Open Access
Full Length Article

ORIGINAL ARTICLE

Text-Based Sentiment Analysis Using CNN-GRU Deep Learning Model

Ahsan Nadeem[1], Naeem Aslam[2], Muhammad Kamran Abid[3], Muhammad Fuzail[4]

[1,2,3,4] Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan

ABSTRACT

Sentiment analysis identifies both positive and negative viewpoints from sources like social media, surveys, and reviews by automating text analysis with artificial intelligence (AI). Using data to inform decisions is made easier by this. Deep Learning (DL) has gained a lot of interest in recent years from academia and industry because of its outstanding performance. Convolutional neural networks (CNN) and recurrent neural networks (RNN) are the two deep learning designs that are most frequently utilized. Because they can examine enormous amounts of data, neural networks have the potential to be more accurate in sentiment analysis. Our work utilizes a hybrid model that combines Word2Vec preprocessing with the Gated Recurrent Unit (GRU) from Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) for sentiment analysis on movie and smartphone reviews. Achieving 99.99% precision, 99.84% recall, and 99.92% accuracy for movie reviews, and 99.08% accuracy, 98.93% precision, 99.40% recall, and 98.93% F1-score for Amazon mobile phone assessments. This paper presents a CNN-GRU-based Word2Vec algorithm for sentiment categorization, which addresses the challenge of evaluating vast amounts of user-generated text data with 99.50% accuracy.

Keywords: *Sentiment Analysis, Machine Learning, Deep Learning, Sentiment Score, CNN, GRU*

Author`s Contribution

[1,2,3,4] *Data analysis, interpretation, and manuscript writing, Active participation in data collection, Conception, synthesis, and planning of research. Interpretation and discussion*

Address of Correspondence
Ahsan Nadeem
Email: ahsan.nadeem@gmail.com

Article info.
Received: August 05, 2023
Accepted: December 11, 2023
Published: December 30, 2023

Cite this article: *Nadeem A, Aslam N, Abid M, Fuzail M. Text-Based Sentiment Analysis Using CNN-GRU Deep Learning Model. J. inf. commun. technol. robot. appl.2023; 14(1):8-15*

Funding Source: *Nil*
Conflict of Interest: *Nil*

INTRODUCTION

Social media platforms have become essential components of contemporary life, giving people a forum to interact with others on a variety of subjects, share experiences, and voice their opinions. A significant and quickly growing source of data is the enormous volume of content shared on these platforms, including text messages [1], photos, and videos. People establish online communities where they share opinions, feelings, and responses to a range of events and subjects through electronic communication. Because of this trend, social media has emerged as a major global source of a wide range of information that shapes discussion and public opinion. Over the past ten years, the use of social media has surged at an unparalleled rate; billions of people use Facebook, Instagram, Twitter, and WhatsApp alone. The widespread use of social media has changed how people interact, communicate, and use information, which has had a dramatic impact on people's lives. Social media is becoming a necessary tool for marketing, knowledge sharing, and self-expression [2].

Businesses utilize social media channels to communicate with clients, market their goods and services, and get feedback. Businesses can obtain important insights into consumer preferences, market trends, and brand impressions by keeping an eye on social media interactions. Social media allows businesses to interact with customers in real-time, which expedites problem-solving and fortifies ties with the audience. On social media, politicians, celebrities, and organizations can interact with their followers, disseminate information, and ascertain public opinion [2][3]. Social media platforms such as Twitter and others are valuable resources for political debate because they give politicians a direct line of communication with the public and help them discover what views the public has on a range of subjects [4]. In the same manner, celebrities interact with their followers on social media, provide updates about themselves, and advertise their endeavors.

Analyzing public emotions, spotting new trends, and projecting future consumer behavior all depend on social media analysis. Through a detailed analysis of data created on social media platforms, analysts and researchers may find important insights into people's beliefs, proclivities, and attitudes. Decision-making processes in a range of industries, such as marketing, public relations, and policy development, can benefit from this information. Facebook acts as a platform, which is essential for social networking, and has been leveraged as a key tool to influence the way people relate and communicate on the internet. Facebook was the first social network that offered an immediate and exclusive online presence that was based on instantaneous messaging, text publication, and sharing of multimedia content. It is one of the world's most-used and talk platforms with over 2 billion active users. Friends and family of experienced users could live anywhere yet together authorize.

Facebook as a marketing equipment that organizations use to win the hearts of new customers, boost brand visibility, and grow sales is notoriously known. Companies stand a good chance of developing and managing promising marketing efforts in the long run through techniques such as targeted marketing and involvement. Facebook also furnishes with businesses beneficial insights about the demographics, interests, and 'behaviours of customers which can be applied to the analysis of marketing strategy. Social media Facebook creates a platform for the educational and corporate community to teach each other, and enlighten others through sharing information among themselves. The creation of groups and pages as per individual school subjects and interests will create an opportunity for educators and learners to act together on projects and have in-depth discussions. They may also exchange materials. Does this act of sharing information on Facebook contribute to the planning of various events, instigate, and propagate activism, and diffuse social issue awareness as well [7]?
Apart from that, Twitter is also a popular social media site where real-time conversations, transactions, and breaking news come up. Twitter is the most chosen medium among millions of users unanimously since Twitter has evolved to be a platform for real-time, fresh information about news, pop culture, and public conversations. People contribute their opinions on a range of issues, like to participate in the discussion, and follow news providers. The key role of Twitter is to communicate businesses with their customer, brand building, and online reputation management.

By searching for consumer questions and related hashtags, businesses can provide feedback and solutions, and promote their products and services in Twitter trends and conversations. Companies through Twitter can link with audiences of their targeted choices by choosing among the various advertising options [7].

Consequently, social interactions and the spread of information in social media have been affected and are different from before. These sites, which have billions of users worldwide, are rapidly transforming into main gateways for information as well as insight hunters, by people, companies, and organizations. Academics and experts can have a lot to gain by looking at how social media has given us a closer insight into people's views, interests, and conduct.

These results could then be applied to guide decisions in many contexts. Sentiment analysis [8], or just SA for short, is an artificial natural language processing (NLP) method that classifies textual input into three categories: neutral, positive, and negative. SA helps organizations understand what customers think about their products and services by using machine learning algorithms to analyze large quantities of text. Unstructured data, which mainly originates from text-based sources including articles, emails, social media, and chats, makes up about 80% of all data. Automating text processing [9] with sentiment analysis saves time and yields useful information. It helps with web presence management, customer support, and marketing plan assistance because it is based on sentiment research.

