International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I5P216
Volume 70 Issue 5, 131-144, May 2022

Review Article

A Bi-LSTM and GRU Hybrid Neural Network with BERT Feature Extraction for Amazon Textual Review Analysis

Madhuri V. Joseph

Research Scholar, Department of Computer Science, Banasthali University, Rajasthan, India.

Madhuri.joseph2005@gmail.com

Received: 12 February 2022 Revised: 04 April 2022 Accepted: 09 April 2022 Published: 22 May 2022

Abstract *- Nowadays, businesses move towards digital platforms for their product promotion and to improve their overall profit margin. Customer reviews determine the purchase decision of the specified products in the e-commerce system in this digital world. In this case, reviewing products before buying is the common scenario in this current world. It will help the buyers to buy quality products at affordable prices. On this basis, it is necessary to implement deep learning techniques to analyze the sentimental tweets of customers based on their product ratings. The study planned to propose an enhanced BERT algorithm in feature extraction on sentiment analysis with a large data set and hybridized deep Bi-LSTM-GRU neural network for the classification process.*

Consequently, amazon products' customer review datasets are employed for the analysis. The review about mobile phones is retrieved from amazon for the sentiment analysis. Initially, the data was preprocessed to increase the accuracy and performance of the classifier. Further, the feature extraction is done to reduce many data into accurate ones with BERT (Bidirectional Encoder Representations from Transformers) algorithm. It isn't easy to evaluate the review sentiments without efficient classification. After this process, the study analyses the performance of the deep Bi-LSTM-GRU neural network method with the existing method. Finally, the study concluded that the proposed algorithm achieved 97.87, 98.36, 98.89, and 98.47, respectively, based on the accuracy level, frequency, precision, and recall measures. These performance measures are higher than the existing algorithms developed in the studies. The study achieved more accuracy and efficiency through the proposed deep Bi-LSTM-GRU neural network method in sentiment analysis on mobile phone reviews in the amazon e-commerce system.

Keywords *- Sentiment analysis, Feature extraction, BERT, Bi-LSTM, GRU neural network.*

1. Introduction

In past decades, the world has moved towards online platforms, especially in the covid-19 pandemic situations. The e-commerce system has emerged, and the people are moved to online buying based on customer reviews and ratings; in this case, reviewing products before buying is the common scenario in this current world. It will help the buyers to buy quality products with affordable processes. It will reduce the cheating happening in the e-commerce system [1]. Based on this current situation, it is necessary to implement deep learning technique's to polarize the large data of reviews into positive, negative, and neutral. It will help the customers understand the product reviews easily to take their buying decisions on the particular product. On this basis, it is necessary to implement deep machine learning techniques to analyze the sentimental tweets of customers based on their ratings of products [2]. The study planned to propose a hybrid of deep Bi-LSTM-GRU neural networks for the classification process. It enhanced the BERT algorithm's feature extraction on sentiment analysis with many data sets [3]. The enhanced BERT algorithm will perform faster than the existing developed algorithms. The BERT protocols are commonly used in the ATSA (Aspect term sentiment analysis) task. The BERT (bidirectional encoder representations from transformers) algorithm is used in this stage for the accurate feature extraction process. It trained on 2.5 billion words and applied in bidirectional learning to acquire the context of words for both right to left and left to right simultaneously [4]. It was used in the bi-directional transformers to pre-train the unlabeled data and later fine-tuned the pre-trained data set. It is widely used to improve several natural language processing tasks like sentiment analysis, word segmentation, question answering, and named entity recognition. In the pre-training stage, the BERT algorithm has been applied to optimize the training process by this enhanced semantic process [4].

The Bi-directional LSTM is an extension of the existing LSTM method, which improve the performance of the sequence classification problem. Bi-LSTM trained two LSTMs instead of one LSTM method for forwarding and backward processing. For the data set training process in terms of speed, GRU (Gated recurrent units) are 29.3% faster than the existing methods for processing the data sets [5]. For this purpose, this study combined the Bi-LSTM method and GRU method for better efficiency and faster process in many data sets. Based on these advantages of each method, the study proposed a novel deep Bi-LSTM-GRU neural network method, which has been developed to train the dataset for sentiment analysis to get accurate and efficient results for the study[6].

1.1. Contributions of study

The following are the major contribution of the study,

- To enhance the feature extraction process in sentiment analysis through the enhanced BERT method
- To propose hybridized deep Bi-LSTM-GRU neural network for an effective classification process in sentiment analysis
- To compare the effectiveness of enhanced deep learning techniques with the existing methods in terms of accuracy, precision, frequency measure, and recall,
- To obtain high accuracy in sentiment analysis with a large number of data sets based on the customer reviews and ratings

1.2. Paper organization

The structure of the proposed study is organized in the following sections. Section II presents the related literature studies for the sentiment analysis, Bi-LSTM, GRU, and BERT algorithm. Section III discusses the proposed work to extract and classify data sets. Section IV elaborates the analysis of the results from the proposed method, and its conclusion is expressed in Section V.

2. Review of Literature

A product review plays an important role in e-commerce, which allows the customers to get information about product quality, price, and durability to take appropriate decisions on buying goods[7]. Using MRC (machine reading comprehension) on the formal documents, this paper targeted the possibility of turning customer reviews into a wide source of knowledge that will be exploited to answer the user questions. It is stated in nature RRC (review reading comprehension). There are no existing studies done on this area of RRC. The data set has been selected for the sentiment analysis in this process. Since the review, RC has limited training examples for the RRC. The study explored the novel post-training method with the popular language protocol BERT developed to enhance the performance of fine-tuning for RRC.

The BERT algorithm has been implemented in the post-training process on review-based tasks like aspect sentiment classification and aspect extraction in the sentiment analysis. Finally, the result of the study stated that the proposed post-training method with the BERT algorithm is effective compared to other existing protocols [8]. Similarly, [9] has investigated extraction from documents, which aimed to retrieve the association between the words in the paragraph. The relation extraction contains longer and more entities than the next level and sentence level. It leads to the document-level extraction process being the hardest one. The document-level- entity mask method with information types has been developed to solve these problems. It will mask each entity with special tokens. In this process, the BERT algorithm has been developed based on the model that predicts the relationship between the entities by processing the text once. The study tested the proposed model on the DocRED data set, the large-scale open domain-document level data set on relation extraction entities. However, the result of the study indicated that the manually annotated part of the DocRED dataset, the proposed method obtains 6% improvements compared to the existing state-of-the-art model, and the performance of the BERT protocol improved by 5% compared to the existing protocols. Accordingly, [10] has stated that the early sentiment analysis on the level of sentence and text believed that the text has only one sentence. The phenomenon would cover the details of sentiments, and it isn't easy to fine-tune the people's sentiments fully. It leads to wrong information in decision-making about the products.

Moreover, the existing sentiment analysis method mainly focused on the recurrent neural network and the attention mechanism. It lacked emotional sensitivity on aspect words. This study performed an effective sentiment analysis using ALM-BERT to construct the feature location model to overcome these gaps. In this study, the pre-trained BERT model has been developed to extract the information from the comment context. Aspect-based sentimental feature extraction method has been used to learn features from the aspect words and the information about the aspect words.

Nowadays, the BERT protocols have been widely used in the ATSA task. Furthermore, [11] suggested that the semantic relevance of aspect word hiding vectors and context hiding vectors are poor. These are easily affected by irrelevant words. Hence, the novel CGBN model has been proposed with the BERT algorithm in the sentiment analysis. Using this protocol, the aspect word hiding vectors and context hiding vectors have effective semantic relative data extracted instantaneously.

Moreover, the paper proposed a new interactive gating mechanism, the co-gate proposed method. Compared to the existing methods, it effectively reduced the noisy words and captured the emotional semantic features effectively. In this paper, the pre-training and post-training have been implemented in the BERT protocol to enhance the suggested model. Finally, the experiments and analysis proved that the BERT algorithm performed effectively in the pre-training and post-training process in the sentiment analysis. Consequently, [12] stated that aspect-based sentiment analysis, which is intended to predict the sentiment polarities of the provided information, is the challenging one in the research area. Recently, the enhanced BERT proposed method has been used in sentiment analysis. In many contexts, the input form for the BERT algorithm is only the word sequences that don't provide the extra contextual data.

