Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 2459–2467

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Sentiment Analysis of Customer Feedback and Reviews for Airline Services using Language Representation Model

AKSH PATEL, PARITA OZA, SMITA AGRAWAL

Department of Computer Science and Engineering Nirma University, Ahmedabad, India

Abstract

The competitive airline sector has grown at a breakneck pace in the last two decades. A useful source for collecting consumer feedback and performing various forms of analysis on it is proper data collection. This collection of data can be used for sentiment analysis. Sentiment analysis is a type of analysis that involves extracting sentiment to find attitudes and emotions associated with the text or data supplied. It's a classification approach in which machine learning techniques are used to identify positive and negative words or reviews in text-driven databases. Further to explain the reasons for negative comments, a word cloud and a bar graph are used. Sentiment analysis is used to analyze the Airline reviews dataset in this paper. To test the performance of sentiment analysis, many Machine Learning (ML) algorithms have been utilized, such as Naive Bayes, Support Vector Machine, and Decision Tree (DT), and each of these approaches has produced distinct results. The performance of Google's BERT algorithm has been evaluated to that of other machine learning algorithms in our research. Furthermore, this paper explores the Bert architecture, which has been pre-trained on two NLP tasks: Masked language modeling and Sentence prediction. The "Random Forest" is used as a baseline against which the results of the "BERT Model" are compared because its performance is the best among the machine learning models. In terms of performance criteria such as accuracy, precision, recall, and F1-score, it is discovered that BERT outperformed the other ML techniques.

Keywords: Sentiment Analysis; BERT; Classification; Review; Machine Learning;

1. Introduction

Customer feedback is extremely important to the firm since it allows to enhance the services and facilities it give to their clients. According to Md. Hasib in 2021, the airline industry's competitive market has grown at a breakneck pace over the last two decades [1]. This study can forecast the public's reaction to the firm and determine whether or not customers are satisfied with the price and service given. Sentiment analysis, as defined by Spraha Kumawat in 2021, is an analysis that involves extracting sentiment in order to uncover the attitudes and feelings associated with the input text or data [2]. People nowadays rely on reviews and personal recommendations. As a result, a significant number of favorable evaluations strongly influence both customers and the airline firm in this scenario. Therefore, people will naturally choose items with a more significant number of favorable evaluations. Similarly, items with unfavorable assessments will try to discourage buyers from using them, resulting in a loss for the corporation. Therefore, sentiment Analysis is critical for determining the customer's relationship with the airline. Many research has been done on

E-mail address: 18BCE010@nirmauni.ac.in, parita.prajapati@nirmauni.ac.in, smita.agrawal@nirmauni.ac.in

1877-0509
10.1016/j.procs.2023.01.221

sentiment analysis using various machine learning algorithms such as Naive Bayes, Support Vector, and others, and practically every approach has shown promising results. SVM outperformed Naive Bayes for certain analysts, but Random Forest outperformed all other classification methods for others.

This paper proposed a framework to perform sentiment analysis on airline databases. The database is collected from KAGGLE [3]. Various ML approaches [5] such as Logistic Regression, SVM, DT, Random Forest, and Adaboost are used. Their results are compared either. Our main contribution is to evaluate the performance of the BERT framework against existing ML approaches.

The rest of the paper is organized as follows: In section 2 we mentioned literature of the domain. The dataset used for this work is briefly discussed in section 3. Section 4 gives an insight into the BERT framework for the classification process. We proposed our approach in section 5. Section 6 discusses the result of experiments carried out. Finally, the paper ends with the conclusion in section 7.

1.1. Motivation and Research Contribution

Sentiment analysis is employed to determine whether or not it performs well in product reviews. On the basis of customer reviews, research is currently being conducted in the airline business too. However, the researcher cannot fully analyze the data provided by the clients because it may be incomplete or erroneous. BERT has an advantage over other models since it creates word embeddings that are dynamically changed by the words or phrases around them, as opposed to models like Word2Vec, GloVe, or TF-IDF. The encoder scans the entire sequence of words or phrases in a single pass rather than reading the input in order. BERT can read the context from left to right and vice versa, therefore we can claim that it is bidirectional. BERT has an advantage over other algorithms or methods because of this. Therefore, BERT can be useful in comprehending the reviews and also in analysing them, which in turn aids the airline business in performing well and improving.

