Aspect based sentiment analysis using fine-tuned BERT model with deep context features

Abraham Rajan, Manohar Manur
Department of CSE, CHRIST (Deemed to be University), Bangalore, India

Article Info

Article history:

Received Sep 30, 2022
Revised Nov 2, 2023
Accepted Dec 5, 2023

Keywords:

BERT
DC-BERT
Interaction layer
NLP
Sentiment analysis

ABSTRACT

Sentiment analysis is the task of analysing, processing, inferencing and concluding the subjective texts along with sentiment. Considering the application of sentiment analysis, it is categorized into document-level, sentence-level and aspect level. In past, several researches have achieved solutions through the bidirectional encoder representations from transformers (BERT) model, however, the existing model does not understand the context of the aspect in deep, which leads to low metrics. This research work leads to the study of the aspect-based sentiment analysis presented by deep context bidirectional encoder representations from transformers (DC-BERT), main aim of the DC-BERT model is to improvise the context understating for aspects to enhance the metrics. DC-BERT model comprises fine-tuned BERT model along with a deep context features layer, which enables the model to understand the context of targeted aspects deeply. A customized feature layer is introduced to extract two distinctive features, later both features are integrated through the interaction layer. DC-BERT mode is evaluated considering the review dataset of laptops and restaurants from SemEval 2014 task 4, evaluation is carried out considering the different metrics. In comparison with the other model, DC-BERT achieves an accuracy of 84.48% and 92.86% for laptop and restaurant datasets respectively.

Corresponding Author:

Abraham Rajan
Department of CSE, Research Scholar, Christ University
Bangalore, India
Email: abraham.rajan@res.christuniversity.com

1. INTRODUCTION

Natural language processing (NLP) is a part of machine learning that provides the ability of computers to learn and understand text, script, and words that are spoken. The understanding of these systems is in the same way as human understanding. Sentiment analysis is a technology that is applied to unstructured texts, where the sentiment in this information is extracted [1]. Sentiment analysis is a branch of natural language processing that is mainly applied in the field of data mining and machine learning. This is applied widely in news, politics, and the educational field [2]. The expanding social media platforms also increase the demand and use of sentiment analysis worldwide. There are the same words that can be used in various contexts. This challenging task of retrieving the sentiment information along with the use of similar words in different contexts is achieved due to the growth of NLP over the years [3]. The detection of emotions and the sentiments expressed in any written or spoken text is also referred to as open mining, which is termed sentiment analysis. It is mainly bifurcated into sentiment information that is expressed as neutral, negative, or positive about any statement [4]. There are also physiological risks that are present in various social media platforms, which can be avoided with the use of sentiment analysis. Reviews that are collected from customers for various events such as movies, restaurants, merchandise, foods, or various applications can be automatically detected with the growth of sentiment analysis [5]. The classification of a document or text is done on the document level, this level of classification is not applicable practically. Whereas the emotional classification is based on an aspect level. This level of classification has more applications in the real-time world [6]. The level of subjective information that flows on the web on various social media has a massive impact that leads to huge consequences. The positive benefits of this textual information are the retrieval of information from these reviews and comments that are posted on the web. The growth of a business, prediction of politics, education, society, and medical fields relating to psychology. cause the need for this subjective information to be automatically detected and segregated [7]. The emerging e-commerce platforms due to digitalization in every field also pose a need for this methodology to be developed and improved over time. There occur comments and reviews that may have both negative and positive polarity in them, whereas not all reviews need to be negative or positive. The reviews can also be neutral. Hence, this requires the analyses to be performed on an aspect level [8], several studies have been carried out to analyze and classify the sentiment based on aspects. Figure 1 shows the general framework for aspect-based sentiment analysis.

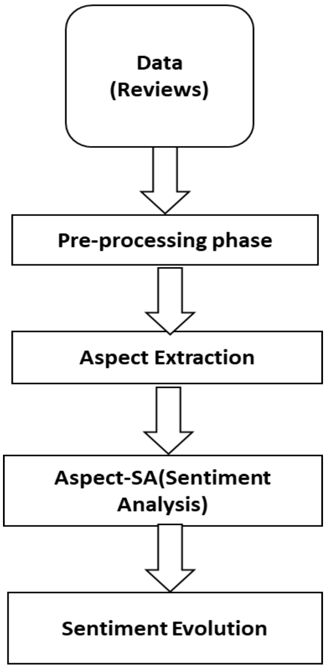

Figure 1. General framework for aspect-based sentiment analysis

Figure 1 shows the general framework of aspect-based sentiment analysis, it comprises five modules, the first module is where the dataset is designed, in this case, review data is selected along with targeted aspects. The second module includes the pre-processing phase; the third module includes the identification of targeted aspects from given sentences or documents, the fourth module shows the analysis of sentiment and the fifth module discusses the sentiment evaluation and classification into the different categories as positive, negative, or neutral. Moreover, bidirectional encoder representation from transformer (BERT) [9] has been one of the successful adoptions in NLP for sentiment analysis based on aspects. However, despite the effectiveness of the BERT model, aspect-based sentiment analysis remains a major challenge in the real-time scenario based on major three reasons. The first reason is that there is the enormous growth of social media data, which causes substantial barriers as aspect level- sentiment adoption for a new domain, is challenging due to limited labelled data. The second reason is existing BERT approach utilized a uniform model across domains such as “appearance”, and “performance”, for the laptop dataset and “service”, “food”, and “price” in the case of the restaurant domain. The third reason is the role of contextual information as it has been given little attention. BERT model is designed for pre-training deep bidirectional representations through unlabeled text by joining the conditioning on right context and left context in all layers, thus BERT model can be fine-tuned by adding the additional layer for various ranges of tasks [10].