In many domains that have had recent breakthroughs, such as computer vision and natural language processing (NLP), deep learning models have replaced antiquated techniques [10]. These models make substantial progress by utilizing labeled data and the enhanced capabilities of deep learning. Deep learning techniques take an end-to-end modeling approach, in contrast to conventional methods that manually extract features before classifying them. This method eliminates the requirement for manual feature engineering by allowing classification activities to be completed while training classification models simultaneously. In tasks including sentiment analysis, information retrieval, machine translation, word embed-ding, named entity recognition in NLP, image classification, object recognition, segmentation, and image production, deep learning-based methods have shown impressive results. Because of their superior performance in both image classification and natural language processing tasks, convolutional neural networks, or CNNs, have become a popular option for sentiment analysis on text data. CNNs can extract significant features by convolving across sections of the input data, which is an effective way of capturing spatial correlations between features, in contrast to Recurrent Neural Networks (RNNs), which need label-ing of whole sequences.

CNNs are excellent at identifying global patterns in data and extracting local characteristics, which makes them a good choice for sentiment analysis applications. To maximize model performance, researchers experiment with different CNN designs by changing factors like the number of convolutional layers, filter counts, and filter sizes. The efficacy of the suggested CNN model is further assessed through comparison with alternative machine learning techniques. The remarkable capabilities of deep learning, in particular CNNs and RNNs, have attracted a lot of interest from academia and industry. Because these neural network architectures can handle enormous amounts of data and achieve higher accuracy than previous methods, they show promise in sentiment categorization applications. Deep learning is anticipated to become increasingly important as it develops, propelling discoveries and breakthroughs in sentiment analysis and other areas.

Social media sites like Facebook, Instagram, YouTube, and many more have become incredibly popular in the last ten years. Companies have come to grasp that these platforms can be a priceless supply of data that improves their ability to interact with and comprehend their clientele. The abundance of emails, posts, comments, and other types of interaction, however, makes it very challenging to keep track of how satisfied a consumer is with a business. One of the most widely used application sites nowadays, Twitter, for example, boasts over 350 million active users who send out 500 million tweets daily. Nevertheless, an automated method to assist in processing and extracting useful data from these massive volumes of data may be provided by Deep Learning and other Machine Learning (ML) techniques. This paper presents a deep learning-based sentiment analysis method and combines it with Word2Vec embedding to determine and examine the polarity of the text-based content.

LITERATURE REVIEW

The study by Rodrigues et al.'s [11] objective was to detect live Twitter spam messages and SA of saved and live tweets. The proposed method made use of two different databases, one for spam identification and one for SAs. The LSTM-DL model achieved a classification performance of 96.78%, while the multivariate regression naive Bayes classifier had a validation accuracy of 98.74 percent for the classification of Twitter spam. The categorization process demonstrated that information taken from tweets may be used to reliably determine whether a tweet on Twitter is spam. Using the features that were taken from the tweets, it can accurately determine the sentiment value of those tweets, based on the classification results. The long short-term memory classifier achieved an accuracy of 73.81 percent for the Twitter sentiment classification, whereas the SVM classifier had an accuracy of 70.50 percent.

Using the categorization of tweets collected from the social networking site Twitter, Sitaula et al. [12] investigated the emotions of Nepali people. At first, they advise doing this using three different feature extraction methods. Three substitute CNNs are proposed to implement the recommended features. Lastly, use an ensemble network that functions end-to-end to put these three convolutional models together to arrive at the results. These feature extraction strategies have the necessary discriminating power for sentiment categorization, as the research results on this dataset show. Compared to the worst method (59.5%), this methodology improves classifier accuracy by a large margin of 68.7%, or 5.8% more than the leading method (62.9%). There has been a significant improvement in classification accuracy with the same size range and a 95% F-1 score.

Using an efficient Bidirectional LSTM Network (BiLSTM) and Forward-Backward encapsulating context data from Arabic features sequencing, Hanane et al. [13] suggested a model. The suggested framework performs better than the most recent DL [14] models as well as the industry-standard conventional ML techniques, according to the results obtained from six benchmark SA databases. On this dataset, the system achieves multiple accuracies of 0.7205, 0.918, and 0.926 while surpassing state-of-the-art approaches. In a study by W. Hazim, G. Guad, I. Mahmood, et al. [15], the Arabic Tweet dataset from the University of California, Irvine (UCI) Machine Learning Repository database was utilized to evaluate the LSTM deep learning method. With the help of the LSTM, Arabic tweets' positive and negative sentiments were distinguished. With an accuracy score of 89.8%, the analysis revealed that the LSTM model outperformed the previous study using the same dataset and traditional machine learning algorithms.

The deep learning model for Arabic Twitter sentiment categorization proposed by A. Ombabi, W. Ourada, and A. Alims [16] consists of two layers of the LSTM algorithm and one layer of the CNN method. Moreover, it was considered that the model's input layer was the Fast Text (Skip-gram) word embedding, which converted words into vectors. The SVM approach was applied to the model to improve the classification results. The experiment demonstrated that the suggested model produced a very good outcome with an accuracy of 90.75%. L. A. Alhuri, H. R. Aljohani, et al. [17] demonstrated their deep learning algorithms, LSTM and GRU, for analyzing sentiment in Arabic tweets. They leveraged an already-existing collection of COVID-19 tweets collected using Saudi hashtags. The authors annotated the polarity of tweets using the already accessible 2.2 million tweet corpus known as AraSenTi. The word embedding technique Global Vectors (GloVe) was employed in this work to extract features. The F1 score of 81% in the results indicated that the GRU classifier outperformed the LSTM.

Basiri et al. [18] used a bi-directional CNN–RNN model based on the attention mechanism to get better assessment results. This model has a separate BiGRU layer for context extraction and an independent BiLSTM layer. Cheng et al. [19] introduced an improved method for categorizing Chinese short texts using the Ernie_BiGRU model. The text categorization task performed well when Ernie and BiGRU were used together, according to many assessment criteria. Li Yang and colleagues [20] used the features obtained from CNN and BiLSTM to leverage the advantages of each model's feature extraction and greatly improve the model's classification performance. Important terms should be given more consideration because the informational value of the text varies from the results of the classification. In support of supervised machine learning methods like Naive Bayes and Linear Support Vector Machine, Rachana Bandana [21] presented sentiment analysis at the document level for the classification of movie reviews using heterogeneous information. With the help of, the author gained knowledge and created a list of both favorable and bad text evaluations.

