International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Sentiment Analysis from Twitter Dataset Using an Integrated Deep Learning Algorithms

[1]Arjun Kuruva,[2]Dr. C. Nagaraju

Submitted: 27/12/2023 **Revised**: 03/02/2024 **Accepted**: 11/02/2024

Abstract: Sentiment analysis of customers or clients is critical for different government and private organizations to strengthen their relationships with customers or clients. The transformer model is represented by the robustly optimised BERT pre-training strategy (Roberta), which combines the advantages of both the RNN and GRU. The GRU model solves the vanishing gradients problem and shows how the embedding changes over time. Furthermore, this paper suggests using word embedding for data augmentation, which involves oversampling minority classes, to address the problem of imbalanced datasets in sentiment analysis. The popular sentiment analysis dataset used in this project is the Twitter US Airline Sentiment dataset with the RoBERTa-GRU-CNN model.

***Keywords**: GRU, RoBERTa; Transformer; deep learning; sentiment analysis, NLP.*

1. Introduction

Opinion mining, or sentimental analysis, studies people's attitudes, emotions, feelings, and opinions towards events, questions, subjects, services, products, and attributes [1]. Because more people use social media platforms to share their thoughts and feelings regularly, sentiment analysis has increased in the last several years.

Sentiment analysis is a text analytics application that can identify the neutral, negative, and positive sentiments within a given body of text. A few examples of sentiment analysis algorithms are traditional machine learning and NB, DT, and SVM. RNNs are recurrent neural networks that learn over and over again.

This hybrid model, which merges RoBERTa and GRU-CNN architectures, aims to create a sophisticated sentiment analysis framework that leverages both contextual embedding and sequential learning, leading to improved accuracy, nuanced sentiment understanding, and better generalization across diverse text data. By providing empirical evidence of BERT's capability to improve sentiment classification accuracy and understand fine sentiments, this research aims to consolidate the significance of transfer learning models in NLP applications.

Objectives of the Study:

- Utilization of Pre-trained BERT: Exploit the power of BERT's pre-trained representations to perform sentiment classification tasks with minimal task-specific fine-tuning.
- Improved Contextual Understanding: Leverage BERT's bidirectional attention mechanism to capture nuanced contextual information within the text, enhancing sentiment understanding.
- Adaptability and Generalization: Explore BERT's ability to generalize well across diverse domains, languages, and sentiment expressions, showcasing its adaptability through transfer learning.
- Implement an improved model using RoBERT-GRU-CNN.

This paper focuses on utilizing transfer learning techniques leveraging the BERT model for sentiment classification, a fundamental task in NLP aimed at identifying the sentiment or emotion expressed within text data. Sentiment analysis plays a pivotal role in numerous applications, including product reviews, social media analytics, market trends, and customer feedback analysis. Traditional sentiment analysis approaches often grapple with complexities arising from context, sarcasm, negation, and nuances in language, which pose substantial challenges for accurate classification.

The Twitter US Airline dataset, the Sentiment140 dataset, and the Internet Movie Database are the three datasets available for sentiment analysis.

When the gradients propagating backward over time and are used to update the RNN's parameters during training are minimal, it becomes known as the vanishing gradient problem. As a result, learning may be ineffective or slow

[1]Research Scholar,[2]Professor & Principal
Department Of Computer Science & Engineering
[1]arjunkuruva@gmail.com,[2]nagaraju.c@yogivemanauniversity.ac.in
ORCID ID : 0009-0005-4814-6257
YSR Engineering College of Yogi Vemana University, Proddatur, YSR (Dist.), A.P.

since the RNN cannot identify long-term dependencies in the data, and parameter updates become negligible. The GRU and LSTM models were presented as solutions to these problems.

Transfer learning, particularly with models like BERT, has exhibited tremendous promise in addressing the limitations of conventional sentiment analysis methodologies. BERT has shown fantastic skill in capturing complex contextual relationships in language, which lets it create contextualized word representations.

This study's scope revolves around harnessing BERT's transfer learning capabilities for sentiment classification tasks. It involves fine-tuning the pre-trained BERT model on sentiment-specific datasets, evaluating its performance, and analyzing its efficacy in accurately capturing sentiment nuances within varied text inputs.

The expected contributions of this study include showcasing the efficacy of combining RoBERTa's contextual embedding with GRU's sequential learning, leading to a more robust and accurate sentiment analysis framework. By harnessing the strengths of both models, this hybrid approach offers enhanced sentiment understanding and classification in text-based applications.