In this study, a framework for doing sentiment analysis on airline databases was proposed. Database data is gathered from KAGGLE [3]. Different machine learning (ML) techniques [5] are utilised, including Logistic Regression, SVM, DT, Random Forest, and Adaboost. Their outcomes are contrasted. Our primary contribution is to assess how well the BERT architecture performs. The Bert architecture, which has been pre-trained on two NLP tasksMasked language modelling and Sentence predictionis also explored in this study. Because it performs the best among the machine learning models, the ”Random Forest” serves as a standard against which the outcomes of the ”BERT Model” are measured.

2. Related Work

The classification task is currently widely studied in a variety of fields, including cancer detection and diagnosis [24, 26, 27], mammogram classification [28] , agricultural and environmental biology science [25], education, graph theory, Geoscience data [22], live streaming applications [6, 7] and so on. In addition, the airline industry has developed, and research is now undertaken based on consumer input or feedback. The article [8, 9] was published on Amazon Product Reviews. A supervised learning model was used to polarise many unlabeled product data. Pre-processing included tokenization and the elimination of stopwords. TF-IDF and chi-square algorithms extracted features from a bag of words. Later, the accuracy, precision, recall, and F-measure of different machine learning approaches employed in this experiment were calculated [10]. A new Adaboost technique for sentiment analysis of US airline twitter data is proposed in [11]. The data for this study paper was gathered from Skytrax from 2014 to 2017. Unwanted data is removed by data preparation. The Adaboost method is compared to several machine learning methods. For statistical processes like correlation and regression, data mining is used. As a result, Adaboost did exceptionally well in the trial. Another study on sentiment analysis Using Language Models was presented in [2]. To analyze US airline sentiment, authors have used Bert, Roberta, and Electra. The technique extracts data from a vast collection of unsupervised data to cover textual qualities. For the models mentioned above, various learning curves were created. Bert was believed to be more accurate than previous models in terms of accuracy. Twitter Data Sentiment Classification System for US Airline Service Analysis is shown in [12]. This study is entirely based on machine learning approaches for text categorization that are commonly utilized. The authors employed a technique called the doc2vec model, which is a phrase embedding, and then translated each sentence into a vector. Another research study based on a review analy-

sis where machine learning techniques for text categorization are utilized regularly is A Novel Deep Learning-based Sentiment Analysis of Twitter Data for US Airline Service [1]. Authors employed CNN and DNN neural network models in this case and translated the review into metadata then used TF-IDF to apply the data to the DNN's four layers. Authors compared machine learning approaches and neural network models based on model correctness in this research study. A research study [13] based on opinion mining on the airline twitter dataset, with the suggested model, split into three phases: pre-processing, classification, and validation. First, stopwords, punctuation signals, and stems are removed during pre-processing [14]. Following that, numerous machine learning approaches were applied in the categorization step. Next, the data was split 70 percent into training and 30 percent into testing for validation, and then 10-fold cross-validation was used. In one more study [15], consumers' comments were gathered in textual forms, such as thumbs up or thumbs down, emoji, and other symbols. This is then processed using pre-processing techniques as well as natural language processing. Finally, feature modeling and performance evaluation are carried out. Accuracy, precision, and recall are used to evaluate performance. The authors of [16] compared Naive Bayes and Support vectors for sentiment analysis on review datasets. The pre-processing is done first, and then the machine learning algorithms are implemented. Next, the probability of various characteristics is utilized to assess the dataset in Naive Bayes. The support vector is then employed, and the results are compared. Consequently, as compared to Nave Bayes, the support vector was more efficient. In order to assess customer sentiment forecast using online data, the study [23] tested six classification systems on the Bangladesh Airline Review dataset. For prediction, three deep learning algorithms (CNN, LSTM, and BERT) and three machine learning methods (Decision Tree, Random Forest, and XGBoost) were applied by the authors.