Using sentiment analysis of movie tweets, M. Ali Fauzi [22] demonstrates the ensemble technique using Naive Bayes. Together with textual, lexicon-based, part-speech, and Twitter-specific features, the ensemble features also comprise Bag of Words. To save the fraction of positive and negative tweet terms according to their part of speech, they developed lexicon-based characteristics. The sentiment analysis of a social network for a military life board on the biggest online communities in Taiwan was offered by Liang-Chu Chen et al [23]. For this online community, text mining was used methodically with the use of a self-organized military emotion lexicon. For this military life PPT board, a web crawler technology is utilized to extract the material from military blogs and postings. When it comes to customer reviews in mobile applications, Barkha Bansal et al.'s [24] work compares the CBOW model to the skip-gram model. Semantic features for every mobile phone input are calculated using the cosine distance measure.

Hou et al. [25] published an attention mechanism recognition model based on BiGRU that was used to detect ship problems. The bidirectional GRU network outperformed earlier models in terms of speed, accuracy, and recall rates, according to experiments. Marquez et al. [26] suggested a time-varying sentiment lexicography based on incremental word vectors. By training an incremental word sentiment classifier with a dynamic word vector, the aim is to automatically update the sentiment lexicography. K. Kobs, A. Zehe, A. Bernstetter, et al. [27] suggested a hierarchical fusion cross-modal complementary network for multi-modal sentiment network analysis. The feature extraction module from text and picture was used to learn the attention features of text and image generated by the image-text generator, to create a hierarchical fusion framework that could fully merge diverse modal elements and accurately interpret the emotion of text and image. P. Mishra [28] proposed a three-layer convolution architecture for sentiment analysis that makes use of a bidirectional encoder and CNN to examine the relevant corpus that was collected from Twitter. For its study, it used a corpus of 17,000 tweets. The results reveal its high degree of accuracy.

RESEARCH METHODOLOGY

In addition to the standard operations of data preprocessing, each data set will now undergo the operations embedded in the BOW. This simulates the process of the R programming environment in creating sequences of documents in the Term Document Matrix which is commonly referred to as the TDM. This technique enables thorough digging into the patterns of word repetition and the connections between sentiment with term counts by logging the glimpse of term occurrence in the dataset. Through applying the second analytics step, the top 20 phrases that are highly repetitive among all the reviews of both datasets will be revealed. Even though the TDM matrix function in the R programming environment allows for the generation of word frequency distribution in the tabular and graphical form, it does not consider each individual word's psycho-dynamic meaning. Data based on the sentences that are used most often in the review are presented at this step, thus showing the things that are regarded to be popular and the feelings that customers hold.

The sentiment rating and polarity (positive or negative) of each review are first extracted utilizing the Syuzhet and Tidytext algorithms in the R environment. This translates to seeking to establish whether the content suggests an emotion and then sorting it depending on its emotional tone. They make a thorough analysis of the stance and points of view as inferred from these assessments possible by sentiment quantification. Next, the sentiment analysis performs the examination process into words and phrases present in the text. A Word cloud tool is utilized to display the results of the analysis visually. The TDM matrix is used to scale the word cloud of the resulting word cloud according to the frequency of each word by transforming it into color, size, or font style. The pictograph technique represents the most often perceived emotions and themes clustered within the evaluations in a more enjoyable and tractable way.

In the last stage of the process, a CNN-GRU model based on deep learning is used for categorizing the textual data sets by utilizing Word2Vec and camper bone embeddings in the Python computing environment. To accomplish better text sentiment decoding as well as their capability to capture complex inter-correlations and patterns, this model heavily relies on Gated Recurrent Units (GRUs) and Convolutional Neural Networks (CNNs). When in application of word2vec meanings the model makes meaning an increased success in sentiment analysis tasks it brings the context knowledge and semantics representation.