To address these challenges, the study aimed to introduce the novel method of the GBCN model, which used the gating mechanism by the context-aware aspect embedding to control and enhance the BERT proposed method for the aspect-based sentiment analysis. In the first stage, the input text is fed into the BERT, and it refines the context embedding separately. These refined embeddings consist of the most significant and correlated data in the context. Then, the study employed the gating mechanism to control the propagation of features from the BERT output with the context-based embedding. The result of the study revealed that the achieving rate of the suggested model and enhanced BERT is 88% and 92.9%, respectively, compared to the existing methods. Likely, [13] stated that deep learning techniques have emerged as significant in artificial intelligence to derive insights and performance prediction processes from a large volume of data. Among many succeeding processes in deep learning methods, sentiment analysis is one of the most important [14]. Opinion mining or sentiment analysis is finding people's opinions about the goods and services of a particular business. This paper investigated sentiment analysis of the Amazon e-commerce system based on customer reviews using LSTM and GRU methods. The hybrid of the LSTM-GRU method has provided more efficiency and effectiveness in sentiment analysis. Finally, the study concluded that the prediction of sentiment tweets on Amazon Company had been done effectively using LSTM-GRU compared to the existing algorithms done in sentiment analysis [15].

The review by customers serves as feedback on the quality of products, enhancing business opportunities. This study aimed to predict consumer opinion about mobile phones based on their reviews. Moreover, it analyzes the most significant factors behind reviews which segregates even positive, negative, and neutral. It will increase the company's business and helps the buyers to make the right decisions based on the reviews. In this study, the data has been preprocessed, then converted into text to vectors using the feature extraction process, and moved to the classification process. In this stage, the following algorithm has been developed: logistic regression, SGD, naive base, and CNN.

Furthermore, the study applied the lime technique to suggest the analytical reason for reviews classified as negative-positive and neutral. Finally, the experiments revealed that CNN with the word2vec feature extraction technique provided the best result for balanced and unbalanced data sets. Moreover, the analysis for handwriting recognition using the 8 LSTM showed the scale of the analysis. The hyperparameters used during the analysis needed optimization, which was done by the random search technique. Later it was evaluated using the ANOVA framework. The results revealed that the model had improved performance compared with other relevant models.

Moreover, the tuning of hyperparameters can be adopted effectively by using them individually[16]. In another study, the hybrid approach of GRU and LSTM was adopted for the sentiment analysis. And as a result of the study, the approach had shown better performance than the current studies [17]. Deep Bidirectional Wavelet Sequence (DBLSTM-WS) is used for the classification along with the LSTM network. And evaluated with other variants of LSTM, and the results from the experiment showed the approach was effective for the classification [18].

3. Proposed System

This section explains the proposed flow based on the preprocessing, feature extraction, and classification method. The amazon product's customer reviews are the data set taken for the analysis. The review about mobile phones is retrieved from amazon for the sentiment analysis. Initially, the data was preprocessed to increase the accuracy and performance of the classifier. Further, the feature extraction has been done to reduce many data into accurate ones with the BERT (Bidirectional Encoder Representations from Transformers) algorithm. It is difficult to evaluate the review sentiments in the proposed study text mining without a classification process. The hybrid method has been used for classification; it combines deep BI-LSTM and GRU neural networks to analyze the huge amount of data with accurate results and low processing time.

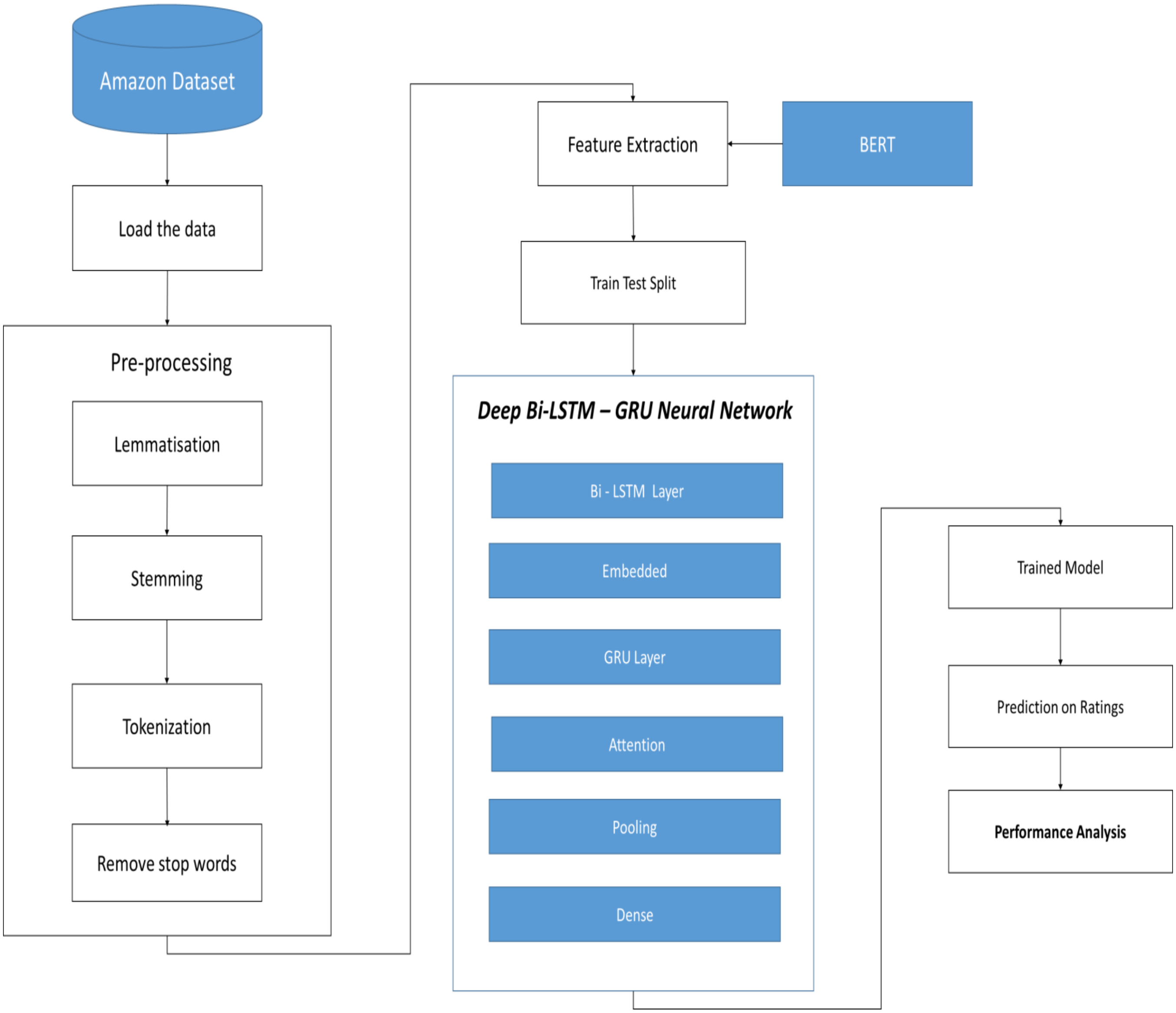

Fig. 1 The proposed flow of sentiment analysis

3.1. Data set Description

The data set has been taken from amazon online shopping customer reviews about unlocked mobile phones. The reviewer comments on their review based on a scale rating from 1 to 5, and they provide their ratings based on their purchase experience and product quality. The mean value has been calculated from the final product ratings. 4410 reviews of mobile phones sold through amazon.com have been taken for analysis. The product name, price, brand, ratings, and review text have been used for the data analysis process. The dataset has been taken from https://www.kaggle.com/. The amazon.com revealed above 4,00,000 reviews, and the 4410 reviews for unlocked mobile phones have been taken from the Amazon unlocked mobile.CSV file, size 132 MB. This data has been used to predict the sentiment reviews of customers based on the quality of unlocked mobile phones.

