International Journal of Intelligent Networks 2 (2021) 64–69

Contents lists available at ScienceDirect

International Journal of Intelligent Networks

journal homepage: www.keaipublishing.com/en/journals/international-journal-of-intelligent-networks

Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model

M.P. Geetha [a,*], D. Karthika Renuka [b]

[a] Department of Computer Science& Engineering, Sri Ramakrishna Institute of Technology, Tamil Nadu, India
[b] Department of Information Technology, PSG College of Technology, Tamil Nadu, India

ARTICLE INFO

Keywords:
Sentiment analysis
BERT
Natural language processing
Product reviews
Machine learning
Opinion mining

ABSTRACT

Nowadays, digital reviews and ratings of E-commerce platforms provide a better way for consumers to buy the products. E-commerce giants like Amazon, Flipkart, etc provide customers with a forum to share their experience and provide potential consumers with true evidence of the product's outcomes. To obtain useful insights from a broad collection of reviews, it is important to separate reviews into positive and negative feelings. In the proposed work, Sentiment Analysis is to be done on the consumer review data and categorize into positive and negative feelings. Naïve Bayes Classification, LSTM and Support Vector Machine (SVM) were employed for the classification of reviews from the various classification models. Many of the current SA techniques for these customer online product review text data have low accuracy and often takes longer time in the course of training. In this research work, BERT Base Uncased model which is a powerful Deep Learning Model is presented to elucidate the issue of Sentiment Analysis. The BERT model gave an improved performance with good prediction and high accuracy compared to the other methods of Machine Learning in the experimental evaluation.

1. Introduction

In today's E-Commerce environment, the success of a product directly depends on the customer such that if the customer likes a particular product then it will be a success. If not then certainly there is a need to improvise the product by making some changes in the product. Sentiments or emotions of the customer about the product can be retrieved by the reviews and ratings. Hence analysis of these sentiments is to be done to analyze the customer's emotions based on their reviews and ratings.

Most of the E-commerce platforms provide the consumers to submit their comments for the products or service. Sentiment Analysis is to be done on the customer's reviews and ratings to extract useful information like customer's opinion on products, description for negative reviews, etc. Customers also have the provision to provide the rating points for the products or service. Amazon.com provides a numerical rating scale ranging from 1 for the worst to 5 for the best [1]. Reviews with mismatched ratings are identified since every individual rating is required to compute the average ratings.

Sentiment analysis can be executed in various fields like consumer comments, movie reviews, and product ratings [2]. Sentiment analysis is being executed by enormous business ventures to analyze customers preferences and interest, perform market basket analysis, track trademark of the product along with its influence and to learn about customer experiences.

The prime objective of Sentiment Analysis is to identify, depending on the review data provided by the customer as they like or dislike it [3]. Sentiment Analysis is text centered analysis; nevertheless to discover the exact sentiment of the sentence, there are factual challenges. Sentiment Analysis comprises of the following stages: (i) aspect based, (ii) sentence based and (iii) document based. A computer program really find it as a challenging task in identifying the sentiment, since there are numerous aspects in predicting the sentiments, even though sentiment analysis aids in the process of identifying the sentiment hidden in the text reviews. For instance, it is much simple in analysing review text to make sense of affirmative (positive) sentiment than the negative sentiment. It also needs governing the intended 'aspect' of the statement that has named as the negative sentiment.

It is viable that the customer has given a negative review due to the reason like time delay in delivering the product, but the customer really liked the product in all aspects [4]. In this case, the negative review outweighs the positive review. And, now is the situation where Aspect Based Sentimental Analysis plays a vital role. Aspect Based Sentimental Analysis (ABSA) is an approach that examines the terms that are linked to the aspects, and associates the sentiments connected to each aspect.

* Corresponding author.
E-mail address: geetha.cse@srit.org (M.P. Geetha).

https://doi.org/10.1016/j.ijin.2021.06.005
Received 29 March 2021; Received in revised form 11 June 2021; Accepted 30 June 2021
Available online 16 July 2021

ABSA model needs aspect categories and its corresponding aspect terms to bring out the sentiment for each aspect from the text corpus.

