 4 depicts the stages involved in gathering data from Twitter. The ensemble approach is used to address the challenges with emotion detection accuracy. The ensemble technique combines a variety of models. The accuracy rises when the ensemble approach is used. First, we used the suggested model to analyze our dataset, and the accuracy was higher than with the old approach. Later, we applied the model to additional datasets used by academics. The accuracy of the model greatly increases whenever we apply it to new datasets.

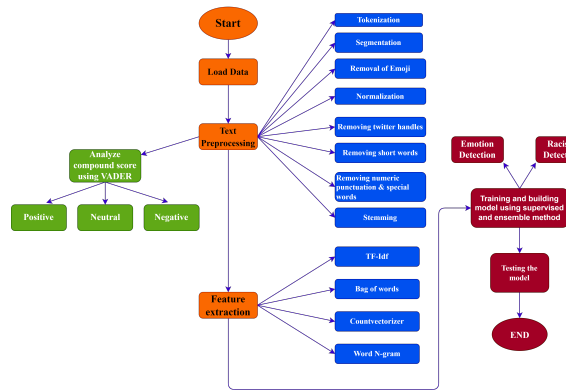


Fig. 1. Overview of the entire system

TABLE I. DATASET SUMMARY

Dataset	Application	Source
Dataset 1	Detect Emotion	Social media(self made)
Dataset 2	Racism Detection	Kaggle
Dataset 3	Sentiment Analysis Racism Analysis	Twitter(Self made)

A. Text Representation

1) *Text preprocessing*: Text preprocessing prepares text data for machines to use for analysis and prediction. Data preprocessing is an essential aspect of machine learning. The algorithms and models need to forecast the result accurately. Besides lowering the extracted feature space, preprocessing can increase classification accuracy [56]. The neat text and some NLP packages are used to text preprocess. The text preprocessing technique applied in our work:

Removing Numeric, punctuation, and special character: We deleted any superfluous punctuation, digits, or symbols to enhance the dataset and replaced them with space. Also, to make the dataset more algorithmically implementable.

Removing Twitter Handles (@user): A was created to eliminate the undesirable pattern of text from the tweets. It takes two arguments, the original and refined strings, which remove the text patterns. The method will give the desired string after removing unwanted text patterns in the output. To withdraw the design from our data, we used the function.

Removing short words: Short and useless words are deleted in this stage. Words with a length of three or fewer were eliminated from this list. As like hmm, oh, and so forth.

Removal of emoji, URL and hashtag: We must remove emojis from text to gain accurate output in our work. Therefore, we also eliminated URLs and hashtags from our text.

Segmentation: Text segmentation is the technique of splitting written material into valuable components, including words, sentences, or themes. We applied the segmentation technique. First, any kinds of commas, dots, and hyphens were deleted.

Tokenization: The tokenizer breaks up the input text into little chunks known as tokens. For example, if elements of a word are more prevalent than the word itself, there can be more tokens than words.

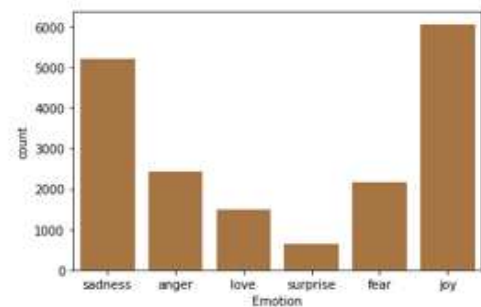


Fig. 2. Number of each class in the dataset emotions

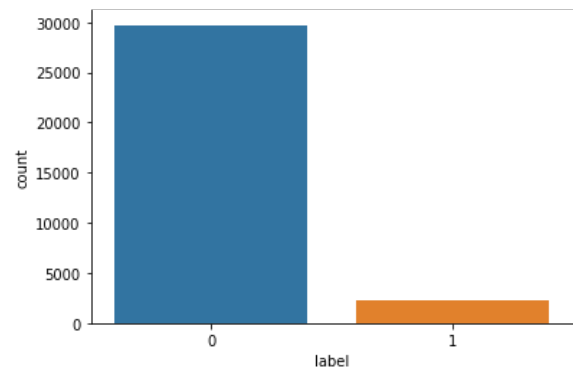


Fig. 3. Number of class in the dataset of racism detection

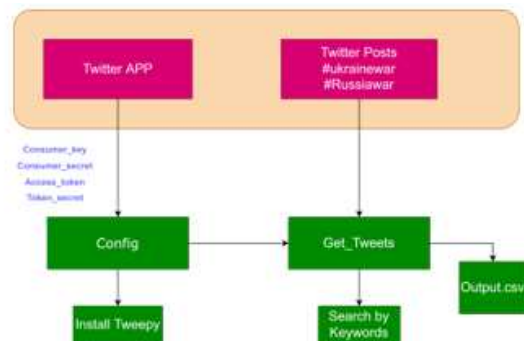


Fig. 4. Process of data collection from Twitter

Streaming: Data streaming is sending a continuous stream of data to extract useful information. A data stream is a collection of pieces arranged in time. Stemming is the process of removing suffixes (“,” “ly”, “es”, “s”, and so on) from a word using a set of rules. For Example, play, playing, and player, are the different variations of the word – “play”.

Lemmatization: It seems the term yet ensures that it retains its meaning. Lemmatization has a pre-defined vocabulary that records the context of words and checks the word in the dictionary while decreasing the term’s length. Twitter users occasionally utilize lengthy words, such as “loooovveee, greeeeat”, when they purposely type or add more letters that

TABLE II. EXAMPLE OF BAG-OF-WORDS FEATURE

She	He	Good	Girl	Will	Guy	
1	1	2	1	0	0	B1
0	0	1	0	1	1	B2

repeat more frequently. We replace the long and needless character

2) *Feature extraction*: The most crucial phase in detecting emotion is feature selection, which influences the task's overall outcome. Therefore, it is critical to choose features carefully to improve classification because better feature selection leads to accurate prediction. We used various features to analyze our procedure data after completing the pre-processing step. We applied several feature extraction techniques to find the suitable one for our proposed system. Combining several strategies to achieve the best potential result is beneficial [17]. Feature extraction techniques applied:

Bag-of-Words Features: An approach for extracting characteristics from text documents is a Bag of Words. Machine learning algorithms may be trained using these features. It develops a vocabulary of all the unique terms in the training sets. Take a corpus, for example, shown in Table II (a textual compilation) named A comprising B documents b1, b2, bB, with X unique tokens retrieved from the corpus A. The list of X tokens (words) will be formed, and B X X will determine the size of the bag-of-words matrix Y. In document B(i), the frequency of the token is represented in each row of the matrix Y. For instance, suppose we have two documents-

B1: She is a good girl. He is also good.

B2: Will is a good guy.

It starts by building a vocabulary from the documents' unique terms. ["She", "He", "good", "girl", "Will", "guy"] Here, B=2, X=6. The 2 X 6 matrix Y will be expressed as follows:

The scenario above shows the training characteristics, which include the frequency of each word on each page. This's strategy known as the bag-of-words approach since it's based on the number of occurrences rather than the sequence or order of words.

Word n-grams: Several component evaluation approaches are available for AI-based, more aesthetic implementation, such as n-gram, tf-idf, count vectorizer, and word integration[57]. The n-grams feature combines n consecutive words or characters that are considered beneficial for categorizing texts. Here, we looked at the performance of uni-gram and bi-gram to acquire the optimum model. To achieve the entire situation, we integrated the bi-gram and tri-gram features, providing a substantially superior outcome in our study. We also observed scikit-learn to explore features [58]. A series of N words or characters is referred to as an n-gram. Consider the following sentence: **"Emotion detection from text"**.