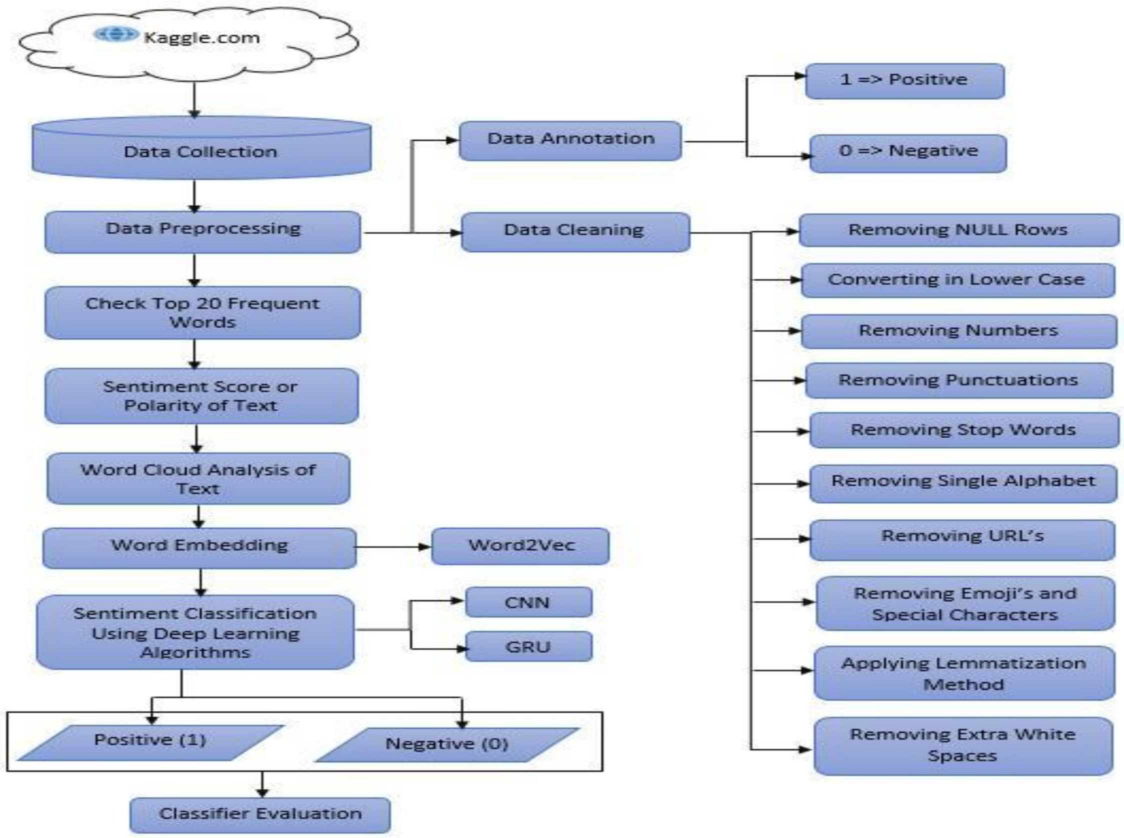

Figure 1: Proposed Methodology of Sentiment Analysis

A dataset is organized data in rows and columns with the purpose of statistical analysis, modeling, and strategic planning that are common in most of the education fields. The reviews in the IMDB-Re-views dataset, which consist of 50,000 reviews (accessible on platforms such as Kaggle), contain positive or negative opinions that have been determined by the grammatical mood of the reviews. On the contrary, unlike the categorization in the social media dataset, the user's sentiment is reflected in the 50,000 user reviews of mobile phones which make up the Amazon Mobile Phone Dataset (also from Kaggle), and this is easy to do because all the reviews are labeled by sentiment tone. An established trend is that these datasets serve as essential assets for the researchers and the practitioners engaged in sentiment analysis because they are collections of text data with annotations meant for the models' formation and testing. The secondary data consists of the word-count and word-tokenization columns used to count words in each document, and the Document_IDs that are used to uniquely name each document found in both datasets. The evaluations column contains user-provided feedback like comments or ratings.

The Sentiment column uses the sentiment represented in the text to generate a binary label (1 for positive, 0 for negative). These features make it easier to distinguish between independent and dependent components, which facilitates the analysis and simulation of the sentiment patterns and trends seen in the datasets.

Preprocessing is done on the provided CSV data to extract pertinent data. The process of reviewing and cleaning text data involves several processes aimed at improving its quality and usefulness. To verify the integrity of the data, null rows are first eliminated. We then transform text reviews to lowercase so that the corpus environment is more uniform. Then, to concentrate on important text, punctuation, stop words, single alphabets, URLs, emoticons, and special characters are eliminated. By reducing words to their base forms, stemming and lemmatization processes are used to further enhance the text. The goal of these preparation procedures is to provide a consistent and uncluttered corpus environment for analysis. The final dataset is more suited for sentiment analysis and other natural language processing applications [33][34] by removing unnecessary information and normalizing the text. Eliminating superfluous and insignificant textual parts improves the precision and efficiency of ensuing analysis and modeling procedures. Furthermore, removing superfluous white spaces from the dataset guarantees uniformity and readability. Preprocessing, in general, is an important step in getting text data ready for sentiment analysis. This helps practitioners and academics get valuable insights from the processed data and make wise judgments.

We found the top twenty words, which are essential for text analysis, in the text reviews of each dataset. Word frequency analysis helps to analyze customer service problems and product reviews by identifying common terms and trends in datasets. As shown in Figure 2, we identified the top 20 frequently occurring words for each dataset review using TDM in RStudio. We conducted a sentiment analysis on two datasets, IMDB Reviews, and Amazon Mobile Phone Reviews, using a series of procedures in the R programming environment. We used NLP [35] techniques to sanitize the text after creating a text review corpus using the supplied datasets [36]. We computed sentiment scores using the Syuzhet package [37], which is shown in Figures 3 and 4 using ggplot2. Similarly, sentiment scores were acquired using the Tidytext package [38], which is shown in Figure 5. Sentiment analysis using word cloud analysis highlights frequently used terms and the associated sentiment, providing a visual summary of the sentiments represented in the text. This visual aid provides a succinct synopsis of the prevailing viewpoints and sentiments. Figures 6 and 7 display sentiment analysis of reviews for the Amazon Mobile Phone and IMDB using the word cloud package in R [39][40]. Finally, sentiment analysis was done in Python utilizing Kaggle's Jupyter environment on IMDB and Amazon Mobile Phone reviews. This included importing, cleaning, preprocessing, and splitting the datasets, tokenizing [41], padding the text [42], using Word2Vec [43], and analyzing the outcome with CNN-GRU [44][45].

word	frequency
movie	98989
film	92087
good	70862
one	53318
see	47083
make	43988
like	43102
get	35199
just	34883
time	29762
character	27602
watch	27265
bad	25759
even	25002
story	24158
think	23992
really	22951
show	21397
scene	20707
great	19762

word	frequency
phone	58533
good	21688
work	15447
great	12694
use	12628
get	11840
buy	8809
one	8439
like	8182
screen	7570
love	7265
battery	7008
just	6891
time	5770
new	5586
come	5318
camera	5051
product	4732
price	4716
app	4632