The paper is organized as follows: Section (1) describes about the Introduction. Section (2), represents the related work with respect to the research. Section (3) describes about the methodology and explains about the BERT algorithm with implementation. Section (4) shows the evaluation and performance analysis of the algorithm. Conclusion and Future enhancements are represented in Section (5).

2. Related work

The content of the following documents is based on the ongoing works in the domain of sentimental analysis and Deep learning [5]. Decision Tree, Logistic Regression and many more of the traditional Machine Learning classifiers have been used in the past researches.

The dominance of user reviews can be seen by observing the use of online shopping platforms by the people on a daily basis. Reviews [6] are also key as they are the first thing a customer turns to when they want to purchase the product. Hence, making the review data more dynamic is of utmost importance. In this research, a supervised learning model that can segregate the negative and positive reviews was proposed. In this paper sentimental analysis were executed by using several feature selection techniques. With Naive Bayes as the classifier, after collecting Amazon dataset, stop words, digits, punctuations and special characters are removed by performing the preprocessing steps [7,23]. Naïve Bayes Classification confirmed that phrase level, aspect level feature selection technique provides comparatively good output than single word and multiword. The inability to obtain sufficient result is the main disadvantage of using Naive Bayes.

3. Methodology

Amolik, A et al. [8] described that Machine learning based supervised algorithms have been utilized for performing sentimental analysis techniques on the input dataset. The following represents the steps of sentiment analysis using machine learning techniques based classification.

- Selecting and loading of dataset
- Pre-processing of selected dataset
- Feature Extraction
- Classification based on evaluated semantics
- Performance Evaluation of classification

3.1. Loading the dataset

Amazon, being the most used, is one of the most noted, and considerable leading E-commerce sites [21]. This is because of the countless user reviews that can be observed. Amazon product dataset was issued by the researchers. The unlabeled data had to be labeled to be of use as a supervised learning model. The structure of the metadata to be uploaded is in JSON format. Using Python, input review text data file in JSON format is converted into CSV file.

Table 1
Metadata description.

Feature	Description
reviewerID	• ID of the reviewer
asin</monospace>	• Product ID
reviewerName	• name of the reviewer
reviewText	• Review text content
Overall	Product Rating ranges from 1 to 5
Summary	• Summary of the review
unixReviewTime	Time of the review (unix time)
reviewTime	• Time of the review (raw)

Metadata information and Sample data in JSON format is shown in Table 1 and Fig. 1. As for the input review text is concerned, two main categories were picked out from Amazon products Electronics reviews which comprise of about 24,352 of electronics are taken into consideration. To analyze the reviews and to categorize reviewText and Overall attributes are taken into consideration. Review text consists of the reviews and feedback of the consumers who were present all over the world. Then the input dataset is segregated in to testing and training dataset and fed in to the model to perform the classification and sentiment analysis.

3.2. Data pre-processing

3.2.1. Tokenization

Tokenization can be done by segregating series of input text data into tokens. Tokens comprises of individual words, keywords, digits and punctuation marks. During tokenization, punctuation and other special characters in the input data are dropped, since they are not paid attention to, as to amplify the accuracy of the analysis. The tokens represents the input data for each process in building the classification model [9,24].

3.3. Feature Extraction

3.3.1. Bag of words

Bag of Words (BoW) is the representation of unique words from the review text corpus which is popularly used in NLP and Information Retrieval. Feature sets have been extracted using bag of words. Adjectives, adverbs and nouns are identified using Parts of Speech (PoS) tagging to create Bag of Words. Then these are represented using vectors which is fed into the Machine Learning Model.

3.3.2. TF-IDF

TF-IDF is the basic building blocks for many machine learning algorithms [10]. TF-IDF stands for Term Frequency and Inverse Document Frequency are two closely inter related metrics for searching and identifying the relevancy of a given word to the document. Separate Term Frequency and Inverse Document Frequency values are associated with each word. TF*IDF represents the products of weightage of individual score. High TF*IDF score represents the rare occurrence. TF represents the Term Frequency, IDF represents the importance of the term throughout the entire document.