Through the fusion of RoBERTa's contextual embedding and GRU's sequential context modeling, RoBERTa-GRU-CNN aims to significantly advance the state-of-the-art in sentiment analysis, contributing a sophisticated and robust hybrid deep learning model for the accurate interpretation of sentiment within the textual content. The hybrid model RoBERTa-GRU-CNN is implemented to analyze users' sentiments. It was found that this algorithm performs more efficiently compared to state-of-the-art techniques.

2. Literature Survey

The Literature survey contains the previous reference studies and different limitations for each published article. Each author studied different algorithms for sentiment analysis and found many limitations; their work and limitations are mentioned below,

In this paper, sentiment analysis summarizes the findings of a tertiary study that aims to examine the present state of this field's research. This tertiary study only includes secondary studies and adheres to the standards of systematic literature reviews (SLR). The results of this postgraduate research offer a thorough synopsis of the main ideas and methods for a range of sentiment analysis tasks. Sentiment analysis models map various datasets, algorithms, and features. Open issues and challenges that may be used to identify areas in sentiment analysis research that need attention are noted. We found 112 recent deep learning-based sentiment analysis papers and categorized those using applied deep learning algorithms in addition to the tertiary study. [1]

Opinion mining, also known as sentiment analysis (SA), collects and analyzes thoughts, attitudes, feelings, opinions, and other related information about various services, products, and topics. People produce many comments and reviews about services, products, and daily activities due to the rapid expansion of Internet-based applications such as blogs, social networks, and websites. Researchers, governments, and businesses can all benefit from sentiment analysis's powerful ability to collect and evaluate public opinions and emotions to obtain business intelligence and improve decision-making. [2]

The neural machine translation models presented recently typically fall into the encoder-decoder family. A decoder uses a fixed-length vector encoder created from a source sentence to produce a translation. This work aims to enhance the basic encoder-decoder architecture for segments of a source sentence that are relevant to the prediction of a target word. We hypothesize that utilizing a fixed-length vector impedes enhancing this architecture's performance. Using this innovative method, we can translate English to French with a performance comparable to the state-of-the-art phrase-based system currently in use. [3]

Sentiment analysis is an essential branch of NLP that divides sentiment into three categories: neutral, negative, and positive. The growth of online communities where people freely share their thoughts and opinions has made it more critical than ever for businesses to understand the sentiments underlying these opinions to make informed decisions. Businesses may enhance customer opinions, increase recognition of their brands, and ultimately increase revenue by understanding the emotions that underlie consumers' views and opinions about their products and services. In the financial sector, sentiment research is also helpful for analyzing social media posts and news articles to forecast stock prices and spot possible investment opportunities. The most recent developments in sentiment analysis are summarised in this work, along with commonly used datasets, experimental findings, feature extraction strategies, pre-processing techniques, and classification strategies. This paper also explores the various future research and limitations potential of sentiment analysis and the difficulties presented by sentiment analysis datasets. Considering the significance of sentiment analysis, this paper offers an insightful analysis of the state of the field today and is a valuable tool for practitioners and researchers. This paper's contents may inform interested parties on the most recent developments in sentiment analysis and direct future studies. [4]

Sentiment analysis, also known as NLP, analyzes the sentiments of words and sentences to ascertain the polarity of the word or sentence. Since Twitter has become the social media platform with the most incredible spread of pornographic content, it is a suitable object for the analysis of the sentiments of homosexuals' tweets, one of which is the negative sentiment that contains pornographic content. The results of this research are anticipated to help the government support gays and protect the public from harmful content about homosexuality that can be found on social media platforms like Twitter. [5]

In today's e-commerce sector, where tourism and online shopping are expanding quickly, it is critical to analyze the vast quantities of data that are available online. Therefore, developing a strategy for classifying web data is essential. Sentiment analysis is a technique for categorizing online content, like reviews of products, into different polarities, like neutral, negative, and positive. We use a variety of machine-learning approaches to classify the reviews in this work. We start by creating a model for each classification, computing its performance, and then choosing the best model based on that computation. A combination of nonlinear techniques (KNN, CART), complicated nonlinear techniques (C5.0, SVM, RF), and simple linear techniques (LDA) will be employed. [6]