Unigram model(n=1)

["Emotion", "detection", "from", "text"]

Bigram model(n=2)

["Emotion detection", "detection from", "from text"]

In natural language processing, n-grams are an essential but fundamental notion. N-grams play a large part in our system since so many applications need to extract insights from text. The n-gram parameter's range affects the outcome.

Tf-idf Vectorizer: The term frequency multiplied by the inverse document frequency is known as tf-idf. The tf-idf vectorizer [59], which evaluates the presence of a word in a document rather than using only raw numbers, is an excellent example of a suitable input representation [60]. The term frequency refers to the number of times a specific word occurs in a conferred document. On the other hand, Inverse document frequency considers all that include that term. **Term Frequency (TF)**: First, let's define the term "frequent" (TF). It is a metric for the number of times a phrase, t, appears in a document, d:

$$tf_{t,d} = \frac{n_{t,d}}{\text{Number of terms in the documents}} \quad (1)$$

Here, The number of times the phrase "t" appears in the document "d" is represented by n in the numerator. As a result, each paper and word would be assigned a unique TF value. We used the same vocabulary as in the Bag-of-Words model to demonstrate how to compute the TF for B2. B2: Will is a good guy.

Here,

- Vocabulary: 'She', 'He', 'good', 'girl', 'Will', 'guy'
- Number of words in B2 = 5
- TF for the word 'Will' by equation (1)= 1/5 Similarly,
- TF('She',B2) = 0/5
- TF('He',B2) = 0/5
- TF('good',B2) = 1/8
- TF('girl',B2) = 0/8
- TF('guy',B2) = 1/8

Inverse Document Frequency (IDF): The IDF is a metric that determines how essential a phrase is. Therefore, the IDF value is required since simply computing the TF is insufficient to comprehend the significance of words.

$$idf_t = \log \frac{\text{number of documents}}{\text{number of documents with term "t"}} \quad (2)$$

he IDF values for all the words in B2 are calculated by equation 2:

$$IDF('Will') = \log(2/1)$$

TF-IDF: The TF-IDF score for each word in the corpus is now computed. Higher-scoring words are more important, whereas lower-scoring words are less significant. Calculate the tf-idf score:

$$(tf_{i,d} \cdot idf_{i,d})_{t,d} = tf_{t,d} \cdot idf_t \quad (3)$$

Countvectorizer: The scikit-learn module in Python provides a fantastic feature called count vectorizer. The count vectorizer employs a bag-of-words method that avoids textual structures. Instead, only word counts are used to extract information. So, First and foremost, each document will be converted to a vector format. Then, the vector's input measures how many times each word appears in the content [61]. Finally, we combined n-grams with Countvectorizer techniques [62].

B. Proposed Work

We proposed a model for detecting emotion in text. Our model will also see racism. We'll look for and examine racism in the context of the Ukraine-Russia conflict. The public's viewpoint toward the competition is also investigated.

1) *Detect emotion from text:* The dataset description and data preparation techniques, such as text preprocessing and feature extraction, have previously been covered. The dataset is divided into two steps for identifying emotion. During the training phase, 18000 data points were learned, whereas 3000 data points were tested. We used classical machine learning as well as an ensemble model. In training the model to recognize emotion, many classifiers were utilized. Six primary emotions are detected by the system (anger, fear, surprise, joy, love, and sadness). We'll go through the classifier that we used in our system later. Fig. 5 presents the architecture for emotion recognition.

2) *Racism detection:* Racism was detected using data from Kaggle (dataset 2) and Twitter (dataset 3). The Kaggle dataset has labels, whereas the Twitter dataset does not. We combined the two datasets and utilized text preparation and feature extraction techniques. The boosting approach, as well as a machine learning classifier, were used. One-third of the data is tested, and the rest is trained. The best prediction model was determined after using all approaches and algorithms. The model investigated racism in dataset 3, which concerned the Ukraine-Russia war. In the suggested process for detecting racism, the following algorithm was utilized to categorize tweets as racist or non-racist.

Algorithm 1 Racism Detection Thresholds

```
if prediction  $\geq$  0.3 then
    label  $\leftarrow$  1
else
    label  $\leftarrow$  0
end if
```

Where 1 implies racism and 0 denotes non racism. The framework for detecting racism is depicted in Fig. 6.

3) *Sentiment analysis:* The present research's ultimate job is sentiment analysis. Dataset 3 was used to investigate the sentiment. We want to look at how people feel about the Ukraine-Russia conflict. Dataset 3 is obtained from Twitter and is on the conflict between Ukraine and Russia. On Twitter, a tweet is a microblog message. It's only allowed to be 140 characters long. Most tweets include text, embedded URLs, photos, and usernames. There are also misspellings in them. As a result, several preprocessing processes were performed on the tweets to eliminate unnecessary data. The reasoning is that

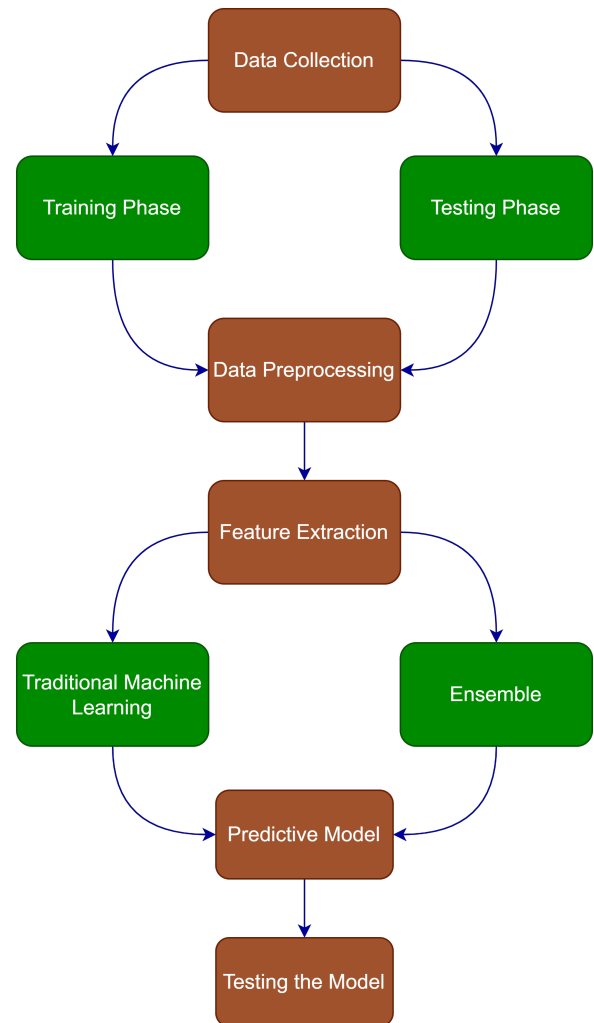


Fig. 5. Emotion detection architecture

cleaner data is more suitable for mining and feature extraction, resulting in more accurate results. We didn't need to use all of the preprocessing approaches outlined above since VADER can understand the sentiment of a text that incorporates emoticons, slang, conjunctions, capital phrases, punctuation, and other idioms. We avoided misspellings, short words, and unusually lengthy letters.