3.4. Sentiment classification

To perform Sentiment Analysis and classification [11] various Machine Learning Models using supervised learning are implemented. Machine Learning models are implemented to predict the sentiment of the reviews and feedback submitted by the customers who bought the products [12]. Python Sci-kit libraries and associated packages are utilized to implement the Machine Learning Model. The actual dataset is divided using random split method into training data and test data in the ratio of 75:25. Training data is used to train the Machine Learning model to accurately predict the sentiment whereas the testing dataset is used to test evaluate the performance of the Machine Learning Models.

3.4.1. Naïve Bayes Classifier

Naïve Bayes Classifier works based on the Bayes theorem. One of the advantage of using Naïve Bayes classifier learns to classify using small quantity of training data. Naïve Bayes Classification uses ratings and review text to train the classifier which helps to understand how the positive and negative review makes sense. It is an arithmetic classifier which correlates the input text data vectors into the output class labels. Each row in the training data set A is represented by a k –dimensional feature vector, $M = m_1, m_2, \ldots m_k$. Output class label consists of N class labels and is represented as $n_1, n_2, \ldots \ldots n_k$. Hence for every input feature vector M, the Naïve Bayes classifier has to predict the output class labels

"overall": 3.0, "verified": true, "reviewTime": "09 22, 2013", "reviewerID": "A1BB77SEBQT8VX", "asin": "B00007GDFV", "style": {"Color:": " Black"}, "reviewerName": "Darrow H Ankrum II", "reviewText": "mother - in - law wanted it as a present for her sister. she liked it and said it would work.", "summary": "bought as a present", "unixReviewTime": 1379808000

Fig. 1. Sample data in JSON format.

by evaluating the posterior probability M belongs to anyone of the output class label n_i.

$$P(n_i|M) > P(n_j|M) \tag{1}$$

where $i,j \in I[1 \quad m]$ and $i \neq j$

$$P(n_i|M) = P(n_i)\prod_{a=1}^{n} P\left(\frac{x_k}{k_i}\right) \tag{2}$$

3.4.2. Support Vector Machine

Feng.X et al. [13] utilized Support Vector Machine which is a non-probabilistic learning classifier. Two labels are being utilized by SVM like 0 and 1. Here positive reviews are represented by 1 and negative reviews are represented as 0. Output class label consists of 0 and 1. All the input feature words are represented as either 0 or 1 specifying negative and positive sentiment respectively. SVM classifier accepts the input feature vector and classifies the output class labels as positive or negative. Optimal hyperplane is utilized to classify the labelled training input data. A hyperplane that segregates the input data of one class from another is represented by the equation y-mx = b

$$\text{In vector form: } (-m \quad 1)(x_1 \quad x_2) = b \tag{3}$$

$$i.e\ wx = b \text{ (product of 2 vectors)} \tag{4}$$

In matrix form:

$$\begin{bmatrix} -m \\ 1 \end{bmatrix}^{*} [x_1 \quad x_2] = b \tag{5}$$

$$i.e\ w^T.x = b \ or\ w^T.x - b = 0 \tag{6}$$

$$w.x - b > 0 \ if\ class : 1 \tag{7}$$

$$w.x - b < 0 \ if\ class : \ 2 \tag{8}$$

Even though SVM works well for large dimensions of spaces, accuracy of classification is not better for large data set since the training time required is high.

3.4.3. LSTM

N.Shrestha et al. [14] Long Short Term Memory (LSTM) is an exceptional category of Recurrent Neural Network. LSTM has the capability to use long memory in the hidden layer as the input for activation function. Review Text Data is given as input to the LSTM and it classifies the input category as output. In between the softmax layer and the hidden layer is the Pooling layer which increases the accuracy [15,25]. LSTM had few drawbacks like they take significantly a longer time to train for the neural networks, not truly bidirectional since the model is learning from left to right and right to left separately and then concatenating the context. As a result of these shortcomings we are moving towards the transformer architec1′ture which addresses some of these concerns.