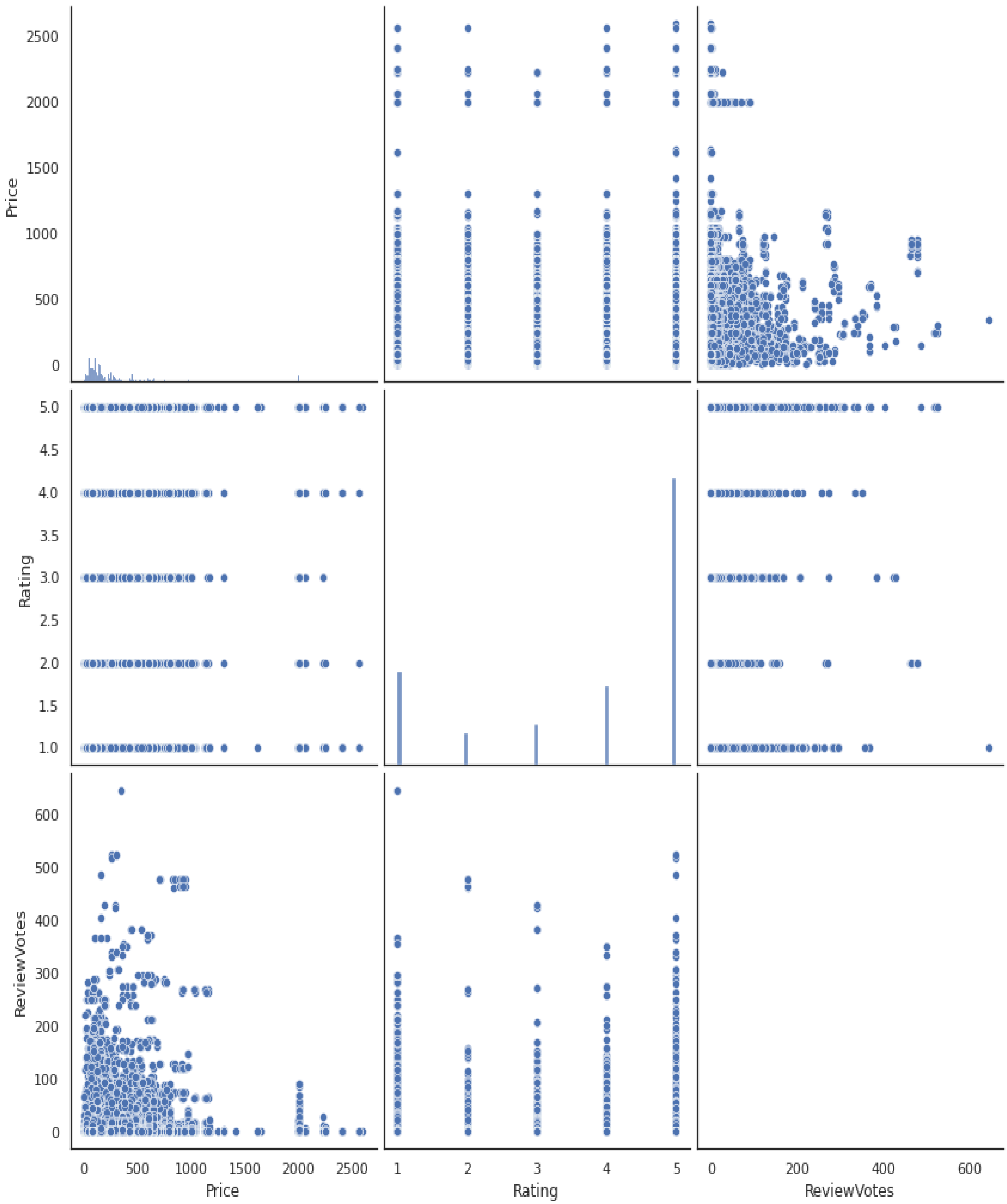

Fig. 2 Review votes distributed for different prices and different ratings

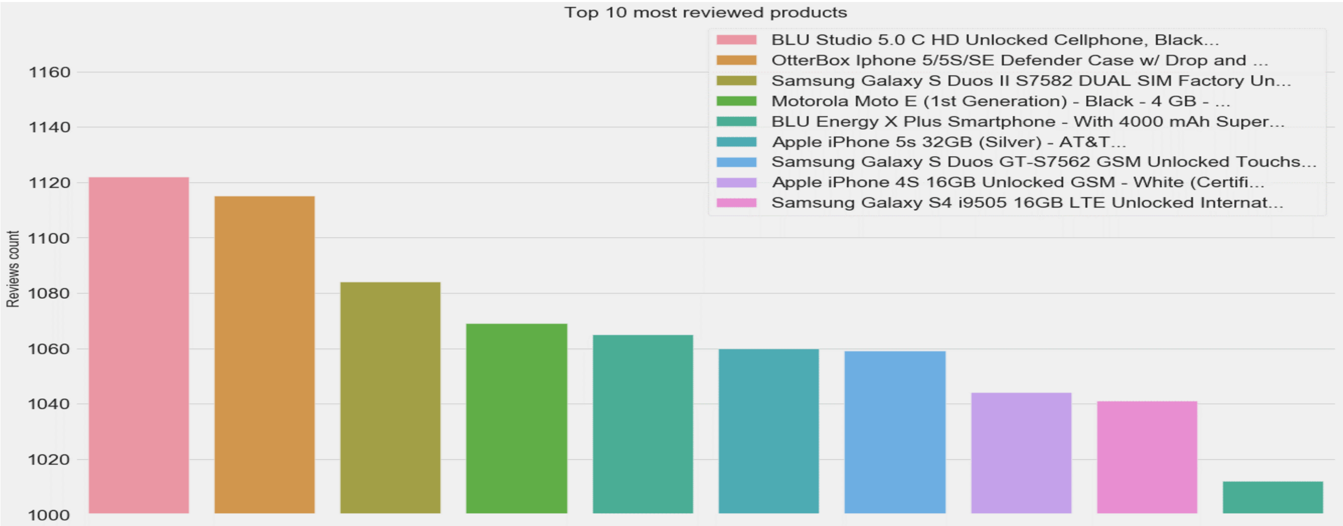

Fig. 3 Top 10 reviews on products

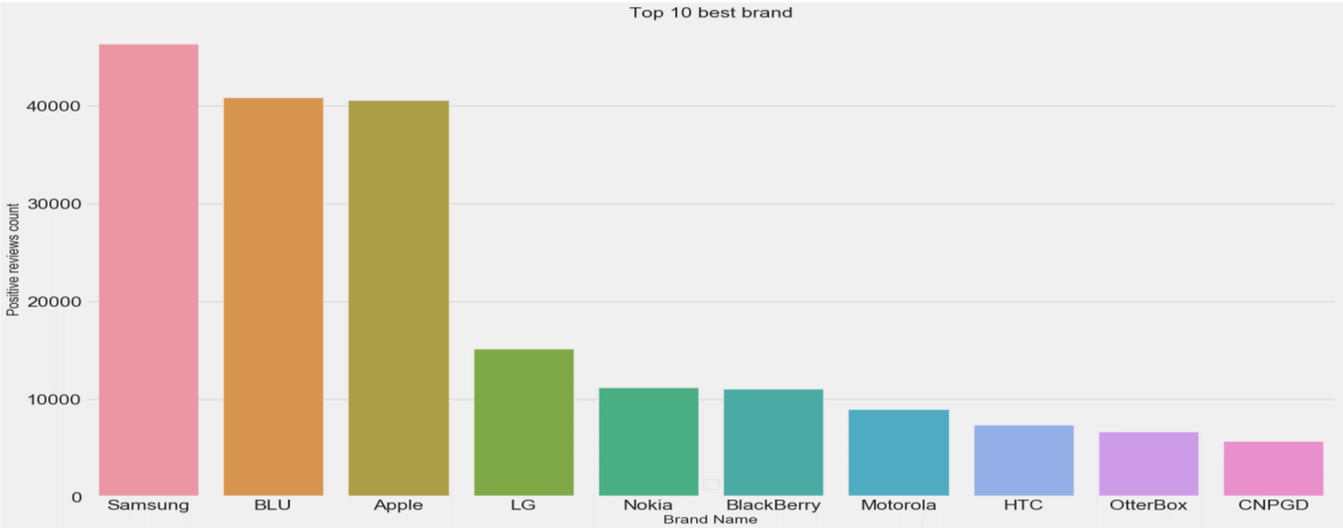

Fig. 4 Top 10 reviews on brands

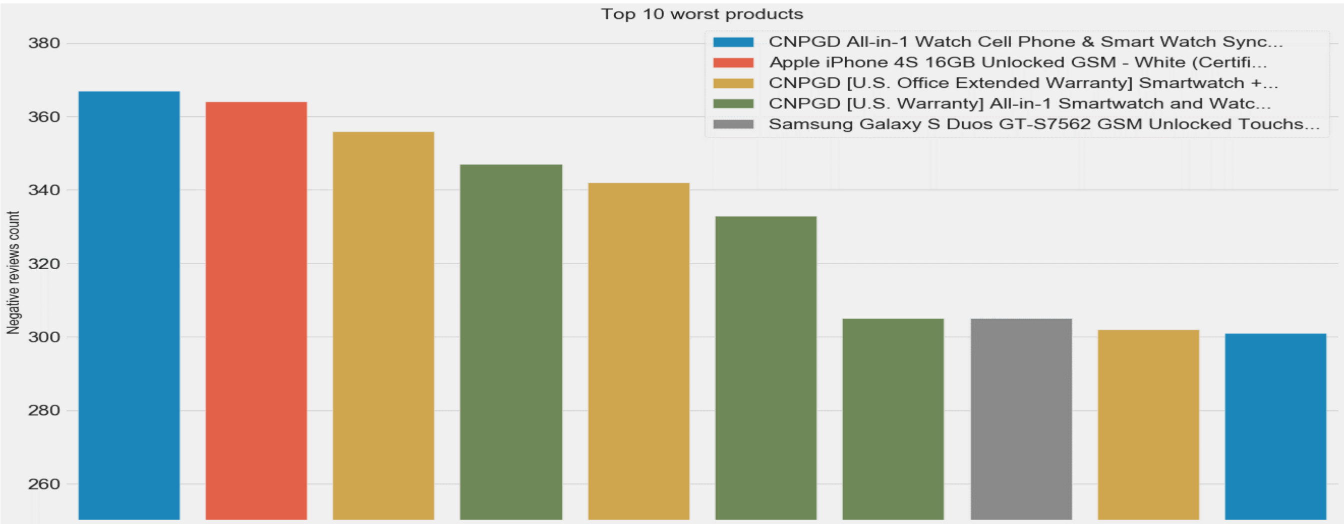

Fig. 5 Top 10 Negative Reviews On Products

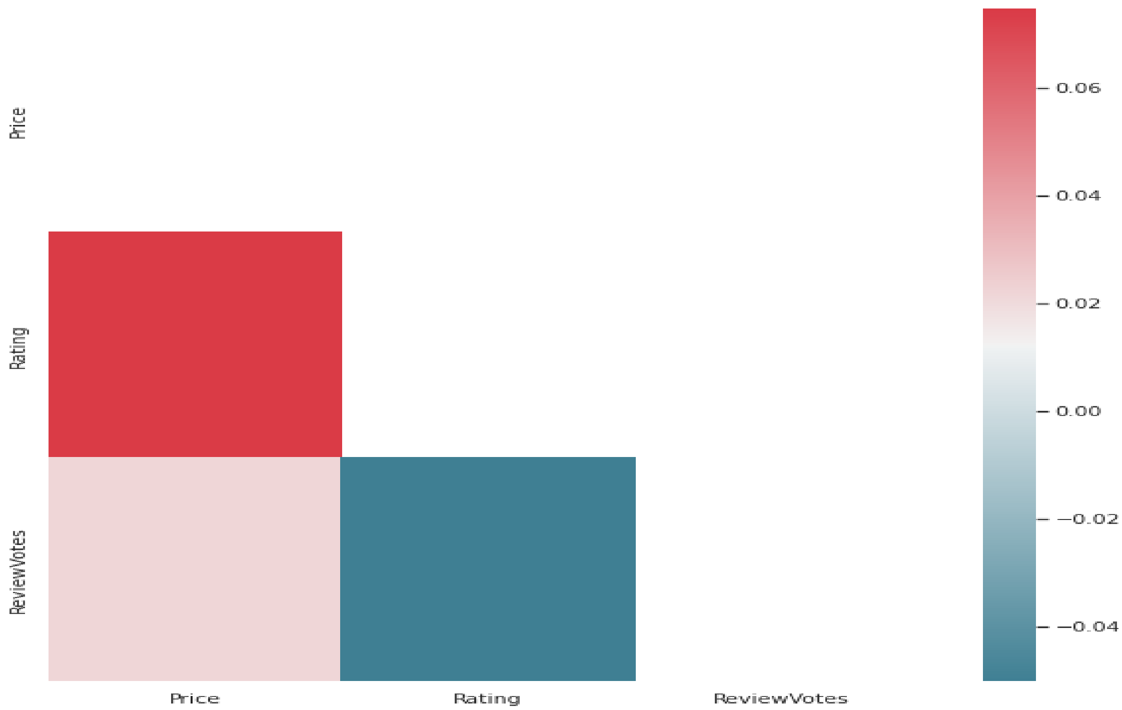

Fig. 6 Correlation matrix on ratings and price

3.2. Preprocessing process

In this process, the data has been preprocessed to increase the accuracy and performance of the classifier. The data has been included in the text reviews in the English language. The unclear text review data has been removed in this process, and the accurate data has been taken for the analysis. The difficult words to interpret and analyze are removed in this stage. Similarly, the punctuations, white spaces, and stop words are removed because it is difficult to process for data classification. The stop words like I, a, is, an, as, are, at, on, in, to, from, was. or, what, this, will. etc, are difficult to interpret for sentiment analysis. These are removed in this stage. All the text has been converted into the lower case in the preprocessing stage.

3.2.1. Lemmatization

Lemmatization is the first stage of text preprocessing; it reduces inflected words accurately, ensuring the root words with similar meanings. For example, sing, singing, and singing are all types of the word sing; it is the root word. Therefore these words changed into root words.

3.2.2. Stemming

Stemming is the second preprocessing process that reduces the words into their stem, affixes to prefixes, and suffixes. Stemming is an important process in natural language processing and natural language understanding. For example, sit and sitting sat are converted into stem words sit.

3.2.3. Tokenization