Algorithm 2 Sentiment Score Range

```
if compoundscore  $\geq$  0.05 then
    sentiment  $\leftarrow$  positive
else if compound_score  $\leq$  -0.05 then
    sentiment  $\leftarrow$  Negative
else if then
    sentiment  $\leftarrow$  Neutral
end if
```

We compute a compound score to analyze sentiment. After assessing the sentiment, we determined the positive, negative, and neutral phrases used in the war. The result and discussion section will provide the top terms used negatively

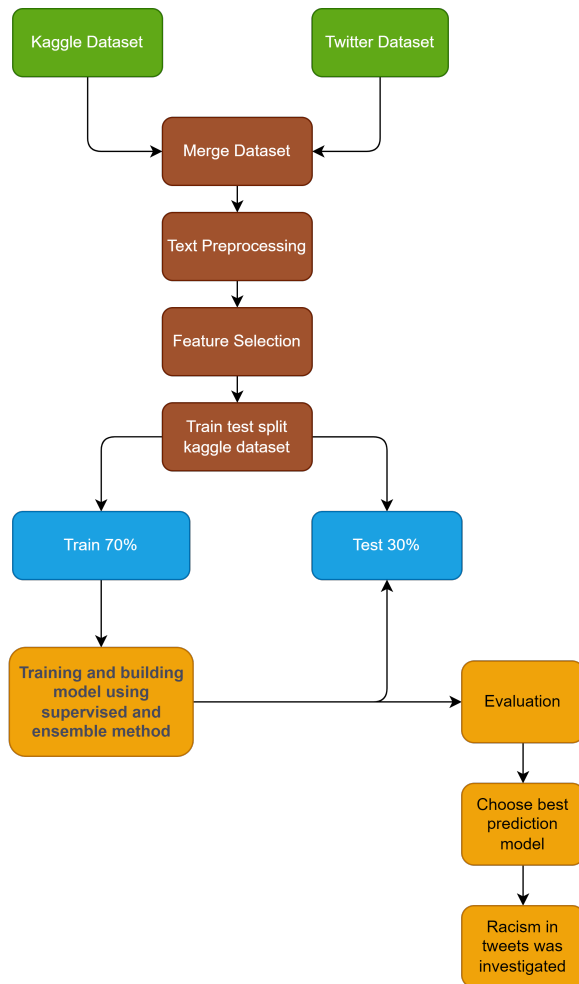


Fig. 6. Racism detection architecture

in the conflict. In addition, positive and neutral words will be highlighted.

C. Proposed Model

1) *Supervised machine learning classifiers*: The learning process is divided into training and testing phases. First, training data samples were used as input, and then the learning algorithm was applied to learn the feature and build the learning model [63]. Finally, for prediction for the data or test, the learning model executes in the testing phase. The approach of supervised learning aims to train the machine. For each of the algorithm, we used parameters based on need. Fig. 7 depicts the model in detail. The dataset is first split into two phases: training and testing. The data were prepared using a feature engineering technique. In order to forecast the label, the baseline method was used. The test data was finally employed to predict the value and contrast it with its actual value. Finally, a number of evaluation criteria are used to gauge performance. Some of the used supervised algorithms in this system:

Decision Trees: A DT [64] is a classifier defined as a recursive split of the instance space. The DT consists of nodes that form a rooted tree, a dispersed tree with no incoming

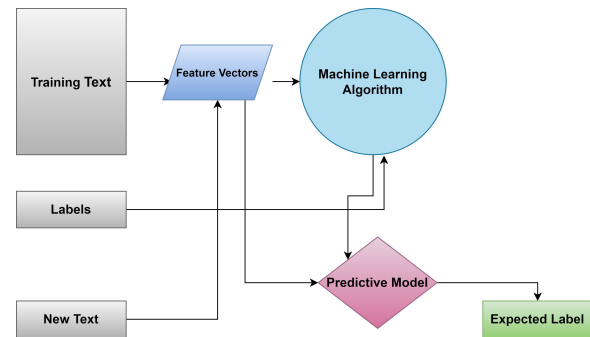


Fig. 7. Supervised learning model architecture

edges, and a root node. Each node has one incoming advantage. An internal node has outgoing edges, whereas leaves are the other nodes. Based on the input values, each internal node in a decision tree separates into multiple sub-spaces. In the simplest case, each test evaluates a single attribute, and the occurrence space is partitioned according to the attribute's value. Every node has a label representing the property being tested, and its branches have starting values. Mini_samples_split, max_depth and class_weight were used as parameter.

Linear Regression: The purpose of LR, a portion of the regression algorithm family, is to discover the interrelationships and dominance of variables. Regression analysis is used to forecast a target variable; predicting an attribute from a limited set is one topic of classification.LR [65] pertains to the supervised learning algorithms area as well.

Naive Bayes: Another supervised learning approach and statistical classification method is Bayesian classification [66]. The Bayesian classification's primary goal is to tackle problems with prediction and solve prediction difficulties. This classification may mix observable data and delivers effective learning techniques. Learning techniques and algorithms have been simplified to understand and analyze using Bayesian classification.

Logistic Regression: LR is the most powerful statistical and data mining technique [67]. LR has several key benefits, including the ability to generate probabilities naturally and the ability to handle multinomial classification issues. A further advantage is that far too many LR model analysis techniques rely on the same fundamentals as linear regression. We determine the likelihood that the output variable's perception corresponds to the proper category using LR [68]. The LR model is used to categorize emotions from inside the input text [69]. With training and testing sets, LR classifies text into several emotion categories [70]. Logistic regression is a broader version of linear regression. Consider the linear regression equation below:

$$x = x_0 + x_1z_1 + x_2z_2 + \dots + x_nz_n \quad (4)$$

Here x is the response variable, and the predictor variables are Z_1, Z_2, Z_3 , and Z_n . So if we apply a sigmoid function to equation 1, we will get a logistic function.

$$l = \frac{1}{1 + e^{-(x_0 + x_1z_1 + x_2z_2 + \dots + x_nz_n)}} \quad (5)$$

Random Forest: RF is among the technique in Machine Learning used to process the data in vast quantities and widely used algorithm based on the results of a DT formed throughout training [71] [72]. The RF algorithm combines numerous trees with training data to generate greater accuracy. The concentrated output from every DT seems to be the forest's output. Each tree has a different set of features and uses typical tree-building methods, and these properties generate the nodes and leaves [73].

Support Vector Machine: SVM is a pattern identification and categorization technique that is pretty simple to use [74]. SVM is among the most extensively utilized cutting-edge machine learning technology (SVM). The SVM approach is related to supervised learning, which requires feature extraction and produces desired results. SVM has the benefit of being very easy to execute and scaling vast amounts of data more efficiently than neural networks [75]. The kernel approach allows SVMs to effectively conduct nonlinear classification by implicitly translating their inputs within high-dimensional feature regions [76]. SVM finds the biggest and best margin hyperplane to categorize the text into numerous emotion classifications. Choosing an appropriate kernel function to get better results is critical because it determines the transformed feature space in which the training set instances will be categorized. Some well-known kernels were used:

$$\text{Linear} : K(A_i, A_j) = A_i^T X_j \quad (6)$$

$$\text{Polynomial} : K(A_i, A_j) = (\gamma A_i^T X_j + r)^d, \gamma > 0 \quad (7)$$

The kernel parameters are

$$\gamma, r, d \\ A_i$$

is a training vector that the function maps into a high-dimensional space. The linear kernel used sigmoid in the research and the n_samples were 100.