3.4.4. BERT

BERT (Bidirectional Encoder Representations from Transformers) is a transformer architecture model for performing Natural Language Processing, Question Answering and Sentiment Analysis. BERT is a pre trained model, developed by the Researchers in Google in 2018. BERT is unique and distinguished from other Machine Learning Models in the sense that it is deeply bidirectional, both from left to right and right to left text representation, pre trained from an enormous unlabeled text corpus which encompasses of Wikipedia and a book corpus. Existing Context free models usually generate a single word embedding for all the words in the vocabulary irrespective of the context like for the word "matches" will have single embedding in the sentences "Ring perfectly matches with the bracelet" and "Toady both the matches are interesting". R. Dong et al. [16] employs BERT, which is represented in deeply bidirectional considering the preceding and succeeding context. BERT is employed in two approaches. One is Masked Language Prediction, in this few words of the input text are masked and fed into the BERT model to predict the masked words. BERT model understands the context of the preceding and succeeding words which is unmasked to predict the masked word. Next approach is Next Sentence Prediction, BERT uses two sentences X and Y as inputs, classify whether sentence X is followed B or it's a haphazard sequence of words.

Z. Gao et al. [17] proposed that BERT Model can be implemented in two phases, the first phase includes pre-training in which the model recognizes the input text data and the context of the data, second phase employs fine-tuning in which the model absorbs and identifies the solution. Fine tuning of Pretrained BERT model can be done by adding one layer to obtain contemporary results. BERTBASE and BERTLARGE are the two categories of BERT Model. Four layers of feed forward size in both the case.

BERTBASE.
Number of Layers of the transformer blocks = 12.
Output Dimensions = 768.
Number of Multi headed Attention = 12.
Total Parameters = 110 M.
BERTLARGE.
Number of Layers of the transformer blocks = 24.
Output Dimensions = 1024.
Number of Multi headed Attention = 16.
Total Parameters =

Classification of text review data is done by installing the BERT Base uncased from Hugging face transformers library. Then the Amazon dataset is uploaded into the Google Colab after changing the runtime to GPU.

3.4.4.1. BERT model structure. Input Review text data format: As per the BERT model's requirement, input token sequence are to be modified. [CLS] (Classification token) and [SEP] (Separation token) should be the first token and last token for every input sequence. Overall input text sequence is classified based on the output embedding related to the [CLS] token.

Preprocessing: Before feeding the review text data as input to the model, the following preprocessing steps [18] are to be implemented.

a) Canonicalization: Initially, all the numerical, punctuation marks, special characters are ignored and other existing uppercase letters are converted to lowercase letters
b) Tokenization: A glossary of 30,522 words are being used by BERT-Base uncased transformer model. Tokenization is done by segregating the input review text data into the certain format of list of tokens as per the glossary. WordPiece tokenization is utilized to handle the terms that are not present in the glossary. WordPiece Tokenizer splits the review word text into sub word or root word by

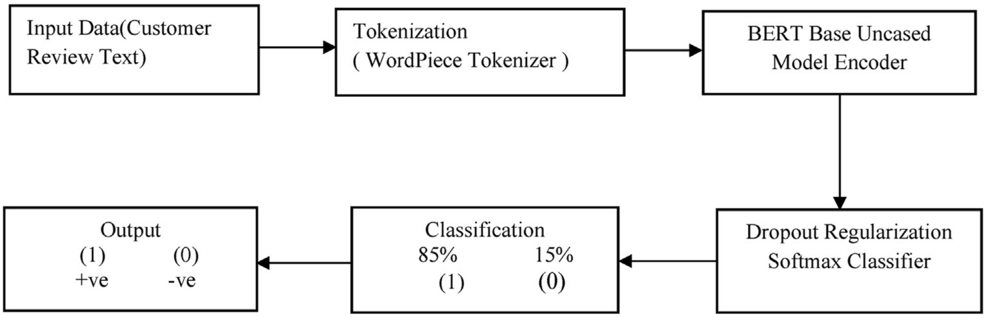

Fig. 2. Bert Base uncased model architecture..

removing the prefix or suffix like for example "looking" is represented as "look + ##ing". Tokenization is represented in Fig. 3.