Figure 2: Top Twenty Frequent Words

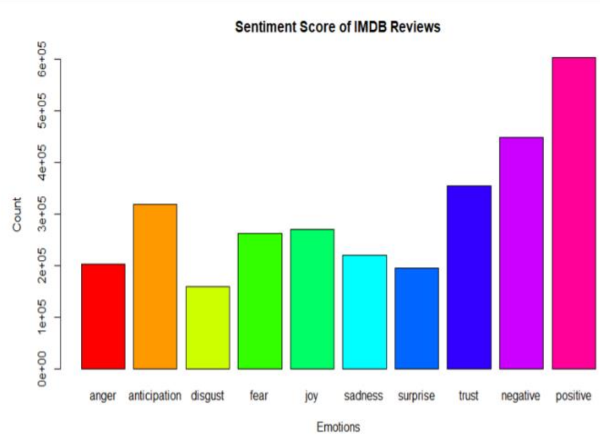

Figure 3: Sentiment Score of IMDB Reviews

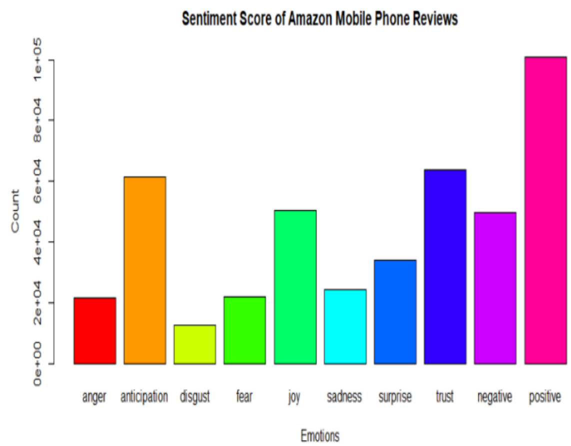

Figure 4: Sentiment Score of Amazon Phone Reviews

RESEARCH METHODOLOGY

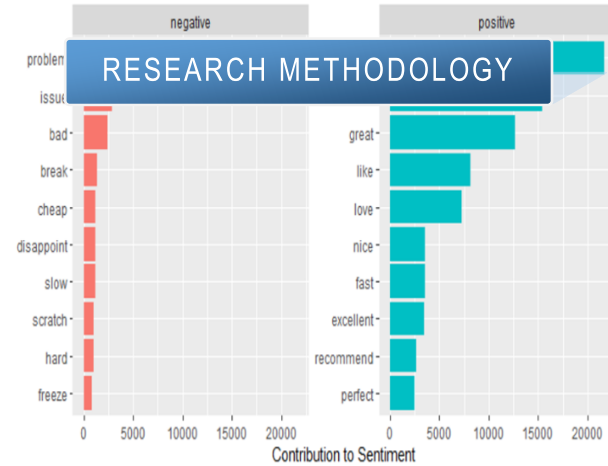

Figure 5: Sentiment Score of Amazon Phone Reviews

Figure 6: Sentiment Analysis of IMDB Reviews

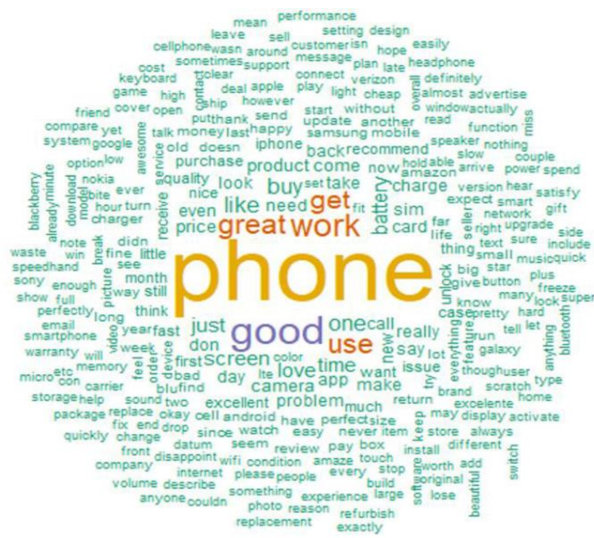

Figure 7: Sentiment Analysis of Amazon Phone Reviews

RESULT AND DISCUSSION

In this study [46][47], Tensor Flow was used to implement and train the models in the Keras framework. Python Jupyter was used on the Kaggle website to train the suggested model. Accuracy, precision, recall, and F1-score are the performance metrics that are used to assess the model's effectiveness.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

where

TN = True Negative
FP = False Positive
FN = False Negative
TP = True Positive

The training accuracy and training loss graph of the IMDB Reviews are shown in Figures 8 and Figure 9. Plotting the accuracy graph between the epoch and the values for training and validation accuracy reveals that when the epoch is at 0.50, the values for training and validation accuracy are 0.9770 and 0.8672, respectively.

Additionally, a loss graph that plots training loss against epoch and validation loss reveals that, for an epoch value of 0.50, training loss is 0.0690, and validation loss is 0.4441.

The training accuracy and training loss graph of the Amazon Mobile Phone Reviews is shown in Figures 10 and Figure 11. Plotting the accuracy graph between the epoch and the values for training and validation accuracy reveals that when the epoch is at 0.50, the values for training and validation accuracy are 0.9601 and 0.9114, respectively. Additionally, a loss graph that plots training loss against epoch and validation loss reveals that, for an epoch value of 0.50, training loss is 0.1139, and validation loss is 0.2375. The graph shows a better convergence of accuracy and loss towards the solution. However, the validation showed poor loss convergence. Lower precision will be the outcome, and this may be viewed as a future work scope. Calculated the F1-Score, Accuracy, Precision, and Recall using the IMDB and Amazon mobile phone reviews' confusion matrix [48] is shown in Figure 12 and Figure 13.