2) *Ensemble method:* Ensemble learning is a procedure that involves integrating many models or classifiers to address a specific issue. Both machine learning and deep learning techniques can benefit from the ensemble approach. It is followed to attain the intended aim and improve system performance. We applied bagging and boosting in our suggested method to discover the outcome. Though many existing systems use the ensemble approach, we attempted to construct this system using a machine learning-based ensemble model. AdaBoost, Gradient boost, and XGBoost have been used to develop and evaluate overall system performance. AdaBoost uses a recursive technique to help poor classifiers better by learning from their mistakes. Although XGBoost is identical to AdaBoost, it outperforms AdaBoost in highly condensed and sophisticated data and system optimization. Ensemble learning is commonly implemented using decision trees, which aid in the solution of quantitative issues. Ensemble learning computes the final classification based on the ensembled findings given by decision trees rather than relying on a single decision tree's predictive analysis and outcome. In our ensemble learning, we used a variety of machine learning classifiers, including LR, SVM, RF, and others, in addition to the decision tree, to train and analyze the model and find the best output. We used

TABLE III. HYPERPARAMETERS OF THE XGBCLASSIFIER

Types	Values
learning_rate	0.2
max_depth	10
n_estimators	200
subsample	0.8

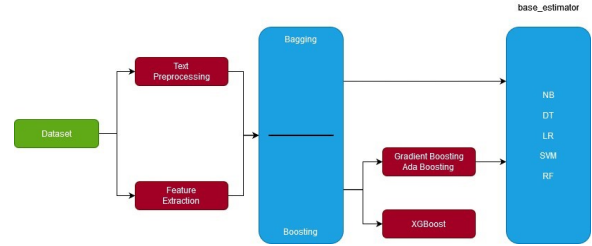


Fig. 8. Proposed ensemble model

several n estimators (10,50,100) to improve performance because the ensemble approaches heavily depend on parameters. Fig. 8 depicts the architecture and base estimator used in our ensemble model. Soft voting and Hard voting classifiers are two types of voting classifiers used to create the results [77]. The test preprocesses ing approach and feature extraction are completed initially. When we deployed the XGB classifier of the boosting ensemble, we employed the grid search parameter tuning strategy to choose the optimum hyperparameters. Table III displays the XGB classifier's hyperparameters.

3) *VADER:* VADER is an analytical program that uses both analyses of lexical and rules-based to analyze sentiment. It generally employs a combination of strategies that classify lexical characteristics according to their semantic orientation, which might be positive or negative. This approach not only detects the polarity of positive or negative attitudes but also determines how the sentiment is [78]. This tool is consistent and has a higher accuracy ratio for performing sentiment analysis from social media. There are several advantages of VADER. It works well with social media content while quickly summarizing various topics. No need for training data though it includes defined vocabulary for analyzing sentiment. Because of its speed and precision, it may be utilized for asynchronous data. The VADER packages display a positive, neutral, or negative value from individual tweets [79]. We used VADER to analyze the sentiment. First, a compound score is calculated based on these values. Next, a compound score is calculated based on these values. The compound score thresholds were used to categorize the sentiment into negative, positive, and neutral, as indicated in Algorithm 2. However, VADER performs better for social media data which is unsupervised learning.

IV. RESULT AND DISCUSSION

The accuracy of all the models is calculated. After finding the best model, we compute each class's precision, recall, and f-score. Combining word-n-gram with a count vectorizer is the most outstanding result to detect emotion. We employ bigrams and trigrams in our implementation. The accuracy could be determined as described in the equation since accuracy was chosen as the assessment criteria, where TP = True Positive,

TABLE IV. CLASSIFIER PERFORMANCE OF SUPERVISED MODEL

Classifier	Training Result	Testing Result
Logistic Regression	0.99	0.88
Decision Tree	0.99	0.86
Naive Bayes	0.96	0.76
SVC	0.98	0.86

TABLE V. BOOSTING METHOD PERFORMANCE(EMOTION DETECTION

Algorithm	Training Result	Testing Result
ADABOOST	0.99	0.89
GradientBoost	0.96	0.78
XGBoost	0.97	0.90

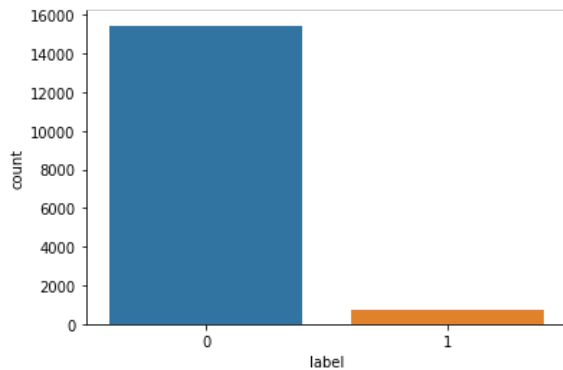


Fig. 9. Racism found in dataset 3

TN =True Negative, FP = False Positive, and FN = False Negative.

The performance of the base model was shown in Table IV. The best testing outcome came from LR. The testing results are the same for DT and SVC. However, NB has the worst results, and overfitting problems are a big deal here. From Table V, XGBoost achieves the desired outcome. Gradient boosting performs the poorest out of all methods, with Adaboosting marginally outperforming the baseline method. The performance of the supervised classifiers for identifying emotion is displayed in Table IV. Since we utilized the supervised model and ensemble model bagging and boosting, Table V shows boosting model performance, while Table VII shows all classes' precision, recall, and f-score. Table VI shows the overall performance of our applied mode to detect emotion. Our technology outperforms the prior method when it comes to seeing emotion. Finally, we compared our model to some of the most recent works in emotion recognition. Table VIII compares our model with others. The accuracy of the proposed ensemble model is higher than that of earlier systems, as demonstrated in Table VIII. Accuracy for gloves, SVM, and LR is under 80%. C-BiLSTM, however, achieved 88% accuracy. The suggested technique, meanwhile, achieves 90% accuracy.

We only use the F-score to detect racism. We used F-score instead of accuracy since we can see how unbalanced our dataset 2 is in Fig. 3. Compared to the values with labels:1, we can observe that the deals with label:0 have a far higher number of differences. As a result, if we use accuracy as our assessment criterion, we may meet a significant proportion

TABLE VI. OVERALL PERFORMANCE OF APPLIED METHOD

Model	Testing Result	Training Result
Supervised	0.88	0.99
Bagging	0.89	0.99
Boosting	0.90	0.97

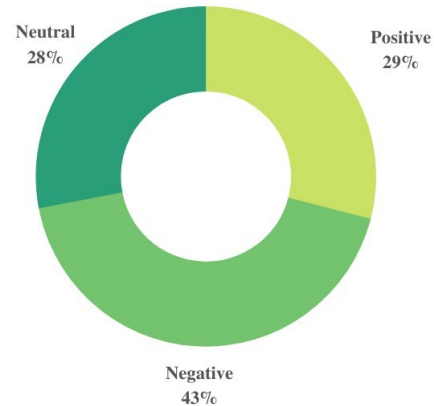


Fig. 10. Sentiment result

TABLE VII. PRECISION, RECALL, F-SCORE OF ALL CLASSES

Emotion	Precision	Recall	F-score
anger	0.90%	0.90%	0.90 %
fear	0.90%	0.83%	0.86 %
joy	0.90%	0.93%	0.91 %
love	0.80%	0.71%	0.75 %
sadness	0.93%	0.94%	0.94 %
surprise	0.67%	0.74%	0.72 %

TABLE VIII. COMPARISON OF RESULT WITH OTHER PAPERS

Ref.	Method/Algorithm	Result
[2]	Glove	F1-69.3 %
[14]	SVM	Accuracy-63.5 %
[30]	C-BiLSTM	F.score-88 %
[33]	LR-SGD	Accuracy-79 %
Proposed system	Ensemble (Boosting)	Accuracy-90.45%

of false positives. The percentage of our relevant findings is referred to as precision. The rate of relevant total results accurately categorized by our algorithm is called recall. We are constantly faced with a trade-off between accuracy and recall, with high precision resulting in low memory and vice versa. F-core is calculated using equation 5. Both bag-of-words and tf-idf perform best when it comes to detecting racism. Table IX shows the results of the racism detection.

$$\text{F-score} = 2 * \frac{\text{Precision} + \text{Recall}}{\text{Precision} * \text{Recall}} \quad (8)$$