Tokenization is the process of segregating the sequence of sentences into individuals like phrases, keywords, and symbols are the tokens. In this process, the punctuation words have been discarded. The tokens are worked as the input process in the text preprocessing stage.

3.3. Feature extraction

Feature extraction is the process of reducing the number of features in the data set to increase the efficiency and performance of sentiment analysis. In this process, the proposed BERT algorithm has been used for an accurate feature extraction process. In this process, feature extraction identified the reviews by customers on the mobile phone ratings and the sentiment predictions that identified the opinions by deciding sentiments as negative, positive, and neutral. It was finally summarized after the classification process.

3.3.1. BERT algorithm

The BERT (bidirectional encoder representations from transformers) algorithm is used in this stage for the accurate feature extraction process. It trained on 2.5 billion words and applied in bidirectional learning to acquire the context of words for both right to left and left to right simultaneously. It was used in the bidirectional transformers to pre-train the large unlabeled corpus and fine-tune the pre-trained data set [19]. It is widely used to improve several natural language processing tasks like sentiment analysis, word segmentation, question answering, and named entity recognition. The BERT algorithm has been applied to optimize the pre-training process by this enhanced semantic task in the pre-training stage. The figure-1 states the overall architecture of the BERT algorithm and the transfer module. The input of the BERT algorithm consists of segment embedding, word piece embedding, and position embedding. The word embedding is initialized randomly, position embedding represents the token's positions, and the segment embedding represents the clauses in the text sequences [20].