3. About Dataset

Airline reviews, gathered from Kaggle [3], were utilized for this study. It displays information from passengers or travelers who have expressed their opinions or feelings on the airlines they have traveled. The data includes categories such as airline feelings (good, negative, or neutral), a list of airlines on which passengers have flown, passenger/traveler information, location, time of comment creation, time zone, and the passenger's remark or text [4]. This dataset comprises 63 percent of data that will be negatively assessed, 21 percent of data that will be neutrally examined, and the remaining 16 percent that will be considered "others" (It's difficult to determine if the dataset is good, negative, or neutral because it hasn't been processed).

4. Bert Architecture For Text Classification

BERT stands for "Bidirectional Encoder Representations from Transformers" and is used to get a pre-trained bidirectional representation of text input by combining left and right context conditioning. Because it has been pre-trained, the model can be fine-tuned more quickly. During the training phase, it may learn from both the left and right sides of the tokens, known as bidirectional learning, which is helpful when learning about the many meanings of the same term. In the example below (see figure 1, you can see the several connotations of the word 'ring.'

BERT architecture is built upon transformers. It has two types. BERT base and BERT large. Both vary in layers of transformer, parameters and attention heads as shown bellow:

- BERT base
  - 12 layers of transformers
  - 12 attention heads
  - 110 million parameters
- BERST large
  - 24 layers of transformers
  - 16 attention heads
  - 340 million parameters

BERT's transformers serve as encoder blocks. The only difference between BERT large and BERT base models is their accuracy. BERT large provides a little higher level of precision than BERT base. However, training data with

Fig. 1. Bidirectional Context Understanding

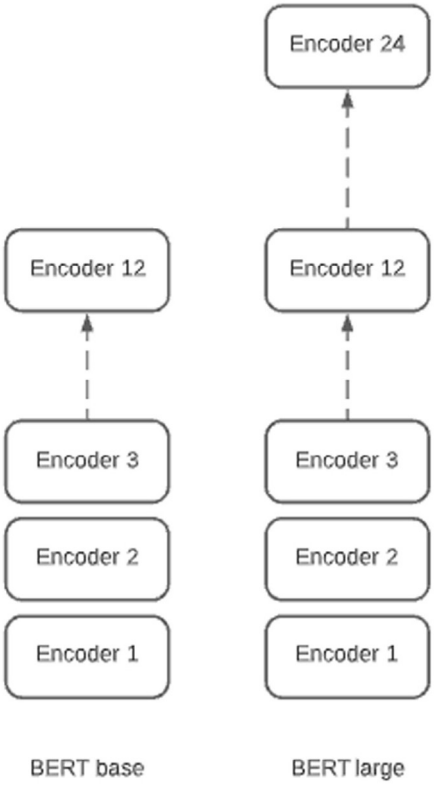

Fig. 2. BERT Architecture

BERT large take longer than training data with BERT base. The data for Sentiment Analysis in the aviation sector is trained using the BERT base architecture. Figure 2 depicts basic architectures of both the variants of BERT. To utilize the context in transformer, there are a few prerequisites that must be met that are enumerated below:

1. Tokens: [SEP] and [CLS] are two special tokens used. A [SEP] token is embedded at the end of each sentence to distinguish between two sentences, and a [CLS] token is inserted at the beginning of each sentence.

2. Text Classification: Although the sentences in the dataset may vary in length, BERT only accepts fixed-length sentences to maintain the consistency of the BERT vectors. It is possible to set a maximum length of 512 tokens. To change the length of the vector, pad, or truncate the tokens. A special token called [PAD] is used for padding.
3. Mask: Finally, there's the idea of an attention mask. It's a list of 1s and 0s that indicate which tokens are padded and which aren't. 0 indicates that it is a padded token [PAD] that does not need to be considered during training.