- Motivation and contribution

Aspect-based sentiment analysis is a fundamental task in SA, it is divided into two categories i.e. aspect extraction and classification. Moreover, it refers to the identification of opinions or feelings about a particular entity. Rapid deployment in a neural network in recent years has shown great growth in deep learning models, BERT model has proven to capture the word features of a particular word in various contexts. However, the selection of an ideal number of a parameter is important for high accuracy, hence this research work proposes fine-tuned BERT model for sentiment classification. Further, the contribution of this research work has been highlighted here.

- This research work utilizes the BERT model and proposes the DC-BERT model for aspect-based sentiment analysis; the DC-BERT model comprises fine-tuned BERT model, which improvises the traditional BERT model; also, it introduces a deep context feature layer combined DC-BERT model.
- DC-BERT model is designed to extract two distinctive customized features; these two features includes a deep understanding of context based on words and a general understanding of the sentence.
- Further, deep context features are adopted to understand the context of targeted aspects; the concatenation layer is used for combining deep features and normal features to enhance the accuracy.
- DC-BERT model is evaluated considering the customer review dataset of laptops and restaurants from SemEval 2014 task 4 considering precision, recall, accuracy, and macro-F1 score; also, comparative analysis is carried out considering accuracy and macro-F1 score.

The organization of this research work is such that the first section discusses the sentiment analysis background and the further section emphasizes the background of character and text recognition systems with the aspect level sentiment information with their feature extraction process along with the motivation and contribution to carry out this work. The second phase consists of the already existing methodologies along with their shortcomings and various techniques that have been applied. The third section focuses on the development of a model for the feature extraction and Network process. The fourth section contains the results that are obtained from this study. This ends with a conclusion stating the outcome of this research.

2. LITERATURE SURVEY

In recent years, several mechanisms have been introduced for aspect-based sentiment analysis (ABSA) tasks; in general, these methods are categorized into traditional machine learning approaches or neural network-based approaches. This section of the research work discusses the related work of aspect-based sentiment analysis and classification approach. The traditional approach of aspect-classification is mainly based on the feature engineering (FE) which indicates the hefty amount of time is used for gathering and analyzing the data, later features are designed based on the dataset characteristics, and further lexicons are constructed. In the case of a traditional approach, it is quite difficult to design the features through a manual process, and in case of change in dataset causes degradation in metrics performance, hence the neural network-based approach is used for feature capturing without feature engineering. The sentiment analysis is performed at an aspect level using the BERT model that is modified to predict the sentiment polarities [10]. Other than sentiment analysis polarity, extra contextual information is also provided by this model. The methodologies applied to sentiment analysis are discussed along with sarcasm analysis [11]. Various aspect level sentiment analysis, dialogue generation along bias in the system of sentiment analysis is discussed. A convolutional neural network (CNN) model is combined along with a bidirectional long short-term memory (Bi-LSTM) model for the analysis of sentiment information from predefined structured datasets. The model proposed in this paper focuses on aspect-level sentiment information [12]. Focuses on the use of recurrent models for sentiment analysis since the use of word sequence for this analysis uses the information through sentiment labels [13]. A graphical neural network is used for sentiment classification for dependency information that is syntactic [14]. The textual information is represented in the form of a graphical tree, the similarities that are present textually are plotted in a dependency graphically network. Segmentation of the text is first performed which a basic task of natural language is processing [15]. The segmentation is performed based on the document that is used as input for the model. The segmentation can take place from a sentence to sequence or from document to document to sentence. After this, a recurrent neural network is applied for sentiment analysis is applied to the segmented text on a sentence level. The sentiment analysis performed in this uses a convolutional neural network based on lexicons [16]. The retrieval of information is performed using sample sequence data of the system, this is termed a lexicon. An adaptive transfer network is used for sentiment analysis on the aspect level [17]. This proposed model focuses on the relationships among multiple domains. Sentiment analysis is performed based on a sentiment dictionary [18]. A sentiment dictionary is constructed that includes different categories of sentiment words. A Bayesian classifier is used to determine the field of the polysemic sentiment words. A convolutional neural network is combined with a bidirectional gated recurrent units (GRU) for sentiment analysis [19]. This combination of the two models is used to extract the sentiment features of the contexts. Previous works consider targeted aspects as auxiliary information or independent information, which not only misses the context information of aspects but also restricts the metrics performance like accuracy and macro-F1 score; hence, the research gap lies in obtaining the context information of targeted aspects. Thus, this research work introduced deep context information alongside with BERT model to enhance the model performance.

3. PROPOSED METHODOLOGY

Aspect-SC and aspect-SA are considered fine-grained NLP task and aims to predict the sentiment polarities with given targeted aspects in particular sentences; also, BERT model has proven to be one of the successful models for NLP-based tasks. BERT is a neural network-based mechanism for NLP; the BERT model has two steps i.e. pre-training and fine-tuning. In the pre-training approach, the model is trained over the pre-trained task on unlabeled data and in the case of fine-tuned; it is trained on labelled data with pre-trained parameters. However, BERT alone fails to achieve high accuracy sentiment polarity detection as it fails to understand context features in deep. Hence, this research work proposes fine-tuned BERT model along with a deep context feature layer known as DC-BERT to enhance the metrics. Figure 2 shows the DC-BERT model.