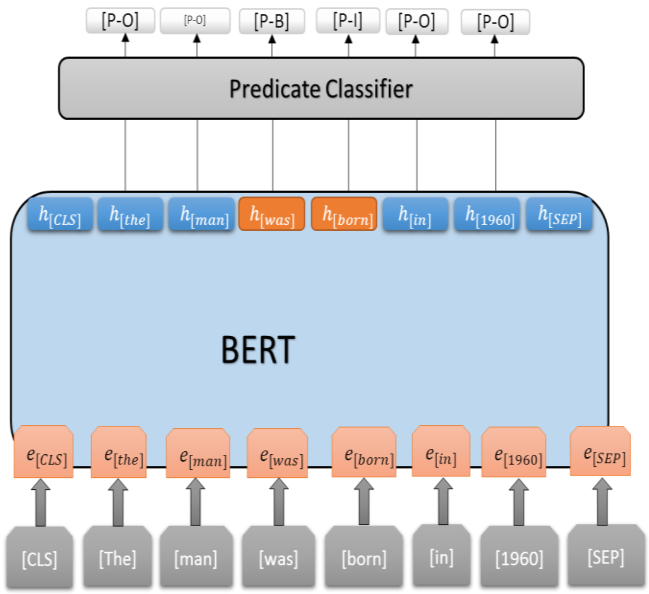

Fig. 7 BERT architecture

The first token in the process is a special classification token {CLS}, for the multi-sentence input, used the special token to separate [21]. The BERT algorithm repeatedly encodes the inputs on the multiple transformer layers and uses feed-forward operations and multi-head attention to extract features.

For the text sequences $a = \{a_1, a_2, \dots, a_M\}$., the input matrix $A \in \mathbb{R}^{M \times d} d = d_w + d_s + d_p$ is the sum of dimensions. In the multi-layer transform module, the primary input is $g_o = a, then$

$$g_l = tf_{block\,(h_{t-1})}, \forall l \in [1, N] \quad (1)$$

The formula (1)tf_{block} above denoted the transformer module, and the N represents the number of layers. Given a corpus $c = \{c_1, c_2, \dots, c_n\}$,$c_i$ is a token. The major goal of this language module is to enlarge the likelihood functions, then

$$N(C) = \sum_i \log P\,(c_i | c_{i-1}, \dots, c_{i-k}; \theta) \quad (2)$$

It is depicted from the above formula; the k is the context window of the next token c_i, which is predicted with the conditional probability p, and the tf_{block} is used to module the distribution of target tokens that follows,

$$P(c) = softmax\,(h_N W_o^T) \quad (3)$$

The above BERT feature extraction process h_N is the state sequence of the last tf_{block} and the W_o^T output weight matrix θ. Then, the model parameter optimized the stochastic gradient descent.

3.3.2. Train test split

After extracting reliable features by BERT protocol, the data set has been split into training and testing data sets. Consequently, the training data set is subjected to the proposed classification process, which involves BILSTM and GRU algorithms. The data sets are subjected to an 80% training process, and 20% are moved to the testing process.

3.4. Classification process

It is difficult to analyze a tweet's sentiments in the sentimental analysis without classification. For this study, in the classification process, the hybrid model of the deep Bi-LSTM and GRU neural network algorithm has been applied for the accurate classification of huge data sets in sentiment analysis. In this stage, the following process is involved, which are embedded, GRU layer, attention, pooling, and dense.

3.4.1. Bi-LSTM protocol

A Bi-LSTM (Bidirectional Long short term memory) is a processing module sequence that contains two LSTMs. One takes inputs in the forward direction, and the other takes the inputs in the backward direction [22]. The Bi-LSTM works better than the other algorithms in capturing the long-term dependencies. It processes a huge amount of data with good accuracy.

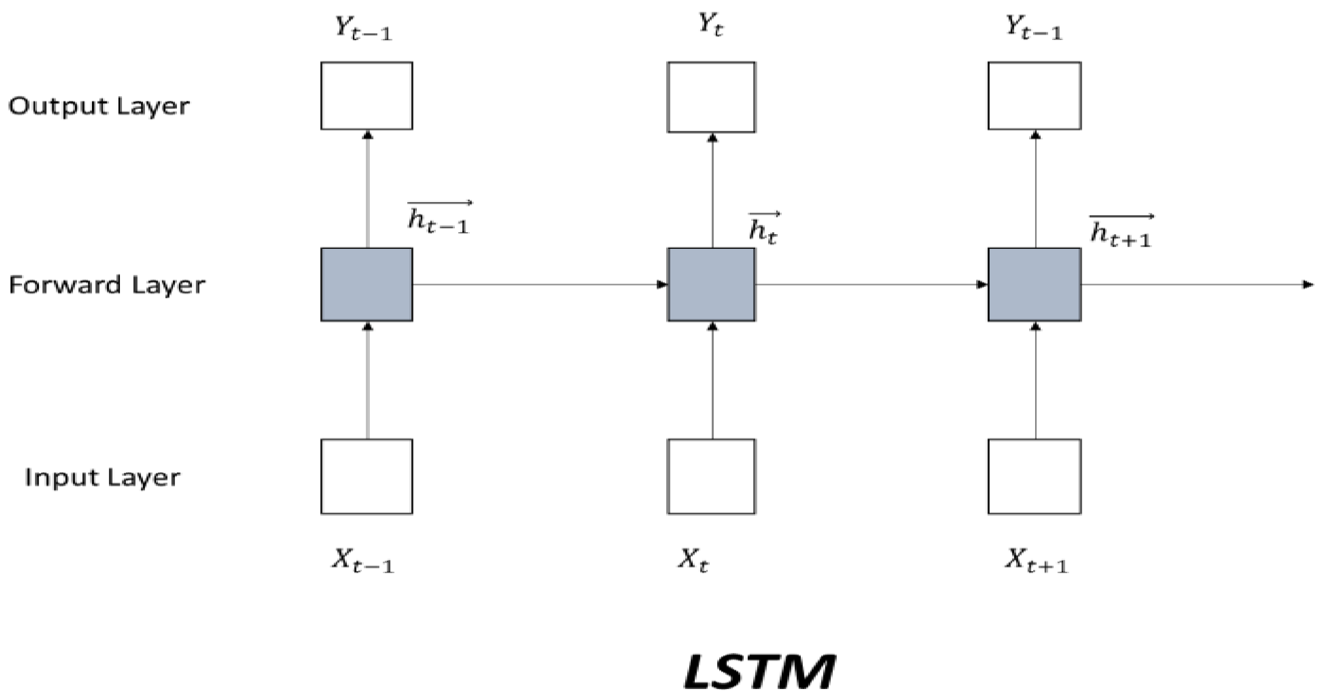

Fig. 8 LSTM architecture

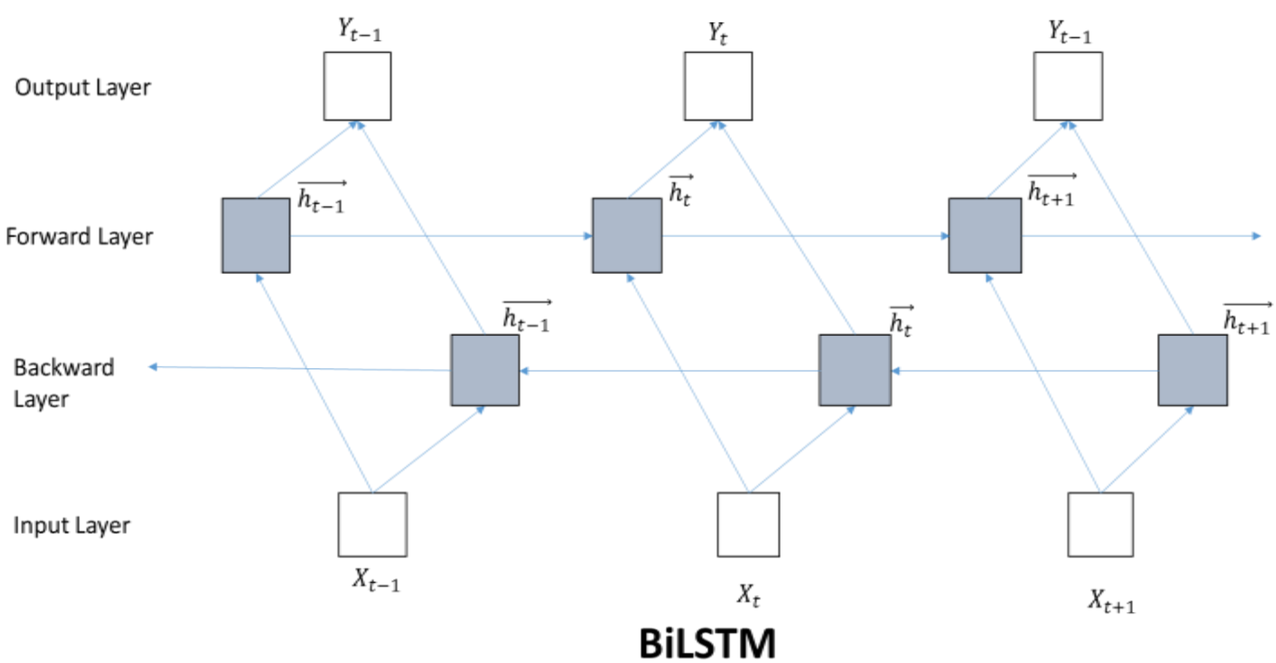

Fig. 9 Bi-LSTM architecture

The mathematical form of the BI-LSTM is depicted below;