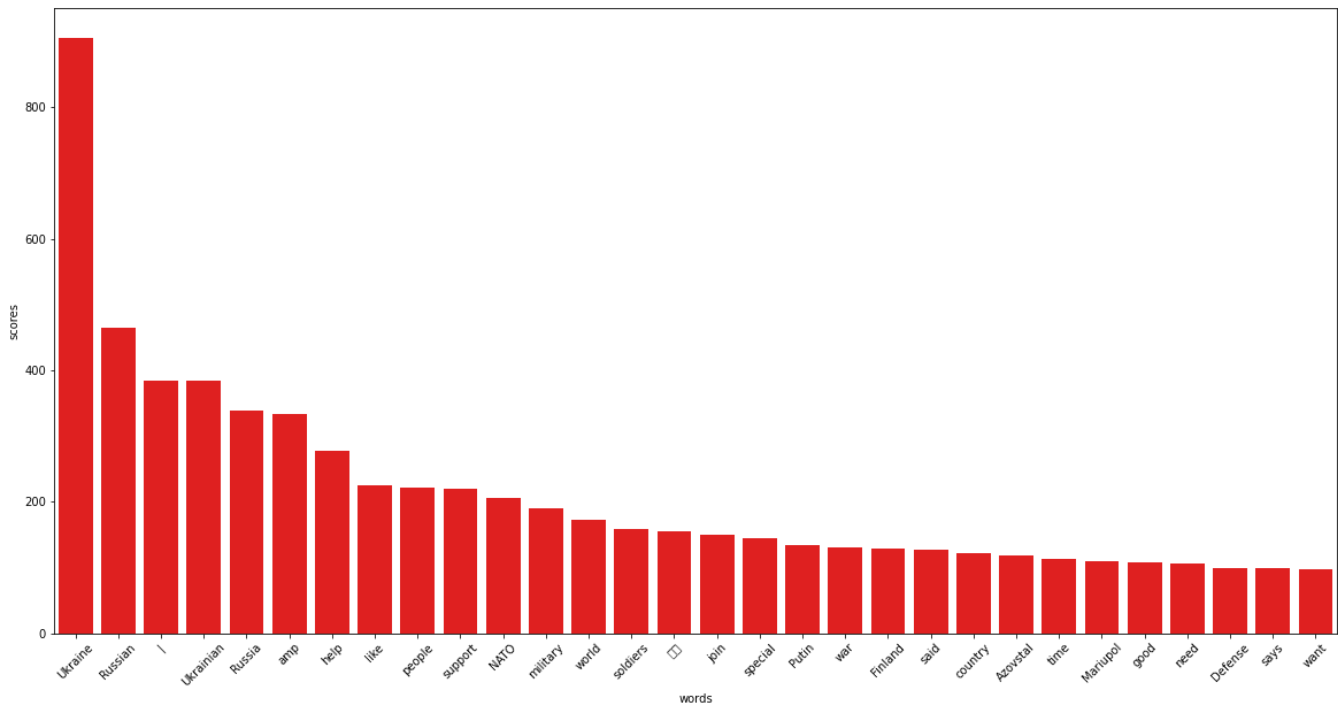


Fig. 11. Most common positive words

TABLE IX. RACISM DETECTION ACCURACY

Algorithm / Model	Feature Extraction	F-score
Logistic Regression	Bag-of-Words	0.552
Logistic Regression	Tf-Idf	0.560
XGBoost	Bag-of-Word	0.563
XGBoost	Tf-Idf	0.547

We chose our best model for identifying racism and used it to examine racism in tweets on the Ukraine-Russia conflict. According to our research, people are less racist regarding this war.

The outcome is shown in Fig. 9. The figure shows that to examine the racism in dataset 3, people are less racist about the war—15459 sentences that are non-racist and 749 sentences that are racist. Finally, when it comes to sentiment analysis, the results show that negative sentiment was identified in 7002, the positive feeling was found in 4639 tweets, and neutral attitude in 4567. Fig. 10 depicts the sentiment discovered in dataset 3.

A. Discussion

We integrate the best model of our used procedures into our website after obtaining it. When a user registers on the website, a unique user name is assigned to them. Then, our system can discern their emotion using that person's textual data. The person's emotional state is kept in the database after the emotion is detected. The system can only store an individual's feelings and restore their output over time. As a result, anyone can examine their feelings throughout time. The system can display a person's emotions at any time. However, in the future, we will expand the system to automatically analyze a person's

feelings and warn the user of any emotional changes. A future study could look into the dynamic alterations.

We have had some issues with the application of detecting racism.

The accuracy of racism detection was low, as seen in the results section. We discovered that an unbalanced dataset caused the problem after further investigation. We detected more nonracist text in our dataset than racist text. When a dataset is unstable, it refers to the data dominance space vs. the dispersed data. Unbalanced learning challenges can be divided into two categories. There are two types of relationships:

- 1) between classes and
 - 2) inside classes.
- Between the classes: This refers to the unequal distribution of data samples between two classes.
 - Inside the class: when the classes of data differ between more than one idea, the unbalanced problem is considered inside the class.

As a result, our dataset issue was sandwiched between the problem of class imbalance. Oversampling and undersampling approaches are two techniques for sorting this problem [80]. Future studies can handle these issues. If we have a sizeable balanced dataset, we can use deep learning in the future. In addition, we may address the issue of data imbalance to achieve better results. The globe is in the grip of the Ukraine-Russia conflict. We attempted to overserve the sentiment of the general public. Geographical sentiment analysis can help decision-makers understand a particular country's sentiment. It will aid in taking any measures necessary to end the problem. Positive, negative, and neutral words are illustrated in Fig.