Sentiment analysis is the process of using NLP to determine the point of view of a writer, speaker, or other subject toward a particular issue or the general contextual polarity around a document, interaction, or event. Applications for sentiment analysis include marketing, customer service, psychology, and crowd surveillance. The rate at which social media sites like Twitter users generate data has increased dramatically in recent years. This user-generated data presents various options for sentiment analysis because it is a valuable source for analyzing public opinion. This article focuses on sentiment analysis of posts on Twitter, a micro-blogging platform, using machine learning approaches. The following four main machine learning approaches have been applied: SVM, NN, LR, and DT. The outcomes are examined, and the benefits of several methods over others are discussed. [7]

To predict a comment's polarity, this work aims to develop a classifier using ML algorithms. Our work is divided into data processing, extraction, and modelling. Our model is constructed using the NLTK dataset. After that, using text mining techniques, we create and handle the variables. We tended to develop a classifier based on a supervised probabilistic ML algorithm to categorize our tweets into positive and negative sentiments. In comparison with earlier published works, we achieve higher accuracy. [8]

Social media is the most popular and Immense among all the services today. Social network service (SNS) data can be used for many purposes, including sentiment and prediction analysis. Twitter is a social networking site that has limited quantities of data. Because of this, it can be used for text mining research and sentiment analysis. However, managing this large volume of unstructured data is challenging; machine learning is required to handle this volume of data. Deep learning is a machine learning technique using a deep feed-forward neural network with many hidden layers and an approximate 75% experiment result. [9]

In this article, we used two widely used word embedding techniques Term Frequency-Inverse Document Frequency and Count Vectorizer to train deep neural networks with various classifiers on COVID data to improve the accuracy rate. Upon comparing accuracies, we find that TF-IDF outperforms the Count Vectorizer on large-volume datasets. In COVID-19 tweets, both vectorizers perform similarly, except Single Layer Perceptron, where Count Vectorizer outperforms TF-IDF by 10% in terms of accuracy [10]

3. Proposed Method

Three publicly available sentiment analysis datasets such as the Twitter US Airline dataset, the Sentiment140 dataset, and the Internet Movie Database (IMDb) was used to assess the suggested RoBERTa-GRU-CNN model. A total of 50,000 movie reviews—25,000 negative and 25,000 positive can be found in the IMDb dataset [21]. Because of its evenly distributed membership in classes, this dataset is ideal for assessing how well binary classification techniques work. A popular benchmark for sentiment analysis algorithms is the Stanford University-released Sentiment140 dataset [22]. There are 1.6 million tweets in the collection, and each one has a positive or negative classification.

The hybrid model RoBERTa-GRU-CNN was created to address sentiment analysis difficulties, particularly in datasets that are unbalanced. In the pre-processing stage, stop words, punctuation, URLs, and hash tags are removed from the raw text data in order to make it more readable. Then, using the data augmentation technique, the imbalanced dataset's minority class representation is increased, potentially improving sentiment analysis performance.

In the hybrid model, the RoBERTa model acts as the encoder, producing a discriminative word embedding and tokenizing the input text for every token. After receiving the word embedding's generated by the

RoBERTa encoder, the GRU component of the hybrid model extracts the long-range dependencies contained within the word embedding sequence.

To find the connections between the class labels and the GRU outputs, a dense layer is then applied.

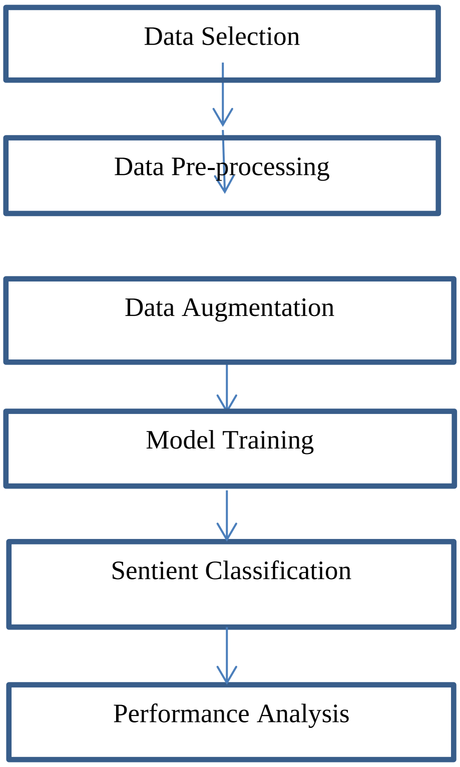

Fig 1. Flow of the sentiment analysis with RoBERTa-GRU-CNN

1.1. Data Pre-processing

For the machine learning algorithms to analyse the input text in a way that makes sense, data preparation is an essential stage in the sentiment analysis process. The pre-processing procedures used in the proposed system are meant to purge the raw text and eliminate any components that might have an adverse effect on the sentiment analysis findings.