$$i_t = \sigma(L_{ji} j_t + L_{hi} h_{t-1} + s_i) \quad (1)$$

$$f_t = \sigma(L_{jf} j_t + L_{hf} h_{t-1} + s_f) \quad (2)$$

$$o_t = \sigma(L_{jo} j_t + L_{ho} h_{t-1} + s_o) \quad (3)$$

$$\widetilde{d_t} = \tan h\,(L_{jc} j_t + L_{hc} h_{t-1} + s_d) \quad (4)$$

$$d_t = f_t \otimes d_{t-1} + i_t \otimes \widetilde{d_t} \quad (5)$$

$$h_t = o_t \otimes \tan h\,(d_t) \quad (6)$$

From the above formula, the f_t, i_t o_t and d_t denoted the value off, I, o, and d consequently. The s denoted the bias vector, and the L denoted the self-updating weights of the hidden layer. $\sigma(.)$*and* $\tanh(.)$They are the hyperbolic tangent and sigmoid functions, respectively. All the gate scores and the hidden layered outputs lie within the level of {0,1}. The operator $\otimes$represents the element-wise multiplications. The graphical representation of the standard LSTM protocol is shown in figure 2. This network can perform only to exploit historical context.

Moreover, it leads to an incomplete understanding of the nature of the problem that arises due to the lack of future context. Hence, the Bi-LSTM proposed algorithm has been developed to access both succeeding and proceeding contexts by merging the backward and forward hidden layers, as shown in the figure-3. The backward and forward pass over the unfolded networks, which is paralleled, and the Bi-LSTM has processed backward and forward hidden states in the network. The Bi-LSTM networks are trained using BPTT (backpropagation through time).

3.4.2. GRU protocol

The GRU (gated recurrent unit) protocol contains two gates: the update gate and the rest gate. These are the two vectors in the protocol that decides what data should be passed to the output, as shown in fig 10. In this process, the GRU protocol trains faster and perform better than the existing algorithms for less training data set in the language modeling process. Mathematically, the GRU proposed method has been described below with the architecture [23].

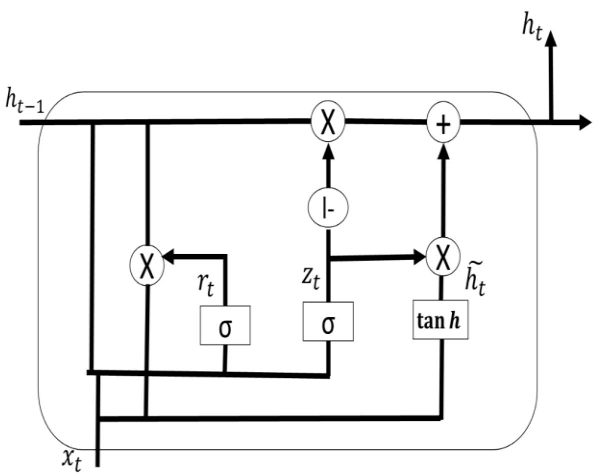

Fig. 10 The architecture of GRU

Calculating the update gate z_t for the time step t to help models to detect how much of the previous data sets need to be passed belongs to the future,

$$z_t = \sigma\left(L^z j_t + U^z h_{t-1}\right)$$

Calculating the reset gate using mathematical formulas to find how much of the previous data is needed to forget,

$r_t = \sigma(L^r j_t + U^r h_{t-1})$

Starting with the uses of the reset gate and the new memory content will use the recent gate to store the related data from the past,

$$h_t = \tan h\,(Lj_t + r_t \odot Uh_{t-1})$$

In this process, the network needs to find the network datah_t, which holds the data set for the current unit, and it is passed to the network. In order, the updated gate is needed for the process. It described what to gather from the current memory and what from the previous stageh_{t-1}.

$$h_t = Z_t \odot h_{t-1} + (1 - z_t) \odot h_t'$$

Finally,

$h_t = \sum_{k=1}^{t}[\prod_{j=k+1}^{t} \sigma\,(L_z I_j)]\,(1 - \sigma(L_z j_k))\ \tan h\,(L_{j_k})$

3.4.3. Deep Bi-LSTM – GRU Neural Network

In this stage, the deep Bi-LSTM-GRU neural network has been developed to train the dataset for sentiment analysis to get accurate and efficient results for the study. The Bi-directional LSTM is an extension of the existing LSTM method, which improve the performance of the sequence classification problem. Bi-LSTM trained two LSTMs instead of one LSTM method for forwarding and backward processing. For the data set training process in terms of speed, GRU (Gated recurrent units) are 29.3% faster than the existing methods for processing the data sets. For this purpose, this study combined the Bi-LSTM method and GRU method for better efficiency and faster process in many data sets. The following are the mathematical representation of the deep Bi-LSTM-GRU neural network method,

$\boldsymbol{FEATURES} \longleftarrow$ [Extracted features from the BERT]
$\boldsymbol{CLASSES} \longleftarrow$ $['Sentiment']$
$\boldsymbol{OUTPUT} \longleftarrow$ [Sentiment Scores]
$\boldsymbol{X} \longleftarrow\ dataset[FEATURES].values$
$\boldsymbol{Y} \longleftarrow\ dataset[CLASSES].values$

$$TRAIN\_DATA,\ TEST\_DATA,\ VALID\_DATA \leftarrow$$
$$TEST\_TRAIN\_SPLIT(X,\ Y,\ 0.80,\ 0.20)$$
$$BATCH\_SIZE \leftarrow 4$$

$$\boldsymbol{MODEL} \leftarrow$$
$$BiLSTM\_MODEL\left(\begin{bmatrix} EMBEDDING\_LAYER(TRAIN\_DATA.length, Output\_length, TRAIN\_DATA.columns), \\ BILSTM\_LAYER(Output\_length), \\ DENSE\_LAYER(units,\ activation = 'relu') \end{bmatrix}\right)$$

$$\boldsymbol{MODEL} \leftarrow$$
$$GRU\left(\begin{bmatrix} EMBEDDING\_LAYER(TRAIN\_DATA.length, Output\_length, TRAIN\_DATA.columns), \\ GRU(Output\_length), \\ DENSE\_LAYER(Output\_length,\ activation = 'sigmoid') \end{bmatrix}\right.$$

$$LOSS \leftarrow 'binary\_crossentropy',\quad optimizer \leftarrow 'adam', EPOCHS \leftarrow 250$$
$$MODEL.compile(LOSS,\ optimizer)$$

$$MODEL.train(TRAIN\_DATA,$$
$$EPOCHS,\ BATCH\_SIZE,\ VALID\_DATA)$$

4. Result and Discussion

In this study, the data set has been subjected to preprocessing, BERT- feature extraction, and classification process using of deep Bi-LSTM-GRU neural network. After this process, the result of the proposed method has been compared with the existing methods developed in sentiment analysis on mobile phone reviews in Amazon Company, as discussed below;

Table 4.1 Comparison of the proposed algorithm with existing studies

Paper title	**Author**	**Description**	**Accuracy**
Sentiment analysis on Mobile phone reviews using supervised learning techniques	[24]	hybrid of RF and TF-IDF method	91%
Aspect context-aware sentiment classification of online consumer reviews	[25]	SGD (stochastic gradient descent), logistic regression, naive Bayes, and CNN (convolutional neural networks)	92.72%
A Comparison of sentiment analysis method on amazon mobile reviews	[26]	SGD classifier, Gradient boosting classifier, multi-nominal NB, NB-SVM, LSTM, CNN, and random forest	85.5%
Proposed method	2021	deep Bi-LSTM-GRU neural network	97.87%