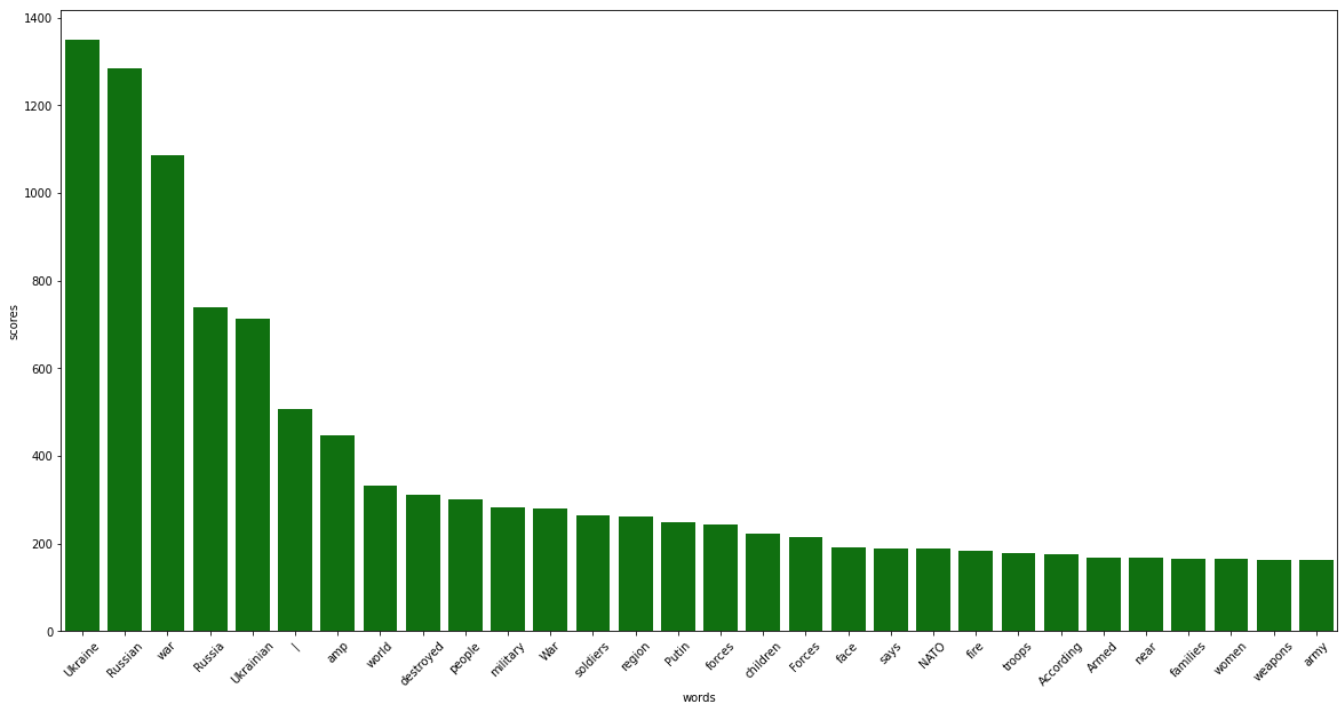


Fig. 12. Most common negative words

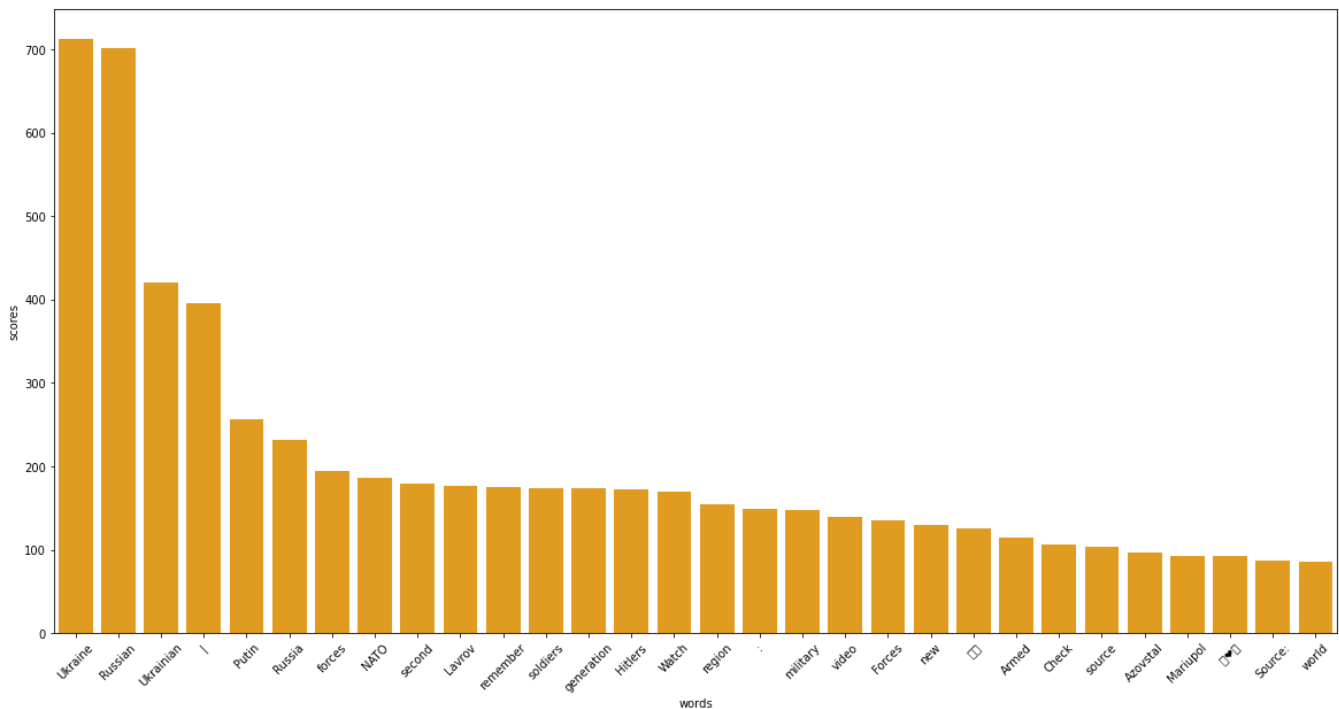


Fig. 13. Most common neutral words

11, 12, and 13. The most frequently used words are Ukraine, Russia, and Ukrainian, which can be used in positive, negative, or neutral contexts. Every term has a mixed connotation, meaning that some people will interpret it positively while others will interpret it negatively. Therefore, more research is required, and each word can be given as a bigram or trigram.

The most common words in the study are visually depicted. Therefore, a thorough analysis is advised.

V. CONCLUSION

This study was designed to accomplish several goals. This study first tries to assess an individual's emotional trajectory through time and to recognize emotion from textual format using machine learning techniques. The ensemble technique works best for detecting emotion. However, there were various difficulties in determining the best accuracy from our dataset. We decided on the ensemble model's optimal accuracy by integrating several data preparation methods. XGBclassifier demonstrated 90% accuracy in identifying the mood in the text.

There are several applications for emotion detection, including marketing, customer reviews, the detection of psychological instability, and others. XGBclassifier and Logistic Regression perform best in identifying racism. However, the ability to recognize racism has to be improved. In our study, we looked at how individuals felt about the conflict between Ukraine and Russia. Nearly half of the Twitter users had unfavorable feelings (43%). The study of this conflict had not received much attention from artificial intelligence researchers. The professional can remedy the problem if we further investigate this. No system is flawless, after all, and we are fully conscious of the constraints of our system. We are looking to implement this system using multiple or large datasets, and deep learning techniques will be applied in the future. Our system comparatively works better than the previous systems to detect emotion. The novelty and future development have already been explored. We hope that this study will contribute to further research and influence the field of emotion recognition, racism detection, and sentiment analysis.

AUTHOR CONTRIBUTIONS

Abdullah Al Maruf: Implementation; Writing: Review and editing; Writing: Original draft; Data collection. Zakaria Masud Jiyad: Writing: Review and editing; Writing: Original draft; Data collection. Md. Mahmudul Haque: Writing: Review and editing; Writing: Original draft; Data collection. Fahima Khanam: Supervision