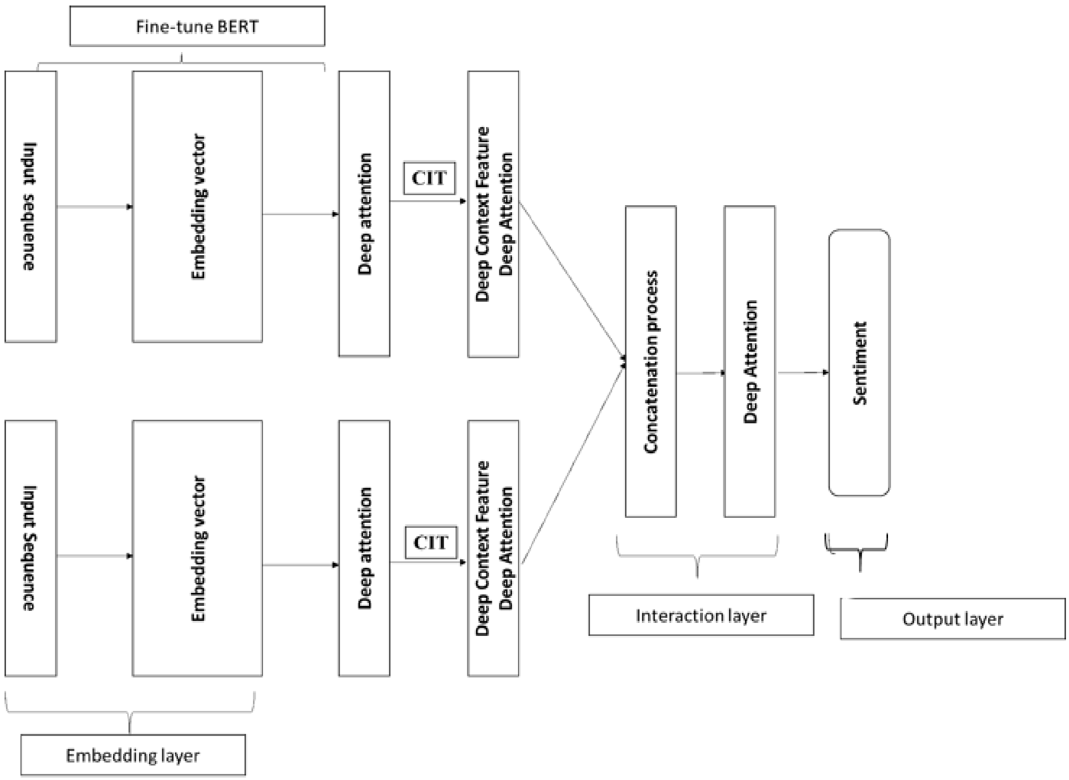

Figure 2. DC-BERT model

Figure 2 shows the implemented design of the DC-BERT model, it includes two distinctive embedding layers, these two layers are included for two customized feature extraction discussed; the first customized feature is related to deep focus on words and the other one is a general focus on words or sentence. At first glove, the model is adopted for embedding which tends to enhance the performance through a learning process. In another embedding layer, the feature extraction layer along with fine-tuned BERT layer is carried out; customized features along with introduced deep context layers are concatenated in the interaction layer to achieve high performance.

3.1. Task designing

Consider any input sequence $U = \{y_0, y_1, y_2 \dots \dots, y_o\}$ along with aspects; this sequence comprises targeted aspects and n words. To design the task for sentiment analysis based on aspect, target aspect sequences are generated as $U^v = \{y_0^u, y_1^u, y_1^u \dots y_o^t\}$. Designed task comprises $\mathbb{T}^u$ subsequence through a given sequence generated with n words.

3.2. Deep contextual relation

The existing approach of the input sequence is divided into the context and aspect to understand their interrelation; this research work develops a deep context that aims to capture efficient context. Deep context counts parameters between each contextual parameter towards particular aspects as parameter-aspect pairs. for instance, “while the movie was so good and entertaining that waiting was worth it”, for an aspect movie, DCR is computed as (1).

According to (1), R_c indicates contextual word position and k indicates the aspect central position and o indicates the total length of a sequence, $Deep\_context\_relation_k$ indicates the deep contextual relationships among the aspect and contextual parameters. The DC-BERT model aims to preserve the original aspect feature and deep context.

$$Deep\_context\_relation_k = |k - R_c| - \lfloor o(1/2) \rfloor \quad (1)$$

3.3. Word embedding layer

Word Embedding layer is considered as the fundamental layer of DC-BERT models, in here each word and taken are mapped to defined vector space through deep embedding layers. In this research, a pre-trained glove model is utilized for enhancing the learning process; let's consider any parameter $N \in T^{f_g X|\tau|}$ as the glove embedding where f_g is vector dimension and $|\tau|$ is the total size of the vocabulary. Considering these parameters each word is embedded into the vector. Further we have explained this.

3.3.1. Fine-tune BERT layer

Fine-tune BERT layer is a pre-trained model for language understanding, it is considered into deep embedding layer. To enhance the performance, the proposed adopts two independent BERT layers to the model with different context features; here two customized features are obtained to understand the context, these custom feature parameters are assigned as $\mathcal{Z}$ and $\mathcal{Y}$, these two parameters are considered and represented through (2). $Q_{FTB}^{\mathcal{Z}}$ and $Q_{FTB}^{\mathcal{Y}}$ $a$re output representation of custom feature context representation.

$$Q_{FTB}^{\mathcal{Z}} = Fine\_tune\_BERT^{\mathcal{Z}}(W^{\mathcal{Z}})$$

$$Q_{FTB}^{\mathcal{Y}} = Fine\_tune\_Bert^{\mathcal{Y}}(W^{\mathcal{Y}}) \quad (2)$$

3.4. Integrated layer

This layer comprises two sub-layers. The first layer is the deep attention layer as the DC-BERT model adopts attention-based deep learning to understand the deep context relation through attention score, and the second sublayer includes contextual information transformation. This is introduced to focus on laptop and restaurant dataset in particular. Later these two are integrated into the integrated layer.