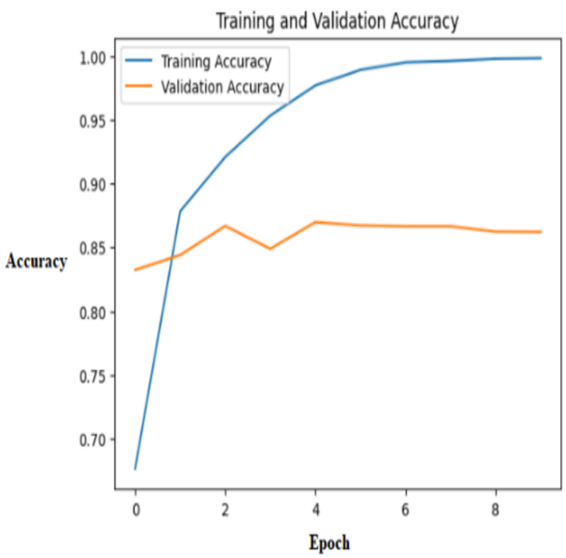

Figure 8: Training Accuracy and Validation Accuracy Of IMDB Review

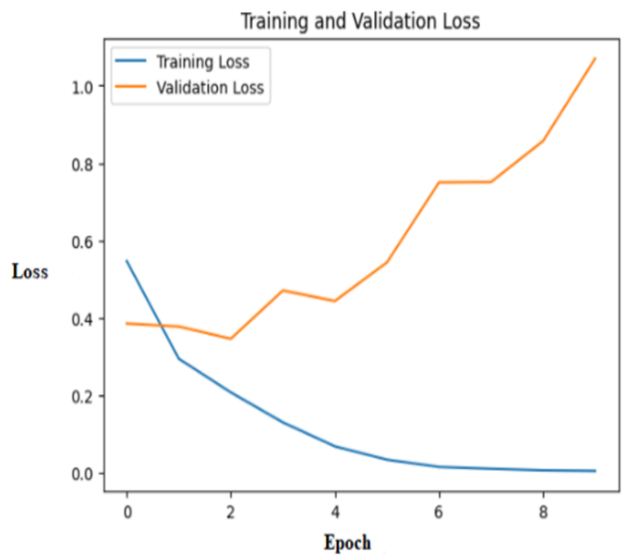

Figure 9: Training Loss and Validation Loss of IMDB Reviews

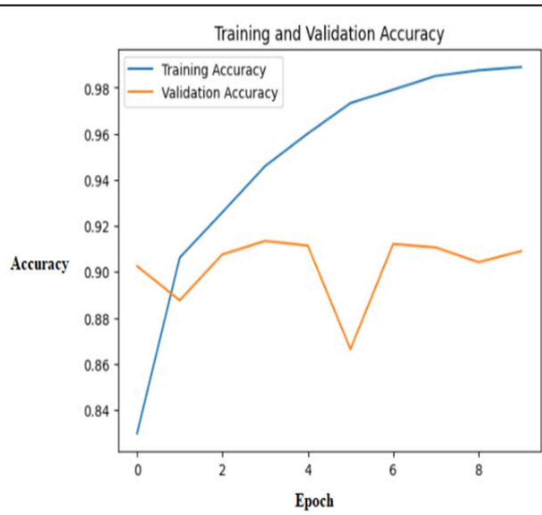

Figure 10: Training Accuracy and Validation Accuracy of Amazon Mobile Phone Reviews

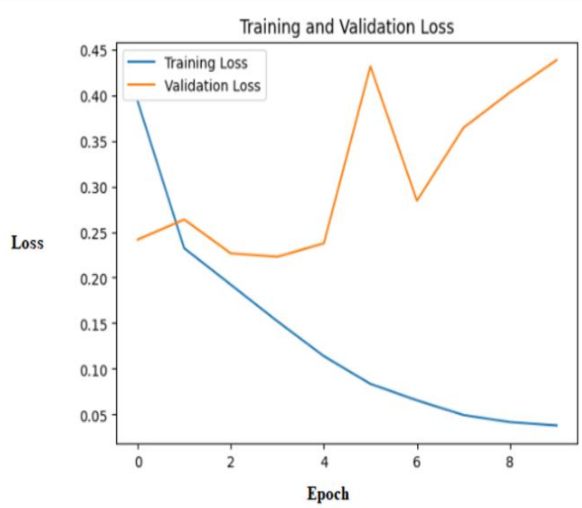

Figure 11: Training Loss and Validation Loss of Amazon Mobile Phone Reviews

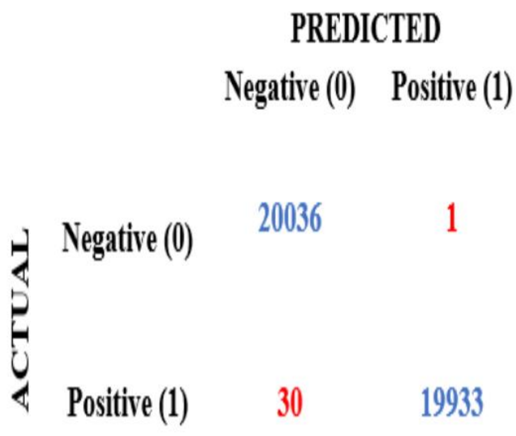

Figure 12: Confusion Matrix of IMDB Reviews Reviews

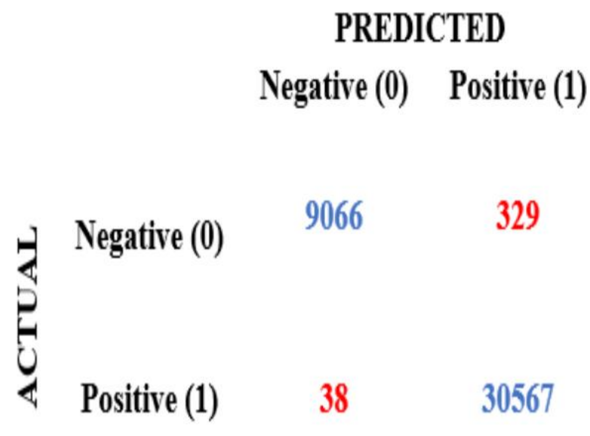

Figure 13: Confusion Matrix of Amazon Mobile Phone

Explaining the above-mentioned Confusion Matrices:

The confusion matrix shows how well the sentiment analysis algorithm classified text into positive and negative categories. The rows display the true sentiment of the text samples, while the columns display the emotion that the model predicted. The number of samples that belong to a given row (actual sentiment), as opposed to those that were expected to belong to a particular column (predicted sentiment), is displayed in each cell. In a well-designed confusion matrix, correct predictions would have high values on the diagonal, and incorrect predictions would have low values off the diagonal. This specific matrix seems to provide a respectable performance, with a significant percentage of appropriate classifications in each of the sentiment categories. Which sentiment analysis technique is used and what training set it utilizes can have a big impact on the confusion matrix.

Now, We Computed Accuracy, Precision, Recall, and F1-Score by using these confusion matrixes with the help of Accuracy, Precision, Recall and F1-Score Formulas:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1 Score} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
IMDB Reviews	99.92	99.99	99.84	99.92
Amazon Mobile Phone Reviews	99.08	98.93	99.40	98.93

Table 1: Results of Proposed Approach Implemented on Two Datasets

The accuracy of our proposed model is compared with BiLSTM, SVM+LSTM, LSTM+GloVe + GRU, LSTM, LSTM+FastText+SVM, Bidirectional CNN-RNN, ERNIE_BiGRU, and ours, which has accuracy of 72.25%, 73.81%, 81%, 89.80%, 90.75%, 93.40%, 94.48%, and 99.50%, respectively, in Table 2, which displays the comparative state of the art of various techniques. Table 1 indicates that the proposed approach [49] [50] has outperformed state-of-the-art models in terms of accuracy.