- ip_ids' – represents the input sentences in the form of sequence of integers. '101' and '102' integer value represents the special tokens to be added and 0 represents the padding token.
- 'attention_mask' – represented in the form of 1's and 0's. The model is informed to have the attention over the masked tokens having the value'1′ and the mask value having '0' can be ignored.

BERT Layer: A simple pretrained BERT Base Uncased Model is built with Dropout regularization and a softmax classifier layer [19]. The architecture of the pretrained BERT-Base Uncased model is represented in Fig. 2. Out of the four main steps, preprocessing and sequence embedding stages are explained already. Third stage is applying Dropout regularization in order to prevent overfitting with a 0.1 factor of probability. Softmax is an activation function used in classification algorithm. Softmax is exponential and it maps the input scores into the output probabilities belonging to the output class labels with 1 as the probability sum. The Softmax function is represented by the equation,

$$S\left(\vec{x}_i\right) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \tag{9}$$

where $S(\vec{x}_i)$ is the softmax function, e^{x_i} represents the standard exponential function for input vector, n represents the number of output classes, e^{x_j} represents the standard exponential function for output vector.

The categorical cross entropy loss function calculates the loss of an example by computing the following sum:

$$Loss = -\sum_{i=1}^{n} x_i . log\hat{y}_i \tag{10}$$

where $\hat{y}_i$'' is the ith scalar value in the model output, xi is the corresponding target value, and 'n' is the scalar values numbers present in the model output.

BERT-Base-Uncased model is optimized using Adam optimizer and various learning rates are verified with 5 epochs. Categorical cross entropy is used to reduce the loss function. During the training phase, our BERT-Base-Uncased model parameters are evaluated. The best model is obtained when the validation loss value is low and this is obtained by tweaking the hyper parameters. Varying the hyper parameters shows a noteworthy impact on the performance of the model. Learning rate value of 2e-05 exhibits better output when compared to other learning rate values. Table 2 and Fig. 4 depicts the performance metrics values with varying learning rate.

Table 2
Performance metrics analysis with varying learning rate.

Max. Sequence Length	Learning Rate	Precision	Recall	F1 Measure	Accuracy
512	5e-05	83.22	85.98	88.72	87.69
	4e-05	82.66	86.02	89.01	87.36
	3e-05	83.17	85.55	88.94	88.15
	2e-05	83.66	86.22	89.41	88.48
	1e-05	83.38	86.12	89.12	88.31

```
x[43]: relatively inexpensive patch cable for electric guitar. i have had it for
a few months and so far it has held up pretty well
ip_ids[43]: [  101  2613 25499 23766  8983  5830  2005  3751  2858  1012  1045
2031
2018  2009  2005  1037  2261  2706  1998  2061  2521  2009  2038  2218
2039  3492  2092   102     0     0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
0     0     0     0     0     0     0     0     0     0     0     0
0     0     0     0     0     0     0     0]
att_masks[43]: [1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0
0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0]
```

Fig. 3. Tokenization.

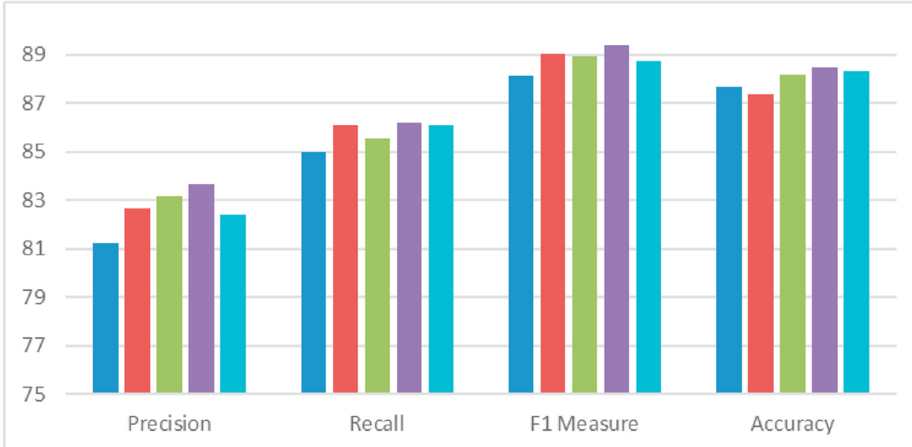

Fig. 4. Performance metrics analysis with varying learning rate.