3.4.1. Deep attention

Deep attention tends to perform several attention functions for attention score computation, identical attention function is computed for efficient computation. Let’s consider $\mathrm{Z_{DPA}}$ as input representation and NEF as a normalized exponential function then an identical attention function is given as (3). Let’s consider any parameter $\mathrm{S}, \mathrm{M}, X$ obtained by multiplying the hidden states of the upper layer through $N^s \in T^{f_j\,X\,f_s}$, $N^m \in T^{f_j\,/\,f_m}$ and $N^x \in T^{f_j\,X\,f_x}$, also these matrices are made trainable in the training process. Further, the attention-based dot product is computed through (4). In the case of the representation learned through each attention head is given through (5). According to (6), integration is updated; later activation function is deployed to enhance the learning capability.

$$(\mathrm{Z_{DPA}}) = \left(\mathrm{S}.\mathrm{M}^{\mathrm{V}} \left((f_m)\right)^{-1/2} \right).X \quad (3)$$

$$\mathrm{S}, \mathrm{M}, X = f_x(\mathrm{Z_{DPA}}) \quad (4)$$

$$f_x(\mathrm{Z_{DPA}}) = \begin{cases} \mathrm{S} = \mathrm{Z_{DPA}}.N^s \\ \mathrm{M} = \mathrm{Z_{DPA}}.N^m \\ X = \mathrm{Z_{DPA}}.N^x \end{cases} \quad (5)$$

$$deep\_attention(\mathrm{Z}) = HT\_func\left(\{J_0; J_1; \dots; J_j\}.Z^{deep\_attention}\right) \quad (6)$$

3.4.2. Contextual information transformation (CIT)

This is one of the modules introduced to improvise the performance on SemEval 2014 task 4 datasets. Input representation of this particular layer is considered as the output of multiple attention. Thus, contextual information transformation can be formulated through (7). According to (7) rectified linear activation function is used where N_1 and N_2 are probable trainable vectors of kernels; d_1 and d_2 are known bias, further output of deep feature layers is given as mentioned in (8). According to (8), $Q^{Z_\iota}_{Deep\_attention}$ and $Q^{y_\iota}_{Deep\_attention}$ are two different features with optimized output as Q_ι^Z and Q_ι^y in a respective manner.

$$CIT(Q_{DA}) = RAF(Q_{DA} * N_1 + d_1) * N_2 + d_2 \tag{7}$$

$$Q^{Z_\iota}_{Deep\_attention} = DA^Z(Q_\iota^{Z_\iota})$$

$$Q^{y_\iota}_{Deep\_attention} = Deep\_attention^y(Q_\iota^Z)$$

$$Q_\zeta^Z = Deep\_attention^Z(Q^{Z_\iota}_{Deep\_attention})$$

$$Q_\zeta^y = Deep\_attention^Z(Q^{Z_\iota}_{Deep\_attention}) \tag{8}$$

3.5. Optimal and deep context feature extractor

This research work deploys a feature extractor for learning the customized features. DC-BERT model approach takes two customized features. The first customized focus on the detailed context of the words, and the second customized feature focuses on the overall sentence. Further deep context relation phenomena are computed in case of each word considering the particular aspects.

3.5.1. Customized feature extractor layer

Figure 3 shows the customized feature extractor layer working. It comprises five steps; first step is accepting the output representation from the previous layer and taking input for this layer. Matrix is designed with a given input sequence and element-wise operation is performed to extract the dep-customized feature, at last, output representation is designed based on customized feature words.

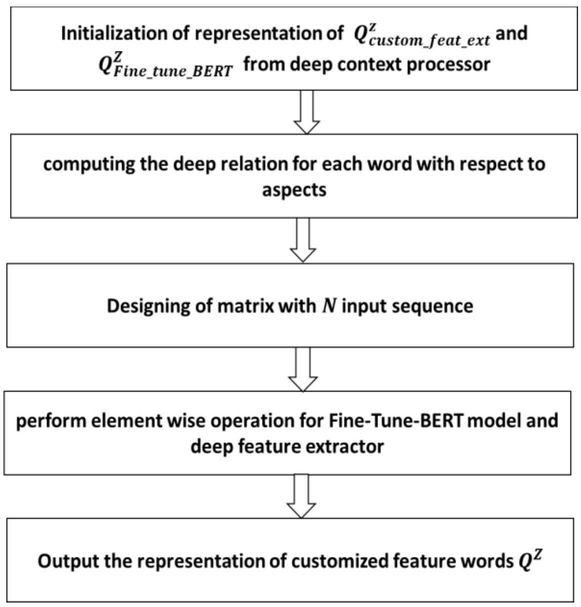

Figure 3. Customized feature extraction

3.5.2. Deep contextual features

Deep contextual features (DCF) layer tends to mask the fewer semantic words learned by fine-tuned BERT model, although it is easy to mask the less relative words in the input sequence, this layer discards such words, with the DCF layer, only the relevant words are masked and correlative among the aspect and less relevant words are stored at the output. At first, the deep feature is set to null vectors another deep attention is utilized to understand the context features, this design improvises the influence of less relevant context but stores the correlation among aspects and less relevant context.

$$X_k = \begin{cases} G & CIT_k \leq \text{threshold} \\ P & CIT_k \leq \text{threshold} \end{cases} \tag{10}$$

For instance, Q_ι^y is an output of a deep feature extractor, DCF focuses on the particular context through designing mask vectors X_0^o, thus matrices M as a mask are formulated as mentioned in (11). Also, with DCF, output representation is given as (12). Further, the output representation of contextual features is attained through the output of the DCF layer and computed as (13). Apart from deep features, the output representation of normal features is given as (14).

$$N = \left[X_0^o, X_1^o, \dots, X_p^o\right] \tag{11}$$

$$Q_{DCF}^z =. \left(Q_{deep\_attention}^z\right).(N) \tag{12}$$

$$Q^z = CIT\left(Q_{CIT}^z\right) \tag{13}$$

$$Q^y = CIT\left(Q_{DCF}^z\right) \tag{14}$$

3.6. Feature concatenation and projection

This layer is deployed for learning the normal features. At first, it concatenates the representation of normal features and deep features and projects them into Q_{encode}^{yz} and further deep attention is applied through encoding operation, moreover, the process has been formulated through (15). According to (15), the bias vector and weight vector are indicated through d^{yz} and N^{yz}.