TECHNIQUE	ACCURACY
BiLSTM [13]	72.25%
SVM+LSTM [11]	73.81%
LSTM+GloVe+GRU [17]	81%
NB+Linear SVM [21]	82%
LSTM+BiLSTM [23]	84.74%
LSTM [15]	89.80%
LSTM+FastText+SVM [16]	90.75%
Bidirectional CNN-RNN [18]	93.40%
ERNIE_BiGRU [19]	94.48%
Our	99.50%

Table 2: Comparative State of Art

CONCLUSION

In this study, we have suggested a CNN-GRU-based Word2Vec sentiment classification method for text data. Global users openly share and express their thoughts on a variety of subjects. Large volumes of such data are very difficult to analyze manually, hence there is a legitimate demand for computer processing of this data. Sentiment analysis examines how individuals feel about various goods and services, political issues, social gatherings, and business tactics. Textual documents comprising reviews (from TripAdvisor, Amazon, and IMDB) and social network posts (primarily from Facebook and Twitter) hold the highest potential for sentiment analysis. Applied to two different datasets, our proposed CNN-GRU with Word2Vec DL Models achieve higher sentiment classification results, with 99.50% accuracy. The CNN-GRU model's Word2Vec embeddings can improve sequential data analysis and natural language processing abilities. The CNN-GRU model can perform better in a variety of applications by generating a more complex comprehension and interpretation of text input by utilizing contextual information from Word2Vec and semantic embeddings. The combination of CNN-GRU and Word2Vec embeddings presents a dynamic research environment with multiple chances to push the boundaries of multi-modal learning, natural language processing, and sequential data analysis. Contextual data, semantic embeddings, and sophisticated modeling approaches enable the CNN-GRU model to solve practical issues and offer fresh perspectives in a range of applications.

REFERENCE

[1] R. R. Karwa and S. R. Gupta, “Automated hybrid Deep Neural Network model for fake news identification and classification in social networks,” J. Integr. Sci. Technol., vol. 10, no. 2, pp. 110–119, 2022.

[2] M. Heikal, M. Torki, and N. El-Makky, “Sentiment analysis of Arabic tweets using deep learning,” Procedia Comput. Sci., vol. 142, pp. 114–122, 2018.

[3] H. Peng, E. Cambria, and A. Hussain, “A review of sentiment analysis research in Chinese language,” Cognit. Comput., vol. 9, pp. 423–435, 2017.

[4] X. Fang and J. Zhan, “Sentiment analysis using product review data,” J. Big Data, vol. 2, no. 1, pp. 1–14, 2015.

[5] A. Yadav and D. K. Vishwakarma, “Sentiment analysis using deep learning architectures: a review,” Artif. Intell. Rev., vol. 53, no. 6, pp. 4335–4385, 2020.

[6] O. Oueslati, E. Cambria, M. Ben HajHmida, and H. Ounelli, “A review of sentiment analysis research in Arabic language,” Futur. Gener. Comput. Syst., vol. 112, pp. 408–430, 2020.

[7] W. Aljedaani et al., “Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry,” Knowledge-Based Syst., vol. 255, p. 109780, 2022.

[8] C. N. Dang, M. N. Moreno-García, and F. De la Prieta, “Hybrid deep learning models for sentiment analysis,” Complexity, vol. 2021, pp. 1–16, 2021.

[9] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. J. Passonneau, “Sentiment analysis of Twitter data,” in Proceedings of the workshop on language in social media (LSM 2011), 2011, pp. 30–38.

[10] Z. Gao, A. Feng, X. Song, and X. Wu, “Target-dependent sentiment classification with BERT,” Ieee Access, vol. 7, pp. 154290–154299, 2019.

[11] A. P. Rodrigues, R. Fernandes, A. Shetty, K. Lakshmana, and R. M. Shafi, “Real-time Twitter spam detection. and sentiment analysis using machine learning and deep learning techniques,” Comput. Intell. Neurosci., vol. 2022, 2022.

[12] C. Sitaula, A. Basnet, A. Mainali, and T. B. Shahi, “Deep learning-based methods for sentiment analysis on Nepali COVID-19-related tweets,” Comput. Intell. Neurosci., vol. 2021, 2021.

[13] H. Elfaik and E. H. Nfaoui, “Deep bidirectional LSTM network learning-based sentiment analysis for Arabic text,” J. Intell. Syst., vol. 30, no. 1, pp. 395–412, 2020.

[14] A. Mabrouk and R. P. D. Redondo, “Deep Learning-Based Sentiment Classification: A Comparative Survey,” vol. 8, 2020, doi: 10.1109/ACCESS.2020.2992013.

[15] W. Hazım, G. Gwad, I. Mahmood, I. Ismael, and Y. Gültepe, “Twitter Sentiment Analysis Classification in the Arabic Language using Long Short-Term Memory Neural

Networks," Int. J. Eng. Adv. Technol., vol. 9, no. 3, pp. 235–239, 2020, doi: 10.35940/ijeat. b4565.029320.

[16] A. H. Ombabi, W. Ouarda, and A. M. Alimi, "Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks," Soc. Netw. Anal. Min., vol. 10, pp. 1–13, 2020.

[17] L. A. Alhuri, H. R. Aljohani, R. M. Almutairi, and F. Haron, "Sentiment analysis of COVID-19 on Saudi trending hashtags using recurrent neural network," in 2020 13th International Conference on Developments in eSystems Engineering (DeSE), IEEE, 2020, pp. 299–304.

[18] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis," Futur. Gener. Comput. Syst., vol. 115, pp. 279–294, 2021, doi: 10.1016/j.future.2020.08.005.

[19] X. Cheng, C. Zhang, and Q. Li, "Improved Chinese Short Text Classification Method Based on ERNIE_BiGRU Model," J. Phys. Conf. Ser., vol. 1993, no. 1, 2021, doi: 10.1088/1742-6596/1993/1/012038.

[20] Y. Li and H. Dong, "Text sentiment analysis based on feature fusion of convolution neural network and bidirectional long short-term memory network," J. Comput. Appl., vol. 38, no. 11, p. 3075, 2018.

[21] R. Bandana, "Sentiment analysis of movie reviews using heterogeneous features," in 2018 2nd International Conference on electronics, materials engineering & nanotechnology (IEMENTech), IEEE, 2018, pp. 1–4.

[22] R. I. Permatasari, M. A. Fauzi, P. P. Adikara, and E. D. L. Sari, "Twitter sentiment analysis of movie reviews using ensemble features based Naïve Bayes," in 2018 International Conference on Sustainable Information Engineering and Technology (SIET), IEEE, 2018, pp. 92–95.