Table 3
Comparative analysis of NB, SVM, LSTM, BERT.

Features & Model	Precision	Recall	F1-Measure	Accuracy
NB	79.32	73.57	69.32	80.12
SVM	82.68	84.31	81.26	81.33
LSTM	80.57	79.28	82.34	83.97
BERT	88.09	86.22	89.41	88.48

4. Evaluation

Various methods for extricating the typical features from the consumer reviews and a different point of view for assigning sentiment to review sentences have been expressed. Our speculation is that the automatic classification of customer specified text reviews into positive and negative reviews by enhancing the classification performance. This speculation is put to test on real-world review data for varying product categories using a number of different classifiers performance metrics like precision, recall, F1-Measure and Accuracy [20]. Performance of various machine learning algorithms along with BERT model is shown in Table 3, in which the results of performance metrics of BERT is comparatively higher than other Machine Learning algorithms. BERT yields higher accuracy of 88.48%. Fig. 5 shows the variation of performance among the different algorithms implemented.

4.1. Precision

Precision represents the correctness of the classification algorithm [22]. It is the ratio of number of correctly predicted positive reviews to the total number of reviews predicted as positive. It specifies how many returned reviews are positive. Increase in precision represents less false positive and vice versa

$$Precision(P) = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{11}$$

4.2. Recall

Recall represents the sensitivity of the classifier. It **is** the ratio of number of correctly predicted positive reviews to the actual number of positive reviews present in the corpus. Increase in recall represents the less false negative and vice versa.

$$Recall\ (R) = \frac{TP}{TP + FN} \tag{12}$$

4.3. F1-measure

It is the harmonic mean of precision and recall. F-measure can have best value as 1 and worst value as 0.

$$F1\ Measure = \frac{2PR}{P + R} \tag{13}$$

4.4. Accuracy

Accuracy is one of the most instinctive performance metrics and it is evaluated as the ratio of correctly predicted outcome to the total outcome.