$$Q^{yz} = [Q^z; Q^y]$$

$$Q_{encode}^{yz} = N^{yz}.Q^{yz} + c^{prim\_sec}$$

$$Q_{IL}^{yz} = deep\_attention\left(Q_{encode}^{yz}\right) \tag{15}$$

3.7. Output layer

In the case of the output layer, representation learned through the feature concatenation layer is pooled through hidden states extraction. This is given position as mentioned in (16). At last normalized exponential function is used for sentiment polarity prediction with d class number and A as sentiment polarity.

$$Q_{pooling}^{yz} = pooling\left(Q_{IL}^{yz}\right) \tag{16}$$

$$A = NEF\left(Q_{pooling}^{yz}\right) = \left(f^{Q_{pooling}^{yz}}\right)\left(\sum_{m=1}^{e} f^{Q_{pooling}^{yz}}\right) \tag{17}$$

3.8. DC-BERT training

Fine-tuned BERT model includes the BERT and glove model. Most of the steps are found to be identical except for the embedding approach and deep contextual layer. Fine-tuned approach utilizes the loss function with regularization, and the loss function is formulated as given in (18). With d as the class number and RP as the regularization parameter, PS indicates the parameter set of fine-tuned BERT model.

$$\mathbb{L} = regularized_{param} \sum_{\theta \in \text{set}_{\text{param}}} \theta^2 + \sum_1^e \hat{A}_k log_{10}(A_k) \tag{18}$$

4. PERFORMANCE EVALUATION

Sentiment analysis has drawn attention due to its broad application and the BERT model has proven to analyze the sentence in a bidirectional manner. This section of the research evaluates the proposed model, proposed model is designed through python as a programming language using spyder as IDE tools. The proposed model is evaluated on the system with Windows 10 platform packed with 8GB RAM and 2GB of compute unified device architecture (CUDA) enables NVIDIA graphics. To evaluate the model, accuracy and macro-F1 score is considered as evaluation parameter, also comparative analysis is carried out with the existing BERT model [20] to prove the model's efficiency.

4.1. Dataset details

Dataset plays a major role in the evaluation of any model, hence this research work considers publicly available datasets from Semeval 2014 task4 [21]. This task comprises two reviews dataset laptop and restaurant, also challenge dataset is categorized into three distinctive categories: positive, negative, and neutral. All three categories, it is divided into train dataset and test dataset. Further details of a dataset are given in Table 1.

Table 1. Dataset description

Dataset	Positive		Neutral		Negative	
Laptop	Train	Test	Train	Test	Train	Test
	994	341	870	128	464	169
Restaurant	2164	728	807	196	637	196

4.2. Metrics

To evaluate DC-BERT, four distinctive metrics precision, recall, accuracy, and the macro-F1 score are considered. These metrics are computed based on four-parameter true positive, true negative, false positive, and false negative from the confusion matrix. Considering the same confusion matrix, the below metrics are computed here.

i) Accuracy: Defined as the ratio of correctly classified sentiments from all the classified sentiments.
ii) Precision: Defined as the ratio of true positive towards the sum of true and false positive.
iii) Recall: Defined as the ratio of true positive towards the total sum of true positive and false negative.
iv) F1-score: Computed by observing the harmonic mean of recall and precision.

4.3. Comparison and comparison method

To prove the model's effectiveness 12 models were considered, and 11 models were referred from an existing model, these methods have been discussed. Several methods are being analyzed here to prove the accuracy. All the models are mentioned along with the technique used by them.

- Long sort term memory (LSTM) [22]: This model is based on the single way sequence model of RNN, it generates a hidden state for a single word and the last state is used for sentiment classification.
- Temporal dependence base long sort term memory (TD-LSTM) [23]: This technique uses LSTM on each side target word, also hidden states are presented on either side of the target and classified in the final representation.
- Attention based long sort term meory (ATAE-LSTM) [24]: It is an extended version of LSTM architecture, once aspect embedding and word embedding is passed through LSTM architecture, hidden states are mapped and the attention vector is computed after combining the hidden states. Final representation is utilized for classification.
- Memory network (MemNEt) [25]: This mechanism adopts a multi-hop attention-based mechanism, their main aim is to achieve contextual relevance and capture the sentiments.
- Interactive attention networks (IAN) [26]: This mechanism assumes that one aspect might have a different meaning; hence input embedding and aspect embedding are given into two different LSTM. Hidden states are averaged for achieving the interactive information of these two. Later, these two are separated and sent to softmax for classification.
- Recurrent attention network on memory (RAM) [27]: This mechanism utilizes the bidirectional LSTM to design memory based on input sequences through their relative positions. Later multiple attention is designed on this weighted memory. At last, a softmax layer is utilized for target emotion prediction.
- Aspect level sentiment classification with attention-over-attention neural networks (AOA-LSTM) [28]: It uses bidirectional LSTM for the conversion of aspect embedding and sentences in hidden states, later the AOA technique is adopted for merging hidden states to compute comprehensive weights. Final state representation is computed through hidden states sequence and weights.
- Multi-generator adversarial network (MGAN) [29]: This technique uses a fine-grained approach of attention mechanism along with multiple attention mechanisms to consider the aspects and sentence relevance.