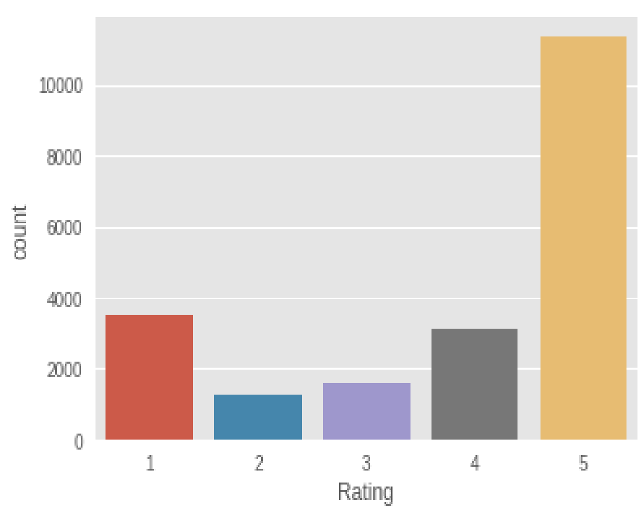

Fig. 11 Number of words used based on rating

Fig. 11 shows the number of words used based on rating. As the study has included the rating of the words used, fig 11 shows the words based on rating, and the words with a rating of five-star have counted more than the other rated words. And the least used rating was two-star has been utilized less than other rated words. And similarly, the exact number of words used without rating based on the count (mean) is shown in fig 12.

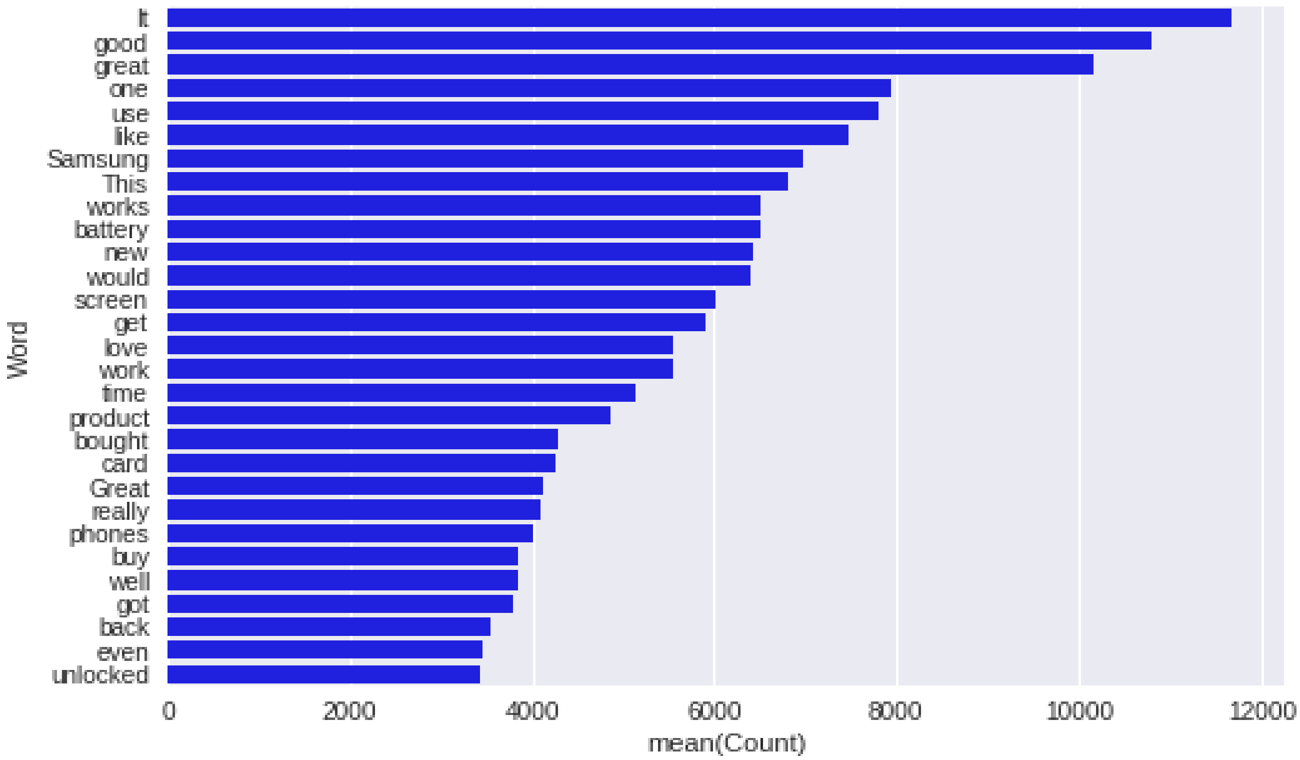

Fig. 12 Words used and their count

Table 1. Comparison of AUC-ROC with other earlier studies

Methods	AUC-ROC
Proposed	0.97
Existing	0.91
SG-Avg-SVM	0.84
SG-Avg-RF	0.86
CBOW-TFIDF-RF	0.85

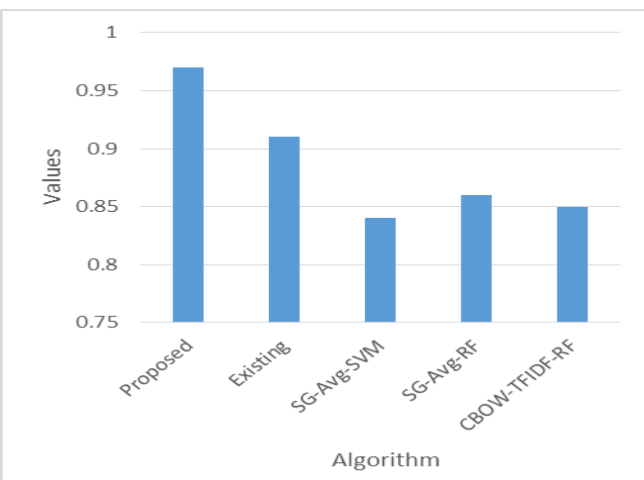

Fig. 13 AUC-ROC accuracy measures compared with existing study

From the above fig 13 and table 1, it is inferred that the proposed deep Bi-LSTM-GRU neural network has achieved a 0.97 accuracy level of AUC-ROC (Area under the ROC curve) compared to the existing method developed in [24]. This paper aimed to analyze aspect content level sentiment classification of optical character recognition (OCR) for deeper analysis. The study has developed a hybrid of RF and TF-IDF methods; based on this method, the study achieved 0.91 accuracies compared to the existing studies. In the study, the proposed method performed a high accuracy rate compared to this method which is 0.97.

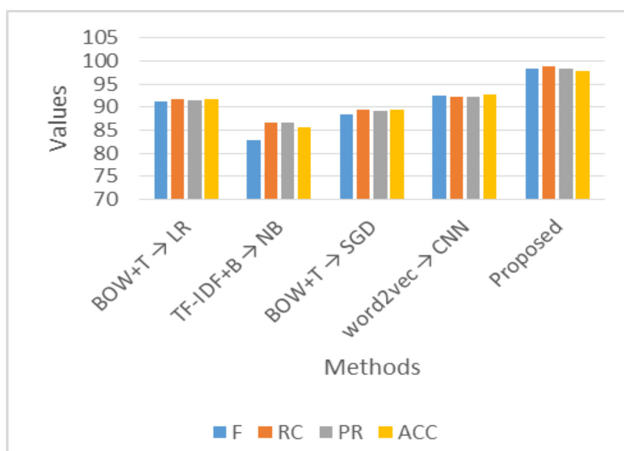

Fig. 14 Performance matrix proposed method and existing method

From fig 14, the proposed algorithm has achieved a 98.47% frequency level, 98.89% recall, 98.36% precision, and 97.87% inaccuracy, respectively. These performance ratings are comparatively higher than the existing study [25], which was developed to compare the sentiment analysis of reviews on mobile phones in the amazon e-commerce system. The study analyzed the performances of sentimental analysis through different machine learning techniques, such as SGD (stochastic gradient descent), logistic regression, naive Bayes, and CNN (convolutional neural networks). Compared to this study, the proposed method achieved high performance than the existing study represented in the above graph.