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \tag{14}$$

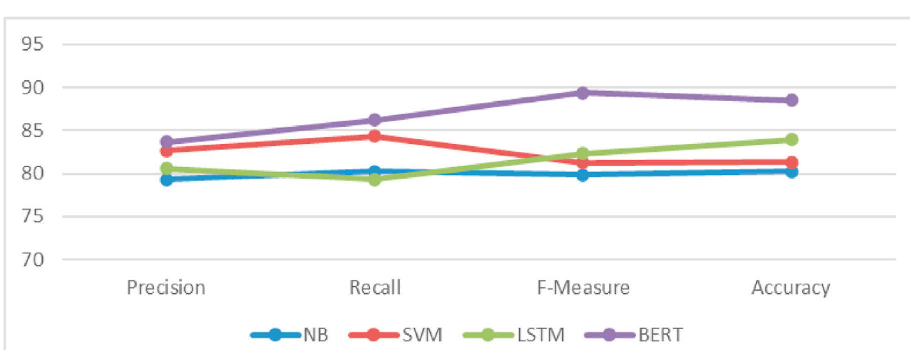

Fig. 5. Comparative analysis of NB, SVM, LSTM, BERT

5. Conclusion

Sentiment Analysis is carried out by using four Machine Learning Algorithms. These Machine Learning Algorithms are used to predict the Sentiment provided by the customers while buying a product. BERT is undoubtedly an ultimate advancement in the field of Machine Learning like Question Answering, Text Classification, and Sentiment Analysis. Since BERT is a pre-trained and allows fast fine-tuning will probably provide solution for the extensive dimensions of applications in the near future. Overall it is visible that all the Machine Learning Classification Models yields good prediction. Depending on the Performance evaluation criteria and comparison, we can conclude that BERT out performs well in accurate prediction than that which is possible using traditional Feature extraction and machine learning models. Performance criteria like F1 Measure, Precision and Recall are evaluated by applying Machine learning Algorithms like Naïve Bayes, SVM, LSTM. Comparative Analysis of these algorithms are made with the BERT Model. Then finally the BERT model is evaluated and the comparative performance variation is analyzed by changing the values of the hyper parameters. Future enhancements can be made by extending the proposed work with Hybridization algorithm. Maximum sequence length of the BERT model is 512 and it can be extended by increasing the input token sequence length.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ongoing Research Projects

- UGC-Minor Research Project (MRP) titled "Multiple Classifier for Email Spam Classification in a Distributed Environment for reducing Network Traffic and improving the Network Performance" for an amount of Rs.3,90,000.
- AICTE –MODEROBS (Data Analytics Tool) Rs. 2,00,000/-
- "Design and Implementation of E-Learning System using Deep Learning Based on Audio-Video Speech Recognition for Hearing Impaired in Native Language" funded by DST - ICPS, for an amount of 70,00,000/-

References

[1] A. Amolik, N. Jivane, M. Bhandari, M. Venkatesan, Twitter sentiment analysis of movie reviews using machine learning techniques, IACSIT Int. J. Eng. Technol. 7 (No. 6) (2016) 1–7.
[2] A. Arthi, K. Aravindhan, Enhancing the performance analysis of LWA protocol key agreement in vehicular ad hoc network, in: IEEE Xplore, 06 June 2019, https://doi.org/10.1109/ICACCS.2019.8728434.
[3] B. Santhosh Kumar, M.P. Geetha, G. Padmapriya, M. Premkumar, An approach for improving the labelling in a text corpora using sentiment analysis, Adv. Math. Sci. J. 9 (No.10) (2020) 8165–8174.
[4] C. Hang, V. Mittal, M. Datar, Comparative experiments on sentiment classification for online product reviews, Proc. Nat. Conf. Artific, Intel. Singapore 2 (2006) 1265–1270.
[5] E. Cambria, D. Das, S. Bandyopadhyay, A. Feraco, A Practical Guide to Sentiment Analysis", Springer, Berlin, Germany, 2017.
[6] D. Gamal, M. Alfonse, E.-S.M. El-Horbaty, A.-B.M. Salem, Analysis of machine learning algorithms for opinion mining in different domains, Mach. Learn. Know. Extrac. 1 (No. 1) (2018) 224–234.
[7] G. Kavitha, K. Aravindhan, Data forfeiture forbidding for multihop wireless broadcast routing in VANET, in: International Journal of Mobile & Adhoc Network 3, November 2013, pp. 280–283. Issue 4.
[8] H. Tang, S. Tan, X. Cheng, A survey on sentiment detection of reviews, Expert Sys. Appl. 36 (No. 7) (2009) 760–773.
[9] J. Panthati, J. Bhaskar, T.K. Ranga, M.R. Challa, Sentiment Analysis of Product Reviews Using Deep Learning," 2018 International Conf. On Advances in Computing Communication and Informatics, India, Bangalore, 2018, pp. 2408–2414.
[10] Jianmo Ni, Jiacheng Li, Julian McAuley "Justifying Recommendations Using Distantly-Labeled Reviews and Fined-Grained Aspects", Empirical Methods in Natural Language Processing (EMNLP), 2019.
[11] K. Aravindhan, DrC. Suresh Gnana Dhas, Destination-aware context-based routing protocol with hybrid Soft computing cluster Algorithm for VANET, in: Soft Computing Springer with ISSN 1432-7643 vol. 23, April 2019, pp. 2499–2507, 8.