- Deep mask memory network basessemantic dependency and context moment (DMNN-SDCM) [30]: This technique is mainly based on the memory network and introduces deep mask-MN (memory network) along with context moment, which provides the background knowledge of target aspects.
- Bidirectional encoder representations from transformers pre-training (BERT-PT) [31]: This technique utilizes machine-reading comprehension and introduces review reading comprehension (RRC); also, post-training approach is used to improvise the aspect knowledge.
- Attentional encoder network based bidirectional encoder representations from transformers (AEN-BERT) [32]: This model uses an attention mechanism for modelling targets and context on the trained BERT approach, it highlights the issue of regularization and label smoothing, and it aims to minimize the fuzzy label consistency.

4.3.1. Laptop dataset

This subsection evaluates the Fine-BERT model on the laptop dataset, Table 2 and Table 3 shows the accuracy and macro-F1 score comparison with the different mechanism. Moreover, through Table 1 it is observed that BERT is one of the successful models, and a variety of BERT models has been used with MGAR-ALBERT with 75.45%, MGAR-ALBERT with 77.98%, BERT-PT with 78.07% of accuracy, AEN-BERT with 79.93%, whereas fine-tune BERT model achieves 84.48% of accuracy in comparison with all these model. In Table 2, other baseline models like the LSTM-based model, DMNN-SDCM, RAM, and MGAN tries to implement deep learning concept but they achieve low accuracy.

Similarly, Table 3 presents the macro-F1 evaluation, in here MGAR-ALBERT no AC-AOA achieves 71.31%, BERT-PT achieves 75.08%, AEN-BERT achieves 76.31%, and MGAR-ALBERT achieves 75.85%. However, in comparison with all these models, fine-tune BERT achieves 83.05%; also, other baseline methods observe very low Macro-F1 score between 60 to 70%. Furthermore, several existing models only computed accuracy and macro-F1, which gives an overall idea of model performance. However, they ignored other metrics like precision and recall; the DC-BERT model achieves a precision of 82.88% and a recall of 83.82%.

Table 2. Accuracy on laptop

Model	Accuracy
LSTM	65.82
TD-LSTM	71.83
ATAE-LSTM	68.65
MemNet	70.33
IAN	72.10
RAM	75.01
MGAN	75.39
DMNN-SDCM	77.59
AOA-LSTM	74.5
BERT-PT	78.07
AEN-BERT	79.93
MGAR-ALBERT no AC-AOA	75.45
MGAR-ALBERT	77.98
Fine-tune BERT	84.48

Table 3. Macro-F1 on laptop

Model	Macro-F1
LSTM	64.02
TD-LSTM	68.43
ATAE-LSTM	62.45
MemNet	64.09
IAN	67.48
RAM	70.51
MGAN	72.47
DMNN-SDCM	73.61
BERT-PT	75.08
AEN-BERT	76.31
MGAR-ALBERT no AC-AOA	71.31
MGAR-ALBERT	75.85
Fine-Tune BERT	83.05

4.3.2. Restaurant dataset

The restaurant review dataset is another dataset from several 2014 tasks 4, this subsection presents the evaluation of fine-tune BERT model by comparing it with different existing models. Table 4 presents the accuracy evaluation and it is observed that BERT based on the model like MGAR-BERT no AC-AOA achieves 82.57%, AEN-BERT achieves an accuracy of 83.12%, BERT-PT achieves 84.95% and MGAR-ALBERT achieves 85.13%; other baseline methods remains less than 80% of accuracy.

Table 5 shows the macro-F1 score evaluation on the restaurant dataset; BERT model like MGAR-ALBERT no AC-AOA achieves a macro-F1 score of 72.23%, AEN-BERT achieves macro-F1 of 73.36%, BERT-PT achieves 76.96%, and existing model MGAR-ALBERT achieves a macro-F1 score of 77.68 %. In comparison with all these models, Fine-tune BERT achieves 90.02%. Other baseline methods remain between 60 to 70%. Similar to the laptop dataset, precision and recall metrics weren't found in these models, the DC-BERT model achieves a precision value of 91.09 and a recall value of 89.14%.

4.3.3. Comparative analysis and discussion

This section analyzes the improvisation over the existing BERT model, Figure 4 and Figure 5 show the classification metrics comparative analysis of both dataset laptop and restaurant. DC-BERT achieves 6.5% better accuracy than the existing model and a 7.2% better macro-F1 score than the existing model. Similarly, in Figure 4, the DC-BERT model achieves 7.73% improvisation in terms of accuracy and 12.34% of improvisation in terms of macro-F1 score.

Although accuracy and macro-F1 give a major idea about classification model performance, there are other metrics like precision and recall that have been ignored by various leading research works. Hence considering Table 2, Table 3, Table 4, and Table 5, it is clear that LSTM based model achieves satisfactory accuracy of 60 to 70 %, and other BERT adopted model achieves good accuracy of around eighty per cent. In comparison with all these models, the proposed DC-BERT model outperforms the other model.

Table 4. Accuracy on restaurant

Model	Accuracy
LSTM	74.61
TD-LSTM	78.00
ATAE-LSTM	77.23
MemNet	78.16
IAN	77.95
RAM	79.79
MGAN	81.25
DMNN-SDCM	81.16
AOA-LSTM	81.2
BERT-PT	84.95
AEN-BERT	83.12
MGAR-ALBERT no AC-AOA	82.57
MGAR-ALBERT	85.13
Fine-tune BERT	92.86

Table 5. Macro-F1 on restaurant

Methodologies	Macro-F1 (in percentage)
LSTM	63.56
TD-LSTM	66.73
ATAE-LSTM	64.95
MemNet	65.83
IAN	67.90
RAM	68.86
MGAN	71.94
DMNN-SDCM	71.50
BERT-PT	76.96
AEN-BERT	73.76
MGAR-ALBERT no AC-AOA	72.23
MGAR-ALBERT	77.68
Fine-tune BERT	90.02