REFERENCES

- [1] R. Rahman *et al.*, "Detecting emotion from text and emoticon," *London Journal of Research in Computer Science and Technology*, 2017.
- [2] U. Rashid, M. W. Iqbal, M. A. Skiandar, M. Q. Raiz, M. R. Naqvi, and S. K. Shahzad, "Emotion detection of contextual text using deep learning," in *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2020, pp. 1–5.
- [3] S. A. Salam and R. Gupta, "Emotion detection and recognition from text using machine learning," *Int. J. Comput. Sci. Eng.*, vol. 6, no. 6, pp. 341–345, 2018.
- [4] C. Pool and M. Nissim, "Distant supervision for emotion detection using facebook reactions," *arXiv preprint arXiv:1611.02988*, 2016.
- [5] M. Hasan, E. Agu, and E. Rundensteiner, "Using hashtags as labels for supervised learning of emotions in twitter messages," in *ACM SIGKDD workshop on health informatics, New York, USA*, vol. 34, no. 74, 2014, p. 100.
- [6] M. Hasan, E. Rundensteiner, and E. Agu, "Emotex: Detecting emotions in twitter messages," 2014.
- [7] S. N. Shivhare, S. Garg, and A. Mishra, "Emotionfinder: Detecting emotion from blogs and textual documents," in *International Conference on Computing, Communication & Automation*. IEEE, 2015, pp. 52–57.
- [8] B. Le and H. Nguyen, "Twitter sentiment analysis using machine learning techniques," in *Advanced Computational Methods for Knowledge Engineering*. Springer, 2015, pp. 279–289.
- [9] S. Kuamri and C. N. Babu, "Real time analysis of social media data to understand people emotions towards national parties," in *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. IEEE, 2017, pp. 1–6.
- [10] R. Pandarachalil, S. Sendhilkumar, and G. Mahalakshmi, "Twitter sentiment analysis for large-scale data: an unsupervised approach," *Cognitive computation*, vol. 7, no. 2, pp. 254–262, 2015.
- [11] P. J. Manamela, M. J. Manamela, T. I. Modipa, T. J. Sefara, and T. B. Mokgonyane, "The automatic recognition of sepedi speech emotions based on machine learning algorithms," in *2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)*. IEEE, 2018, pp. 1–7.
- [12] V. U. Kumar, A. Savithri, M. J. Bhavani, A. M. Priya, K. V. S. B. Jahnvi, and N. D. N. Lakshmi, "Finding psychological instability using machine learning," in *2020 7th International Conference on Smart Structures and Systems (ICSSS)*. IEEE, 2020, pp. 1–4.
- [13] A. D. Souza and R. D. Souza, "Affective interaction based hybrid approach for emotion detection using machine learning," in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 2019, pp. 337–342.
- [14] S. Koppu, K. Lakshmana, P. Venugopal, R. K. Poluru *et al.*, "Automated category text identification using machine learning," in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*. IEEE, 2020, pp. 1–4.
- [15] L. Sana, K. Nasir, A. Urooj, Z. Ishaq, and I. A. Hameed, "Bers: Bussiness-related emotion recognition system in urdu language using machine learning," in *2018 5th International Conference on Behavioral, Economic, and Socio-Cultural Computing (BESCC)*. IEEE, 2018, pp. 238–242.
- [16] S. Grover and A. Verma, "Design for emotion detection of punjabi text using hybrid approach," in *2016 International Conference on Inventive Computation Technologies (ICICT)*, vol. 2. IEEE, 2016, pp. 1–6.
- [17] S. Azmin and K. Dhar, "Emotion detection from bangla text corpus using naive bayes classifier," in *2019 4th International Conference on Electrical Information and Communication Technology (EICT)*. IEEE, 2019, pp. 1–5.
- [18] M. Healy, R. Donovan, P. Walsh, and H. Zheng, "A machine learning emotion detection platform to support affective well being," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2018, pp. 2694–2700.
- [19] B. Sun, Q. Ma, S. Zhang, K. Liu, and Y. Liu, "iself: Towards cold-start emotion labeling using transfer learning with smartphones," *ACM Transactions on Sensor Networks (TOSN)*, vol. 13, no. 4, pp. 1–22, 2017.
- [20] O. Rabie and C. Sturm, "Feel the heat: Emotion detection in arabic social media content," in *The International Conference on Data Mining, Internet Computing, and Big Data (BigData2014)*. Citeseer Kuala Lumpur, Malaysia, 2014, pp. 37–49.
- [21] W. A. Hussien, Y. M. Tashtoush, M. Al-Ayyoub, and M. N. Al-Kabi, "Are emoticons good enough to train emotion classifiers of arabic tweets?" in *2016 7th International Conference on Computer Science and Information Technology (CSIT)*. IEEE, 2016, pp. 1–6.
- [22] S. S. Sadeghi, H. Khotanlou, and M. Rasekh Mahand, "Automatic persian text emotion detection using cognitive linguistic and deep learning," *Journal of AI and Data Mining*, vol. 9, no. 2, pp. 169–179, 2021.
- [23] M. Abdullah, M. Hadzikadicy, and S. Shaikhz, "Sedat: sentiment and emotion detection in arabic text using cnn-lstm deep learning," in *2018 17th IEEE international conference on machine learning and applications (ICMLA)*. IEEE, 2018, pp. 835–840.
- [24] M. Baali and N. Ghneim, "Emotion analysis of arabic tweets using deep learning approach," *Journal of Big Data*, vol. 6, no. 1, pp. 1–12, 2019.
- [25] Y. Rao, H. Xie, J. Li, F. Jin, F. L. Wang, and Q. Li, "Social emotion classification of short text via topic-level maximum entropy model," *Information & Management*, vol. 53, no. 8, pp. 978–986, 2016.
- [26] M. Hajar *et al.*, "Using youtube comments for text-based emotion recognition," *Procedia Computer Science*, vol. 83, pp. 292–299, 2016.