[12] M. Alkalbani, A.M. Ghamry, F.K. Hussain, O.K. Hussain, Predicting the sentiment of SaaS online reviews using supervised machine learning techniques, in: Proc. Of International Joint Conf. on Neural Networks vol. 2016, 2016, pp. 1547–1553. October, Vancouver, BC, Canada.
[13] M. Munikar, S. Shakya, A. Shrestha, Fine-grained sentiment classification using BERT, in: International Conf. On Artificial Intelligence for Transforming Buses and Society (AITB 2019), Kathmandu, Nepal, 2019, pp. 2–5.
[14] M. Rajeswari, M. Mahalakshmi, R. Nithyashree, G. Nalini, Sentiment analysis for predicting customer reviews using a hybrid approach, in: Proc. On Advanced Computing and Communication Technologies for High Performance Applications, 2020, pp. 200–205. Kerala, India.
[15] M.P. Geetha, D. Karthika Renuka, Research on recommendation systems using Deep learning models, in: International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878 (Online), vol. 8, November 2019, pp. 10544–10551, 4.
[16] N. Shrestha, F. Nasoz, Deep learning sentiment analysis of Amazon.Com Reviews and Ratings, Int. J. Soft Comput. Artif. Intel. Appl. 8 (No. 1) (2019), 01–15.
[17] R. Dong, M. Schaal, M.P. O'Mahony, B. Smyth, Topic Extraction from Online Reviews for Classification and Recommendation, IJCAI International Joint Conf. on Artificial Intelligence, Beijing, China, 2013, pp. 1310–1316.
[18] S. Yu, J. Su, D. Luo, Improving BERT-based text classification with auxiliary sentence and domain knowledge, IEEE Access 7 (2019) 176600–176612.
[19] P.K. Singh, M.S. Husain, Methodological study of opinion mining and sentiment analysis techniques, Int. J. Soft Comput. 5 (No. 1) (2014) 11–21.
[20] S Sumathi, M P Geetha, P Ganesh Kumar, K Pushpalatha, A S Shanthi, yntactic and Sementic Based Similarity Measurenent for Plagiarism Detection, in: Int. J. of Innovative Technology and Exploring Engineering 9, (No.2), 2019, pp. 155–159.
[21] T.U. Haque, N.N. Saber, F.M. Shah, Sentiment analysis on large scale Amazon product reviews, in: 2018 IEEE International Conf. On Innovative Research and Development, ICIRD 2018, Thailand, 2018, pp. 1–6.
[22] T. Willianto, A. Wibowo, Sentiment analysis on E-commerce product using machine learning and combination of TF-IDF and backward elimination, Int.l J. Recent Technol. Eng. 8 (No. 6) (2020) 2862–2867.
[23] Y. Wang, X.J. Wang, A new approach to feature selection in text classification, Proc. of 4th Int. Conf. Mach. Learn. Cyber. 6 (2005) 3814–3819.
[24] X. Feng, Y. Zeng, Neural collaborative embedding from reviews for recommendation, IEEE Access 7 (2019) 103263–103274.
[25] Z. Gao, A. Feng, X. Song, X. Wu, Target-dependent sentiment classification with BERT, IEEE Access 7 (2019) 154290–154299.

M.P.Geetha, is currently working as Assistant Professor in Department of Computer Science and Engineering at Sri Ramakrishna Institute of Technology, Coimbatore. Previously she worked as Assistant Professor in Department of Computer Science and Engineering at Sri Krishna College of Engineering and Technology and SNS College of Engineering, Coimbatore. She has more than 9 years of teaching experience. Her areas of interest include Data Mining, Big Data Analytics, Data Structures and Compiler Design. She has conducted workshops and Faculty Development Program in up-coming trends in Computer Science and Engineering. She has published papers in International Journal and National Conferences.

Dr. D. KarthikaRenuka working in Department of Information Technology since 2004. Area of specializations includes Data Mining, Evolutionary Algorithms, Soft Computing, Machine Learning and Deep Learning, Client Server Computing, Computer Networks, Information Security. Organized an International Conference on Innovations in Computing Techniques Jan 22–24, 2015 (ICICT-2015) and National Conference on "Information Processing and Remote Computing 27th and February 28, 2014 (NCIPRC 2014).Reviewer for Computers and Electrical Engineering, Elsevier, Wiley Book chapter, Springer Book Chapters on "Knowledge Computing and its Applications".