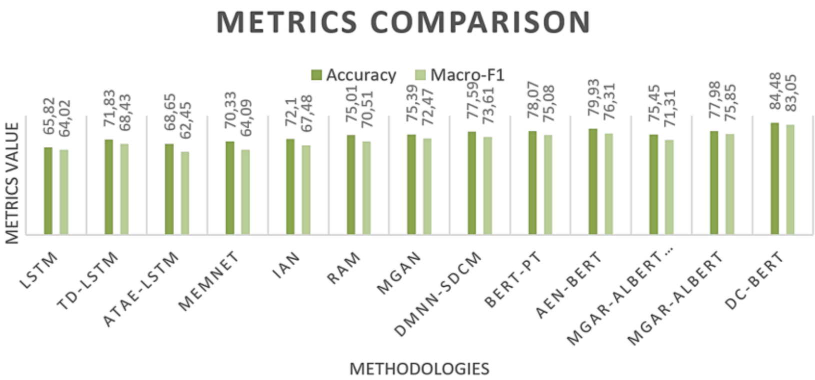

Figure 4. Metrics evaluation on laptop dataset

Figure 5. Metrics evaluation of restaurant dataset

5. CONCLUSION

Aspect-based-SA is considered a fine-grained task to analyze the user sentiment polarity towards particular aspects; it provides valuable knowledge for both consumers and businesses. BERT has been proven to perform well on several natural language processing (NLP) tasks including sentiment analysis and classification. This research work introduces, the DC-BERT model, which improves the BERT model through fine-tuned BERT layer, and a further deep context feature is introduced to enhance the model performance. DC-BERT model extracts the customized features for a deep and better understanding of context based on targeted aspects; these customized features are concatenated in interactive layers for output representation. DC-BERT model is optimized to enhance the metrics on the given dataset. DC-BERT model is evaluated on the review dataset of laptops and restaurants considering the accuracy and macro-F1 score metrics. Comparative analysis of the PC-BERT model with the existing BERT model along with other baseline methods shows the proposed model observes marginal improvisation in terms of accuracy and macro-F1 score. Hence DC-BERT model is proven to achieve the highest metrics in comparison with other models till this research is carried out which provides great scope for future sentiment analysis research. DC-BERT model is fine-tuned model which improvises the metrics considering a particular dataset, however, is a real scenario sentence given can be twisted and most of it could be related to sarcastic comments. Hence, future directions of our work would concentrate on considering more datasets including sarcastic comments.

REFERENCES

[1] Y. Wang, G. Huang, J. Li, H. Li, Y. Zhou, and H. Jiang, "Refined global word embeddings based on sentiment concept for sentiment analysis," *IEEE Access*, vol. 9, pp. 37075–37085, 2021, doi: 10.1109/ACCESS.2021.3062654.

[2] G. Zhai, Y. Yang, H. Wang, and S. Du, "Multi-attention fusion modeling for sentiment analysis of educational big data," *Big Data Min. Anal.*, vol. 3, no. 4, pp. 311–319, Dec. 2020, doi: 10.26599/BDMA.2020.9020024.

[3] W. Ali, Y. Yang, X. Qiu, Y. Ke, and Y. Wang, "Aspect-level sentiment analysis based on bidirectional-GRU in SIoT," *IEEE Access*, vol. 9, pp. 69938–69950, 2021, doi: 10.1109/ACCESS.2021.3078114.

[4] A. Nazir, Y. Rao, L. Wu, and L. Sun, "Issues and challenges of aspect-based sentiment analysis: a comprehensive survey," *IEEE Trans. Affect. Comput.*, vol. 13, no. 2, pp. 845–863, Apr. 2022, doi: 10.1109/TAFFC.2020.2970399.

[5] K. C. Allen, A. Davis, and T. Krishnamurti, "Indirect identification of perinatal psychosocial risks from natural language," *IEEE Trans. Affect. Comput.*, vol. 14, no. 2, pp. 1506–1519, Apr. 2023, doi: 10.1109/TAFFC.2021.3079282.

[6] T. Wang, K. Lu, K. P. Chow, and Q. Zhu, "COVID-19 Sensing: negative sentiment analysis on social media in China via BERT model," *IEEE Access*, vol. 8, pp. 138162–138169, 2020, doi: 10.1109/ACCESS.2020.3012595.

[7] L. Canales, W. Daelemans, E. Boldrini, and P. Martinez-Barco, "EmoLabel: Semi-automatic methodology for emotion annotation of social media text," *IEEE Trans. Affect. Comput.*, vol. 13, no. 2, pp. 579–591, Apr. 2022, doi: 10.1109/TAFFC.2019.2927564.

[8] N. Zhao, H. Gao, X. Wen, and H. Li, "Combination of convolutional neural network and gated recurrent unit for aspect-based sentiment analysis," *IEEE Access*, vol. 9, pp. 15561–15569, 2021, doi: 10.1109/ACCESS.2021.3052937.

[9] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *Comput. Sci. Comput. Lang.*, vol. 1, 2018, doi: https://doi.org/10.48550/arXiv.1810.04805.

[10] X. Li *et al.*, "Enhancing BERT representation with context-aware embedding for aspect-based sentiment analysis," *IEEE Access*, vol. 8, pp. 46868–46876, 2020, doi: 10.1109/ACCESS.2020.2978511.

[11] S. Poria, D. Hazarika, N. Majumder, and R. Mihalcea, "Beneath the tip of the iceberg: current challenges and new directions in sentiment analysis research," *IEEE Trans. Affect. Comput.*, vol. 14, no. 1, pp. 108–132, Jan. 2023, doi: 10.1109/TAFFC.2020.3038167.

[12] J. Zhou, S. Jin, and X. Huang, "ADeCNN: An improved model for aspect-level sentiment analysis based on deformable CNN and attention," *IEEE Access*, vol. 8, pp. 132970–132979, 2020, doi: 10.1109/ACCESS.2020.3010802.