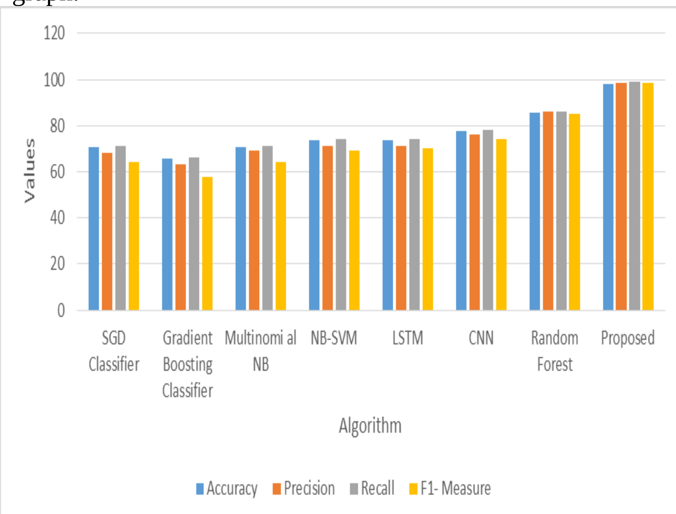

Fig. 15 Performance matrix proposed method and existing method

Fig 15 is depicted the study in terms of accuracy, precision, recall, and f1 measures, and the study has achieved 97.87, 98.36, 98.89, and 98.47, respectively. Compared to the existing algorithms such as SGD classifier, Gradient boosting classifier, multi-nominal NB, NB-SVM, LSTM, CNN, and random forest developed in the paper [26]. Finally, the deep Bi-LSTM-GRU neural network has performed higher than the existing algorithms.

5. Conclusion

The study aimed to analyze the sentimental tweets on mobile phone reviews in the amazon e-commerce system. The dataset has been acquired from amazon reviews provided by the customers. The negative, positive and neutral comments are taken for the analysis based on the ratings by customers. The selected data set has been preprocessed and subjected to a feature extraction process with an enhanced BERT algorithm method. The data set has been divided into training data set and a testing dataset in this process. Then the training data set has been moved to the classification process. And hence the proposed a deep Bi-LSTM-GRU neural network method to polarize the large amount of unlabeled product review data set. After this process, the study has discussed the performance of the deep Bi-LSTM-GRU neural network method with the existing method. Finally, the study concluded that the proposed algorithm achieved 97.87, 98.36, 98.89, and 98.47, respectively, based on the accuracy level, frequency, precision, and recall measures. These performance measures are higher than the existing algorithms developed in the studies. The study achieved more accuracy and efficiency through the proposed deep Bi-LSTM-GRU neural network method in sentiment analysis on mobile phone reviews in the amazon e-commerce system.

Moreover, the study provides an accurate result through this sentiment analysis, which will help the customers find positive and negative reviews. It helps to take correct decisions on buying goods through the e-commerce system. Hence, the study suggested some future works in this area, improving sentiment analysis and being more effective in practical life. The study suggested implementing deep learning and principal component analysis for a fully automated data labeling process.

Conflicts of Interest

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

Funding Statement

The author(s) declare(s) that there is no funding for the publication of this paper.

References

[1] A. S. AlQahtani, Product Sentiment Analysis for Amazon Reviews, International Journal of Computer Science & Information Technology (IJCSIT) , 13 (2021).
[2] H. S. Sharaf Al-deen, Z. Zeng, R. Al-Sabri, and A. Hekmat, An Improved Model for Analyzing Textual Sentiment Based on a Deep Neural Network Using Multi-Head Attention Mechanism, Applied System Innovation, 4 (2021) 85.
[3] D. Mane, P. Srivastava, A. Jain, and D. Chouhan, Assessing Travellers Perception Levels towards Select Online Travel Agents (TA) in India using Netnography and Text Mining, International Journal of Engineering Trends and Technology, 68 (2020).
[4] B. Subba and S. Kumari, A heterogeneous stacking ensemble based sentiment analysis framework using multiple word embeddings, Computational Intelligence, (2021).
[5] S. K. Prabhakar, H. Rajaguru, and D.-O. Won, Performance Analysis of Hybrid Deep Learning Models with Attention Mechanism Positioning and Focal Loss for Text Classification, Scientific Programming, 2021 (2021).
[6] S. Kavatagi and V. Adimule, Bi-GRU Model with Stacked Embedding for Sentiment Analysis: A Case Study, in the Internet of Things, Artificial Intelligence and Blockchain Technology, ed: Springer, (2021) 259-275.

[7] S. Muthukumaran and P. Suresh, Text analysis for product reviews for sentiment analysis using NLP methods, Int. J. Eng. Trends Technol, 47 (2017) 474-480.
[8] H. Xu, B. Liu, L. Shu, and P. S. Yu, BERT post-training for review reading comprehension and aspect-based sentiment analysis, arXiv preprint arXiv:1904.02232, (2019).
[9] X. Han and L. Wang, A novel document-level relation extraction method based on BERT and entity information, IEEE Access, 8 (2020) 96912-96919.
[10] G. Pang, K. Lu, X. Zhu, J. He, Z. Mo, Z. Peng, et al., Aspect-Level Sentiment Analysis Approach via BERT and Aspect Feature Location Model, Wireless Communications and Mobile Computing, 2021 (2021).
[11] Y. Peng, T. Xiao, and H. Yuan, Cooperative gating network based on a single BERT encoder for aspect term sentiment analysis, Applied Intelligence, (2021) 1-13.
[12] X. Li, X. Fu, G. Xu, Y. Yang, J. Wang, L. Jin, et al., Enhancing BERT representation with context-aware embedding for aspect-based sentiment analysis, IEEE Access, 5 (2020) 46868-46876.
[13] G. More, H. Behara, and A. M. Suresha, Sentiment Analysis on Amazon Product Reviews with Stacked Neural Networks, (2020).
[14] C. Du, H. Sun, J. Wang, Q. Qi, and J. Liao, Adversarial and domain-aware BERT for cross-domain sentiment analysis, in Proceedings of the 58th annual meeting of the Association for Computational Linguistics, (2020) 4019-4028.
[15] K. Baktha and B. Tripathy, Investigation of recurrent neural networks in the field of sentiment analysis, in International Conference on Communication and Signal Processing (ICCSP), (2017) 2047-2050.
[16] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, LSTM: A search space odyssey, IEEE transactions on neural networks and learning systems, 28 (2016) 2222-2232, 2016.
[17] B. Athiwaratkun and J. W. Stokes, Malware classification with LSTM and GRU language models and a character-level CNN, in 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), (2017) 2482-2486.
[18] Ö. Yildirim, A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification, Computers in biology and medicine, 96 (2018) 189-202.
[19] O. Habimana, Y. Li, R. Li, X. Gu, and Y. Peng, A Multi-Task Learning Approach to Improve Sentiment Analysis with Explicit Recommendation, in 2020 International Joint Conference on Neural Networks (IJCNN), (2020) 1-8.
[20] C. N. Dang, M. N. Moreno-García, and F. De la Prieta, Using Hybrid Deep Learning Models of Sentiment Analysis and Item Genres in Recommender Systems for Streaming Services, Electronics, 10 (2021) 2459.
[21] A. Oshri, S. M. Kogan, J. A. Kwon, K. Wickrama, L. Vanderbroek, A. A. Palmer, et al., Impulsivity as a mechanism linking child abuse and neglect with substance use in adolescence and adulthood, Development and psychopathology, 30 (2018) 417-435.
[22] S. Sachin, A. Tripathi, N. Mahajan, S. Aggarwal, and P. Nagrath, Sentiment analysis using gated recurrent neural networks, SN Computer Science, 1 (2020) 1-13.
[23] N. Pathik and P. Shukla, An efficient sentiment analysis using topic model based optimized recurrent neural network, Int. J. Smart Sensing Intell. Syst., 11 (2021) 1-12.
[24] B. Bansal and S. Srivastava, Aspect context aware sentiment classification of online consumer reviews, Information Discovery and Delivery, (2020).
[25] S. A. Aljuhani and N. S. Alghamdi, A comparison of sentiment analysis methods on Amazon reviews of Mobile Phones, Int. J. Adv. Comput. Sci. Appl, 10 (2019) 608-617.
[26] M. Shaheen, S. M. Awan, N. Hussain, and Z. A. Gondal, Sentiment analysis on mobile phone reviews using supervised learning techniques, International Journal of Modern Education & Computer Science, 11 (2019).