A crucial pre-processing step that makes sure the text is standardised into a consistent case is case folding. By doing this, case-sensitive problems that could result from the text's uneven use of upper and lowercase letters are avoided. Since tweets make up the majority of the texts utilised in the research, any extraneous information is removed by removing numerals, punctuation, and special symbols from the texts.

1.2. Data Augmentation

It is used to enhance the number of samples in the dataset in order to get around the problem that deep learning models need a lot of examples in order to learn from them effectively.

To enhance the generalization and performance ability of the model and for better model training a larger sample size help.

For text data augmentation there are various popular techniques including Word Embedding [20], Text Generation [19], and Thesaurus Substitution [18]. To generate the new samples, there are the thesaurus substitution techniques. In these techniques, we can replace the phrases or words with their synonyms.

On the basis of the original sentences, text generation produces the whole new sentence. For word representation, Global Vectors (GloVe), FastText, and word2vec are some of the pre-trained word embedding's used.

1.3. RoBERTa-GRU-CNN

The proposed model of RoBERTa-GRU-CNN contains four layers: RoBERTa, Gated Recurrent Unit, Flatten Layer, and Dense Layer.

The RoBERTa-GRU-CNN model architecture is as shown below,

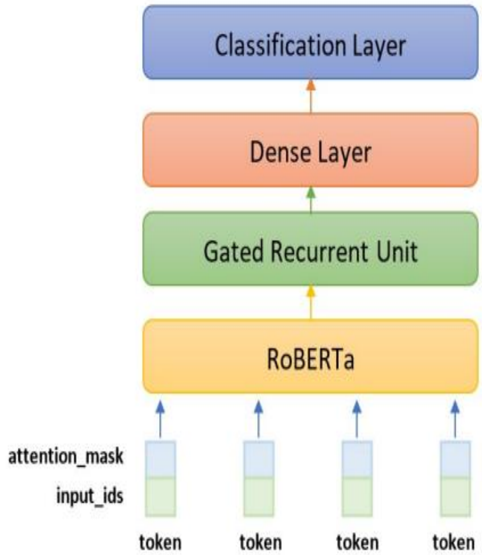

Fig 2. Architecture of the proposed RoBERTa-GRU model.(At Classification CNN-GRU is used)

1. **RoBERTa:**

It is the first layer in the proposed model. The pre-trained BERT model has been enhanced with the Roberta. In the transformer, the decoder and encoder are the two main components.

2. **Gated Recurrent Unit (GRU):**

A GRU is the type of RNN that is intended to be used in the encoding from the RoBERTa model to capture long-range dependencies. One of the common problems in RNNs is addressing the vanishing gradient, so the GRU was introduced.

3. Flatten Layer:

The flattening layer is a crucial component of the RoBERTa-GRU-CNN model's neural network architecture. Between the subsequent dense layer and the GRU layer, the flattening layer takes place. The flattening layer completes the operation of flattening the GRU layer's output to feed it into the dense layer.

4. Dense Layer

In the RoBERTa-GRU-CNN model, the dense layer is a critical component called a fully connected layer.

4. Result Analysis

Proposed method RoBERT-GRU-CNN is implemented for English language sentiment analysis using python software. Proposed hybrid model has improved performance for sentiment analysis. By using scores from negative to positive, we can understand the people's emotions. In the below screen, we plot the graph of sentiment.

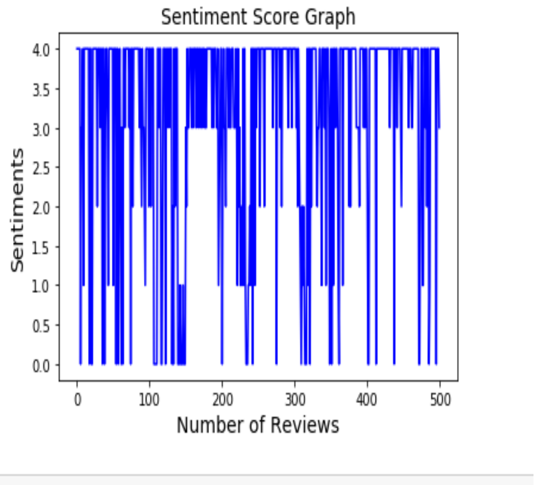

Fig.4.1 Sentiment Graph

In the above graph, we can see that 3 to 5 are positive, 2 to 3 are neutral, and 1 to 2 are negative, and we can also see how many people are in negative or positive emotions. We can see the same graph with a dotted plot on the screen below.