[13] C. R. Aydin and T. Gungor, "Combination of recursive and recurrent neural networks for aspect-based sentiment analysis using inter-aspect relations," *IEEE Access*, vol. 8, pp. 77820–77832, 2020, doi: 10.1109/ACCESS.2020.2990306.

[14] X. Bai, P. Liu, and Y. Zhang, "Investigating typed syntactic dependencies for targeted sentiment classification using graph attention neural network," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 29, pp. 503–514, 2021, doi: 10.1109/TASLP.2020.3042009.

[15] J. Li, B. Chiu, S. Shang, and L. Shao, "Neural text segmentation and its application to sentiment analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 2, pp. 828–842, Feb. 2022, doi: 10.1109/TKDE.2020.2983360.

[16] N. K. Thinh, C. H. Nga, Y.-S. Lee, M.-L. Wu, P.-C. Chang, and J.-C. Wang, "Sentiment analysis using residual learning with simplified CNN Extractor," in *2019 IEEE International Symposium on Multimedia (ISM)*, IEEE, Dec. 2019, pp. 335–3353. doi: 10.1109/ISM46123.2019.00075.

[17] K. Zhang *et al.*, "EATN: An efficient adaptive transfer network for aspect-level sentiment analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 1, pp. 377–389, 2021, doi: 10.1109/TKDE.2021.3075238.

[18] G. Xu, Z. Yu, H. Yao, F. Li, Y. Meng, and X. Wu, "Chinese text sentiment analysis based on extended sentiment dictionary," *IEEE Access*, vol. 7, pp. 43749–43762, 2019, doi: 10.1109/ACCESS.2019.2907772.

[19] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, "Sentiment analysis for E-Commerce product reviews in Chinese based on sentiment lexicon and deep learning," *IEEE Access*, vol. 8, pp. 23522–23530, 2020, doi: 10.1109/ACCESS.2020.2969854.

[20] Y. Chen, L. Kong, Y. Wang, and D. Kong, "Multi-grained attention representation with ALBERT for aspect-level sentiment classification," *IEEE Access*, vol. 9, pp. 106703–106713, 2021, doi: 10.1109/ACCESS.2021.3100299.

[21] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, "SemEval-2014 Task 4: aspect based sentiment analysis," in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2014, pp. 27–35. doi: 10.3115/v1/S14-2004.

[22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[23] D. Tang, B. Qi, X. Feng, and T. Liu, "Effective LSTMs for target-dependent sentiment classification," *Comput. Sci. Comput. Lang.*,

vol. 1, 2015, doi: https://doi.org/10.48550/arXiv.1512.01100.

[24] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based LSTM for aspect-level sentiment classification," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2016, pp. 606–615. doi: 10.18653/v1/D16-1058.

[25] D. Tang, B. Qin, and T. Liu, "Aspect level sentiment classification with deep memory network," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2016, pp. 214–224. doi: 10.18653/v1/D16-1021.

[26] D. Ma, S. Li, X. Zhang, and H. Wang, "Interactive attention networks for aspect-level sentiment classification," *Comput. Sci. Artif. Intell.*, vol. 1, 2017, doi: https://doi.org/10.48550/arXiv.1709.00893.

[27] P. Chen, Z. Sun, L. Bing, and W. Yang, "Recurrent attention network on memory for aspect sentiment analysis," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2017, pp. 452–461. doi: 10.18653/v1/D17-1047.

[28] B. Huang, Y. Ou, and K. M. Carley, "Aspect level sentiment classification with attention-over-attention neural networks," 2018, pp. 197–206. doi: 10.1007/978-3-319-93372-6_22.

[29] F. Fan, Y. Feng, and D. Zhao, "Multi-grained attention network for aspect-level sentiment classification," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2018, pp. 3433–3442. doi: 10.18653/v1/D18-1380.

[30] P. Lin, M. Yang, and J. Lai, "Deep mask memory network with semantic dependency and context moment for aspect level sentiment classification," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, California: International Joint Conferences on Artificial Intelligence Organization, Aug. 2019, pp. 5088–5094. doi: 10.24963/ijcai.2019/707.

[31] Y. Song, J. Wang, T. Jiang, Z. Liu, and Y. Rao, "Attentional encoder network for targeted sentiment classification," *Comput. Sci. Comput. Lang.*, vol. 1, pp. 93–103, 2019, doi: https://doi.org/10.48550/arXiv.1902.09314.

[32] H. Xu, B. Liu, L. Shu, and P. S. Yu, "BERT post-training for review reading comprehension and aspect-based sentiment analysis," *Comput. Sci. Comput. Lang.*, vol. 1, 2019, doi: https://doi.org/10.48550/arXiv.1904.02232.

BIOGRAPHIES OF AUTHORS

Abraham Rajan earned his Bachelor's of Engineering BE degree in CSE from VTU, Belagavi in 2016. He has obtained his master's degree in M.Tech. (CSE) from Reva University in 2018. And currently he is a research scholar at CHRIST (Deemed to be University) doing his Ph.D. in Computer Science and Engineering and also working as assistant professor in Nagarjuna College of Engineering and Technology. He has attended many workshops and induction programs conducted by various universities. His areas of interest are big data analytics and cloud computing. He can be contacted at email: abinrajan1994@gmail.com or abraham_ranjan@rediffmail.com

Dr. Manohara Manur is an associate professor in the Computer Science and Engineering Department at School of Engineering and Technology of CHRIST (Deemed to be University), Bangalore with an experience of 22 years in teaching. He is qualified in Bachelor and Master Degrees in Computer Science & Engineering, and his Ph.D. in Computer Science & Engineering in the area of Data Mining and Big Data. His areas of interest are data mining, computer vision, machine learning, artificial intelligence, internet of things, and image processing. He can be contacted at email: manohar.m@christuniversity.in.