- [27] D. Xu, Z. Tian, R. Lai, X. Kong, Z. Tan, and W. Shi, "Deep learning based emotion analysis of microblog texts," *Information Fusion*, vol. 64, pp. 1–11, 2020.
- [28] M. Karna, D. S. Juliet, and R. C. Joy, "Deep learning based text emotion recognition for chatbot applications," in *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*. IEEE, 2020, pp. 988–993.
- [29] M. Jabreel and A. Moreno, "A deep learning-based approach for multi-label emotion classification in tweets," *Applied Sciences*, vol. 9, no. 6, p. 1123, 2019.
- [30] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and S. Khan, "Classification of poetry text into the emotional states using deep learning technique," *IEEE Access*, vol. 8, pp. 73 865–73 878, 2020.
- [31] D. Haryadi and G. P. Kusuma, "Emotion detection in text using nested long short-term memory," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 6, 2019.
- [32] H. Al-Omari, M. A. Abdullah, and S. Shaikh, "Emodet2: Emotion detection in english textual dialogue using bert and bilstm models," in *2020 11th International Conference on Information and Communication Systems (ICICS)*. IEEE, 2020, pp. 226–232.
- [33] A. Yousaf, M. Umer, S. Sadiq, S. Ullah, S. Mirjalili, V. Rupapara, and M. Nappi, "Emotion recognition by textual tweets classification using voting classifier (lr-sgd)," *IEEE Access*, vol. 9, pp. 6286–6295, 2020.
- [34] A. Sarlan, C. Nadam, and S. Basri, "Twitter sentiment analysis," in *Proceedings of the 6th International conference on Information Technology and Multimedia*. IEEE, 2014, pp. 212–216.
- [35] A. Go, L. Huang, and R. Bhayani, "Twitter sentiment analysis," *Entropy*, vol. 17, p. 252, 2009.
- [36] K. K. Pawar, P. P. Shrishrimal, and R. R. Deshmukh, "Twitter sentiment analysis: A review," *International Journal of Scientific & Engineering Research*, vol. 6, no. 4, pp. 957–964, 2015.
- [37] T. Rao, S. Srivastava *et al.*, "Analyzing stock market movements using twitter sentiment analysis," 2012.
- [38] H. Wang, D. Can, A. Kazemzadeh, F. Bar, and S. Narayanan, "A system for real-time twitter sentiment analysis of 2012 us presidential election cycle," in *Proceedings of the ACL 2012 system demonstrations*, 2012, pp. 115–120.
- [39] O. Adwan, M. Al-Tawil, A. Huneiti, R. Shahin, A. A. Zayed, and R. Al-Dibsi, "Twitter sentiment analysis approaches: A survey," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 15, pp. 79–93, 2020.
- [40] M. Bibi, W. Aziz, M. Almarashi, I. H. Khan, M. S. A. Nadeem, and N. Habib, "A cooperative binary-clustering framework based on majority voting for twitter sentiment analysis," *IEEE Access*, vol. 8, pp. 68 580–68 592, 2020.
- [41] J. N. Rao, Y. DurgaMallewar, K. T. Krishna, C. V. Sai, and Y. R. Sri, "Twitter sentiment analysis for different business applications," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 12.
- [42] E. S. Tellez, S. Miranda-Jiménez, M. Graff, D. Moctezuma, O. S. Siordia, and E. A. Villaseñor, "A case study of spanish text transformations for twitter sentiment analysis," *Expert Systems with Applications*, vol. 81, pp. 457–471, 2017.
- [43] E. Sutoyo and A. Almararif, "Twitter sentiment analysis of the relocation of indonesia's capital city," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 4, pp. 1620–1630, 2020.
- [44] D. Jovanoski, V. Pachovski, and P. Nakov, "Sentiment analysis in twitter for macedonian," *arXiv preprint arXiv:2109.13725*, 2021.
- [45] M. Pota, M. Ventura, R. Catelli, and M. Esposito, "An effective bert-based pipeline for twitter sentiment analysis: A case study in italian," *Sensors*, vol. 21, no. 1, p. 133, 2020.
- [46] M. Wang and G. Hu, "A novel method for twitter sentiment analysis based on attentional-graph neural network," *Information*, vol. 11, no. 2, p. 92, 2020.
- [47] S. C. Nistor, M. Moca, D. Moldovan, D. B. Oprean, and R. L. Nistor, "Building a twitter sentiment analysis system with recurrent neural networks," *Sensors*, vol. 21, no. 7, p. 2266, 2021.
- [48] M. Bibi, W. A. Abbasi, W. Aziz, S. Khalil, M. Uddin, C. Iwendi, and T. R. Gadekallu, "A novel unsupervised ensemble framework using concept-based linguistic methods and machine learning for twitter sentiment analysis," *Pattern Recognition Letters*, vol. 158, pp. 80–86, 2022.
- [49] S. H. Muhammad, D. I. Adelani, I. S. Ahmad, I. Abdulmumin, B. S. Bello, M. Choudhury, C. C. Emezue, A. Aremu, S. Abdul, and P. Brazdil, "Naijasenti: A nigerian twitter sentiment corpus for multi-lingual sentiment analysis," *arXiv preprint arXiv:2201.08277*, 2022.
- [50] D. Antypas, A. Preece, and J. C. Collados, "Politics and virality in the time of twitter: A large-scale cross-party sentiment analysis in greece, spain and united kingdom," *arXiv preprint arXiv:2202.00396*, 2022.
- [51] J. Gautam, M. Atrey, N. Malsa, A. Balyan, R. N. Shaw, and A. Ghosh, "Twitter data sentiment analysis using naive bayes classifier and generation of heat map for analyzing intensity geographically," in *Advances in Applications of Data-Driven Computing*. Springer, 2021, pp. 129–139.
- [52] H. A. Shehu, M. H. Sharif, M. H. U. Sharif, R. Datta, S. Tokat, S. Uyaver, H. Kusetogullari, and R. A. Ramadan, "Deep sentiment analysis: a case study on stemmed turkish twitter data," *IEEE Access*, vol. 9, pp. 56 836–56 854, 2021.
- [53] S. K. Trivedi and A. Singh, "Twitter sentiment analysis of app based online food delivery companies," *Global Knowledge, Memory and Communication*, 2021.
- [54] T. H. McCormick, H. Lee, N. Cesare, A. Shojai, and E. S. Spiro, "Using twitter for demographic and social science research: tools for data collection and processing," *Sociological methods & research*, vol. 46, no. 3, pp. 390–421, 2017.
- [55] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, 2010.
- [56] Y. Bao, C. Quan, L. Wang, and F. Ren, "The role of pre-processing in twitter sentiment analysis," in *International conference on intelligent computing*. Springer, 2014, pp. 615–624.
- [57] P. Shah, P. Swaminarayan, and M. Patel, "Sentiment analysis on film review in gujarati language using machine learning," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 1, p. 1030, 2022.
- [58] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.
- [59] F. Pedregosa *et al.*, "sklearn. feature_extraction. text. tfidfvectorizer," 2013.
- [60] S. R. D'souza and K. Sonawane, "Sentiment analysis based on multiple reviews by using machine learning approaches," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2019, pp. 188–193.
- [61] A. Dadhich and B. Thankachan, "Sentiment analysis of amazon product reviews using hybrid rule-based approach," in *Smart Systems: Innovations in Computing*. Springer, 2022, pp. 173–193.
- [62] S. C. Eshan and M. S. Hasan, "An application of machine learning to detect abusive bengali text," in *2017 20th International conference of computer and information technology (ICCIT)*. IEEE, 2017, pp. 1–6.
- [63] V. Nasteski, "An overview of the supervised machine learning methods," *Horizons. b*, vol. 4, pp. 51–62, 2017.
- [64] O. Z. Maimon and L. Rokach, *Data mining with decision trees: theory and applications*. World scientific, 2014, vol. 81.
- [65] C.-J. Lin, R. C. Weng, and S. S. Keerthi, "Trust region newton methods for large-scale logistic regression," in *Proceedings of the 24th international conference on Machine learning*, 2007, pp. 561–568.
- [66] J. H. Friedman and N. I. Fisher, "Bump hunting in high-dimensional data," *Statistics and computing*, vol. 9, no. 2, pp. 123–143, 1999.
- [67] M. Maalouf, "Logistic regression in data analysis: an overview," *International Journal of Data Analysis Techniques and Strategies*, vol. 3, no. 3, pp. 281–299, 2011.
- [68] T. Rymarczyk, E. Kozłowski, G. Kłosowski, and K. Niderla, "Logistic regression for machine learning in process tomography," *Sensors*, vol. 19, no. 15, p. 3400, 2019.
- [69] F. M. Alotaibi, "Classifying text-based emotions using logistic regression," 2019.

- [70] M. Z. Asghar, F. Subhan, M. Imran, F. M. Kundi, S. Shamshirband, A. Mosavi, P. Csiba, and A. R. Várkonyi-Kóczy, "Performance evaluation of supervised machine learning techniques for efficient detection of emotions from online content," *arXiv preprint arXiv:1908.01587*, 2019.
- [71] K. Nugroho, E. Noersasongko, A. Z. Fanani, R. S. Basuki *et al.*, "Improving random forest method to detect hatespeech and offensive word," in *2019 International Conference on Information and Communications Technology (ICOIACT)*. IEEE, 2019, pp. 514–518.
- [72] M. Z. Asghar, F. Subhan, M. Imran, F. M. Kundi, S. Shamshirband, A. Mosavi, P. Csiba, and A. R. Várkonyi-Kóczy, "Performance evaluation of supervised machine learning techniques for efficient detection of emotions from online content," *arXiv preprint arXiv:1908.01587*, 2019.
- [73] F. Livingston, "Implementation of breiman's random forest machine learning algorithm," *ECE591Q Machine Learning Journal Paper*, pp. 1–13, 2005.
- [74] M. Jain, S. Narayan, P. Balaji, A. Bhowmick, R. K. Muthu *et al.*, "Speech emotion recognition using support vector machine," *arXiv preprint arXiv:2002.07590*, 2020.
- [75] T. Pfister, "Emotion detection from speech," *Gonville & Caius College*, p. 60, 2010.
- [76] B. Mahesh, "Machine learning algorithms-a review," *International Journal of Science and Research (IJSR)*. [Internet], vol. 9, pp. 381–386, 2020.
- [77] O. Sagi and L. Rokach, "Ensemble learning: A survey wiley interdisciplinary reviews: Data mining and knowledge discovery," 2018.
- [78] S. Elbagir and J. Yang, "Twitter sentiment analysis using natural language toolkit and vader sentiment," in *Proceedings of the international multiconference of engineers and computer scientists*, vol. 122, 2019, p. 16.
- [79] K. Mishra, I. Kandasamy, V. Kandasamy WB, and F. Smarandache, "A novel framework using neutrosophy for integrated speech and text sentiment analysis," *Symmetry*, vol. 12, no. 10, p. 1715, 2020.
- [80] M. S. Shelke, P. R. Deshmukh, and V. K. Shandilya, "A review on imbalanced data handling using undersampling and oversampling technique," *Int. J. Recent Trends Eng. Res.*, vol. 3, no. 4, pp. 444–449, 2017.