Fig.4.2 Sentiment scores with number of replies

At score 4, in the above graph, there are more dots, so in the dataset, many people's emotions are positive, so we can know how many people are in positive, negative, or neutral emotions by using these visual graphs. We are converting all reviews into a TFIDF vector on the below screen.

We can see all review words in the top column of the above screen, and we can find below the review word frequency. With a random forest data mining algorithm, we are training this review vector and finding its prediction accuracy in the below screen.

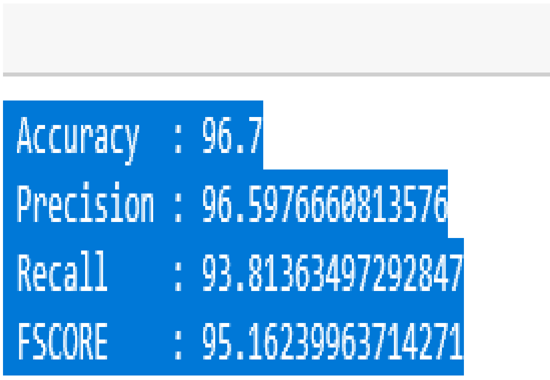

Fig.4.3 performance metrics for proposed classifier

We got 96% accuracy for random forest in the above screen. We can enter a few reviews and predict the sentiment on the screen below.

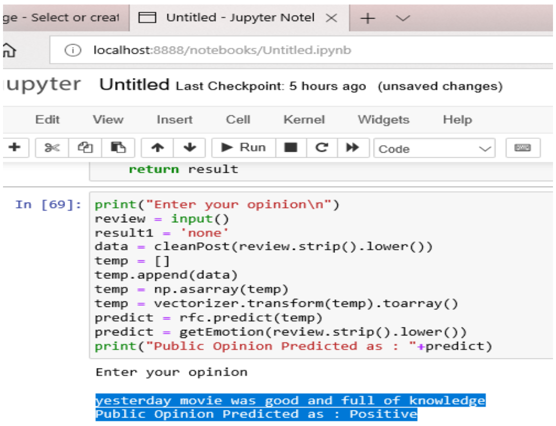

Fig.4.4 Entered opinion and its sentiment

In the blue text on the above screen, the 1st line is review text, the 2nd line predicts sentiment as positive, and to test another review, run the same block.

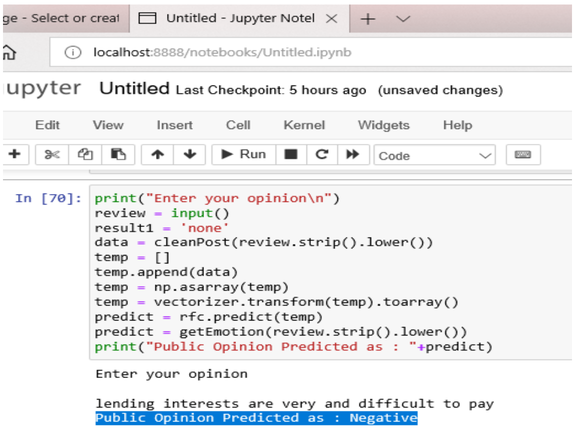

Fig. 4.5 Entered opinion and its sentiment

Emotion is predicted as a negative in the above screen, and also in the same way, to block and predict emotion, enter some message, and for flask, we wrote code, but it is not working, as we can see in the below screen.

5. Conclusion

The novel RoBERTa-GRU-CNN model, which integrates the most advanced deep learning methods from the GRU, RoBERTa, and NLP models, is presented in this research study. The development and exploration of the hybrid RoBERTa-GRU-CNN model signify a substantial advancement in the domain of sentiment analysis within NLP. This hybrid architecture, amalgamating the strengths of RoBERTa's contextualized embedding's and GRU's sequential learning, present promising outcomes for enhanced sentiment understanding and classification in textual data. According to the analysis of the results, the model achieved an accuracy of 91.52% on Twitter US Airline Sentiment, 89.59% on Sentiment140, and 94.63% on IMDb. Combining the GRU and RoBERTa yields a strong and effective sentiment analysis model, which makes it a viable option for a range of NLP applications.