[23] L. Chen, C. L. M. Chen, C. Lee, and M. Chen, "Exploration of social media for sentiment analysis using deep learning," Soft Comput., vol. 8, 2019, doi: 10.1007/s00500-019-04402-8.

[24] B. Bansal and S. Srivastava, "Sentiment classification of online consumer reviews using word vector representations," Procedia Comput. Sci., vol. 132, pp. 1147–1153, 2018.

[25] T. Hou and L. Zhou, "Chinese Ship Fault Relation Extraction Method Based on Bidirectional GRU Neural Network and Attention Mechanism," Comput. Sci, vol. 48, pp. 154–158, 2021.

[26] F. Bravo-Marquez, A. Khanchandani, and B. Pfahringer, "Incremental word vectors for time-evolving sentiment lexicon induction," Cognit. Comput., pp. 1–17, 2022.

[27] K. Kobs et al., "Emote-controlled: obtaining implicit viewer feedback through emote-based sentiment analysis on comments of popular twitch. tv channels," ACM Trans. Soc. Comput., vol. 3, no. 2, pp. 1–34, 2020.

[28] P. Mishra, R. Rajnish, and P. Kumar, "Sentiment analysis by novel hybrid method BE-CNN using convolutional neural network and BERT," Int. J. Adv. Trends Comput. Sci. Eng., vol. 9, no. 4, pp. 5320–5326, 2020.

[29] L. Khan, A. Amjad, K. M. Afaq, and H.-T. Chang, "Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media," Appl. Sci., vol. 12, no. 5, p. 2694, 2022.

[30] H. Ghulam, F. Zeng, W. Li, and Y. Xiao, "Deep learning-based sentiment analysis for roman Urdu text," Procedia Comput. Sci., vol. 147, pp. 131–135, 2019.

[31] N. Aslam, F. Rustam, E. Lee, P. B. Washington, and I. Ashraf, "Sentiment analysis and emotion detection on cryptocurrency related tweets using ensemble LSTM-GRU model," Ieee Access, vol. 10, pp. 39313–39324, 2022.

[32] S. Maity and K. Sarkar, "Topic Sentiment Analysis for Twitter Data in Indian Languages Using Composite Kernel SVM and Deep Learning," Trans. Asian Low-Resource Lang. Inf. Process., vol. 21, no. 5, pp. 1–35, 2022.

[33] A. Ishaq, S. Asghar, and S. A. Gillani, "Aspect-based sentiment analysis using a hybridized approach based on CNN and GA," IEEE Access, vol. 8, pp. 135499–135512, 2020.

[34] T. Siddiqui and A. Y. A. Amer, "A Comprehensive Review on Text Classification and Text Mining Techniques Using Spam Dataset Detection," Math. Comput. Sci. Vol. 2, pp. 1–17, 2023.

[35] V. Dogra, S. Verma, P. Chatterjee, J. Shafi, J. Choi, and M. F. Ijaz, "A complete process of text classification system using state-of-the-art NLP models," Comput. Intell. Neurosci., vol. 2022, 2022.

[36] U. Paralı, M. Zontul, and D. Ç. Ertuğrul, "Information retrieval using the reduced row echelon form of a term-document matrix," J. Internet Technol., 2019.

[37] M. A. Kausar, A. Soosaimanickam, and M. Nasar, "Public sentiment analysis on Twitter data during COVID-19 outbreak," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 2, 2021.

[38] N. Sharma, "Sentiment Analysis using tidytext package in R," in 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), IEEE, 2018, pp. 577–580.

[39] A. I. Kabir, R. Karim, S. Newaz, and M. I. Hossain, "The Power of Social Media Analytics: Text Analytics Based on Sentiment Analysis and Word Clouds on R.," Inform. Econ., vol. 22, no. 1, 2018.

[40] N. Ofek, "Sentiment Analysis for Social Text," in Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook, Springer, 2023, pp. 801–831.

[41] J. Rani, A. R. Shah, and S. Ramachandran, "pubmed. mineR: An R package with text-mining algorithms to analyse PubMed abstracts," J. Biosci., vol. 40, pp. 671–682, 2015.

[42] M. D. Samad, N. D. Khounviengxay, and M. A. Witherow, "Effect of text processing steps on twitter sentiment classification using word embedding," arXiv Prepr. arXiv2007.13027, 2020.

[43] D. Jatnika, M. A. Bijaksana, and A. A. Suryani, "Word2vec model analysis for semantic similarities in english words," Procedia Comput. Sci., vol. 157, pp. 160–167, 2019.

[44] B. Guo, C. Zhang, J. Liu, and X. Ma, “Improving text classification with weighted word embeddings via a multi-channel TextCNN model,” Neurocomputing, vol. 363, pp. 366–374, 2019.
[45] A. B. Oncul, “LSTM-GRU Based Deep Learning Model with Word2Vec for Transcription Factors in Primates,” vol. 11, no. 1, pp. 42–49, 2023, doi: 10.17694/bajece.1191009.
[46] M. Zulqarnain, R. Ghazali, M. G. Ghouse, and M. F. Mushtaq, “Efficient processing of GRU based on word embedding for text classification,” JOIV Int. J. Informatics Vis., vol. 3, no. 4, pp. 377–383, 2019.
[47] U. D. G and G. C. Babu, “Sentiment Analysis on Twitter Data by Using Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM),” pp. 1–11, 2021.
[48] B. Gaye and A. Wulamu, “Sentiment Analysis of Text Classification Algorithms Using Confusion Matrix,”in Cyberspace Data and Intelligence, and Cyber-Living, Syndrome, and Health: International 2019 Cyberspace Congress, CyberDI and CyberLife, Beijing, China, December 16–18, 2019, Proceedings, Part I 3, Springer, 2019, pp. 231–241.
[49] S. Suman and S. Urolagin, “Age gender and sentiment analysis to select relevant advertisements for a user using CNN,” in Data Intelligence and Cognitive Informatics: Proceedings of ICDICI 2021, Springer, 2022, pp. 543–557.
[50] W. Li, Y. Xu, and G. Wang, “Stance detection of microblog text based on two-channel CNN-GRU fusion network,” IEEE Access, vol. 7, pp. 145944–145952, 2019.