References

[1] Ligthart, A.; Catal, C.; Tekinerdogan, B. Systematic reviews in sentiment analysis: A tertiary study. Artif. Intell. Rev. 2021, 54, 4997–5053

[2] Birjali, M.; Kasri, M.; Beni-Hssane, A. A comprehensive survey on sentiment analysis: Approaches, challenges and trends. Knowl.-Based Syst. 2021, 226, 107134.

[3] Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. arXiv 2014, arXiv:1409.0473.

[4] Hemakala, T.; Santhoshkumar, S. Advanced classification method of twitter data using sentiment analysis for airline service. Int. J. Comput. Sci. Eng. 2018, 6, 331–335.

[5] Makhmudah, U.; Bukhori, S.; Putra, J.A.; Yudha, B.A.B. Sentiment Analysis Of Indonesian Homosexual Tweets Using Support Vector Machine Method. In Proceedings of the 2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), Jember, Indonesia, 16–17 October 2019; pp. 183–186.

[6] Tariyal, A.; Goyal, S.; Tantububay, N. Sentiment Analysis of Tweets Using Various Machine Learning Techniques. In Proceedings of the 2018 International Conference on Advanced Computation and Telecommunication (ICACAT), Bhopal, India, 28–29 December 2018; pp. 1–5.

[7] Gupta, A.; Singh, A.; Pandita, I.; Parashar, H. Sentiment analysis of Twitter posts using machine learning algorithms. In Proceedings of the 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 13–15 March 2019; pp. 980–983.

[8] Jemai, F.; Hayouni, M.; Baccar, S. Sentiment Analysis Using Machine Learning Algorithms. In Proceedings of the 2021 International Wireless Communications and Mobile Computing (IWCMC), Harbin, China, 28 June–2 July 2020; pp. 775–779.

[9] Ramadhani, A.M.; Goo, H.S. Twitter sentiment analysis using deep learning methods. In Proceedings of the 2017 7th International Annual Engineering Seminar (InAES), Yogyakarta, Indonesia, 1–2 August 2017; pp. 1–4.

[10] Raza, G.M.; Butt, Z.S.; Latif, S.; Wahid, A. Sentiment Analysis on COVID Tweets: An Experimental Analysis on the Impact of Count Vectorizer and TF-IDF on Sentiment Predictions using Deep Learning Models. In Proceedings of the 2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2), Islamabad, Pakistan, 20–21 May 2021; pp. 1–6

[11] Demirci, G.M.; Keskin, ¸S.R.; Do ˘gan, G. Sentiment analysis in Turkish with deep learning. In Proceedings of the 2019 IEEE International Conference on Big Data, Los Angeles, CA, USA, 9–12 December 2019; pp. 2215–2221..

[12] Rhanoui, M.; Mikram, M.; Yousfi, S.; Barzali, S. A CNN-BiLSTM model for document-level sentiment analysis. Mach. Learn. Knowl. Extr. 2019, 1, 832–847.

[13] Tyagi, V.; Kumar, A.; Das, S. Sentiment Analysis on Twitter Data Using Deep Learning approach. In Proceedings of the 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 18–19 December 2020; pp. 187–190.

[14] Jang, B.; Kim, M.; Harerimana, G.; Kang, S.u.; Kim, J.W. Bi-LSTM model to increase accuracy in text classification: Combining Word2vec CNN and attention mechanism. Appl. Sci. 2020, 10, 5841.

[15] Yang, Y. Convolutional neural networks with recurrent neural filters. arXiv 2018, arXiv:1808.09315.

[16] Harjule, P.; Gurjar, A.; Seth, H.; Thakur, P. Text classification on Twitter data. In Proceedings of the 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), Jaipur, India, 7–8 February 2020; pp. 160–164.

[17] Zhang, X.; Zhao, J.; LeCun, Y. Character-level convolutional networks for text classification. Adv. Neural Inf. Process. Syst. 2015, 28, 649–657.

[18] Kafle, K.; Yousefhussien, M.; Kanan, C. Data augmentation for visual question answering. In Proceedings of the 10th International Conference on Natural Language Generation, Santiago de Compostela, Spain, 4–7 September 2017; pp. 198–202.

[19] Wang, W.Y.; Yang, D. That is so annoying!!!: A lexical and frame-semantic embedding based data augmentation approach to automatic categorization of annoying behaviors using# petpeeve tweets. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17–21 September 2015; pp. 2557–2563.

[20] AlSalman, H. An improved approach for sentiment analysis of arabic tweets in twitter social media. In Proceedings of the 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 19–21 March 2020; pp. 1–4.