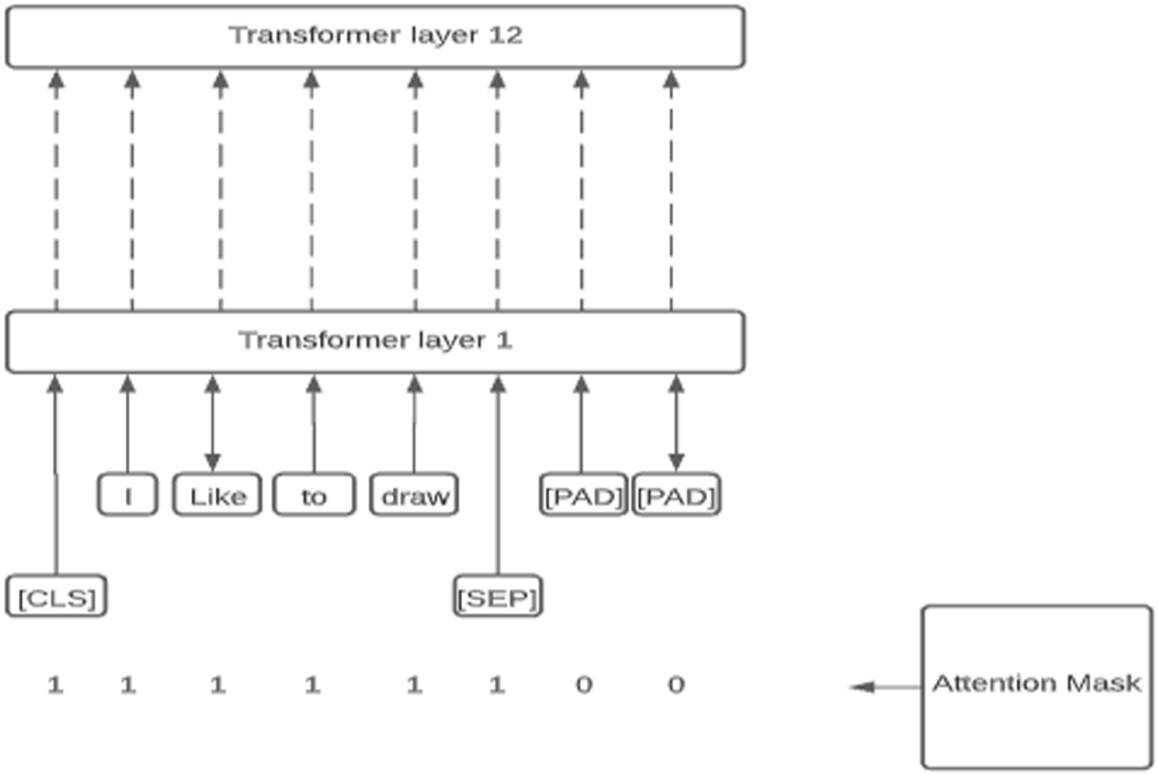

Fig. 3. Padding and Attention Mask

The process of padding and attention mask is pictorially presented in figure 3. Finally, BERT has been pre-trained on two NLP tasks as per the following two steps:

1. Masked Language Modeling: In most cases, models are trained by predicting the following word. Instead of guessing the following word, a bidirectional model can obtain a better result by masking some words from the provided phrase and then attempting to predict that masked word. E.g., Replace "towards" with [MASK] in the statement "I love to read blogs on [MASK] data science" and train the model to predict the word "towards" for the sentence "I love to read blogs on [MASK] data science." Researchers often replace 15% of the word with [MASK] to avoid attention to specific spots. Because the [MASK] token was never used in the fine-tuning process [20], it may be replaced with a random word or left alone.
2. Sentence Prediction: This task aims to find a connection between two sentences. Whatever dataset is used, 50% of the data will be train data, which implies the second sentence will be the next sentence of the first sentence. The next sentence in the left-over dataset will be chosen randomly from the corpus which are labeled as 'IsNext' or 'NotNext.'

5. Proposed Approach

In this paper employed Google's BERT text classification method and other machine learning approaches to train the model for airline review sentiment analysis[17, 18, 19]. BERT is pre-trained on a huge dataset. As a result, less data is required to train it for certain jobs. Using the graph and word cloud, we discovered the explanation behind the unfavorable or negative remarks.

Using the graph and the word cloud, the fundamental reason for the negative feedback supplied by consumers was discovered in this investigation. On the x-axis of figure 4 (left), several terms that are most common among the evaluations or comments offered by customers or passengers can be observed. The y-axis is the same; it displays the frequency of the words on the x-axis.

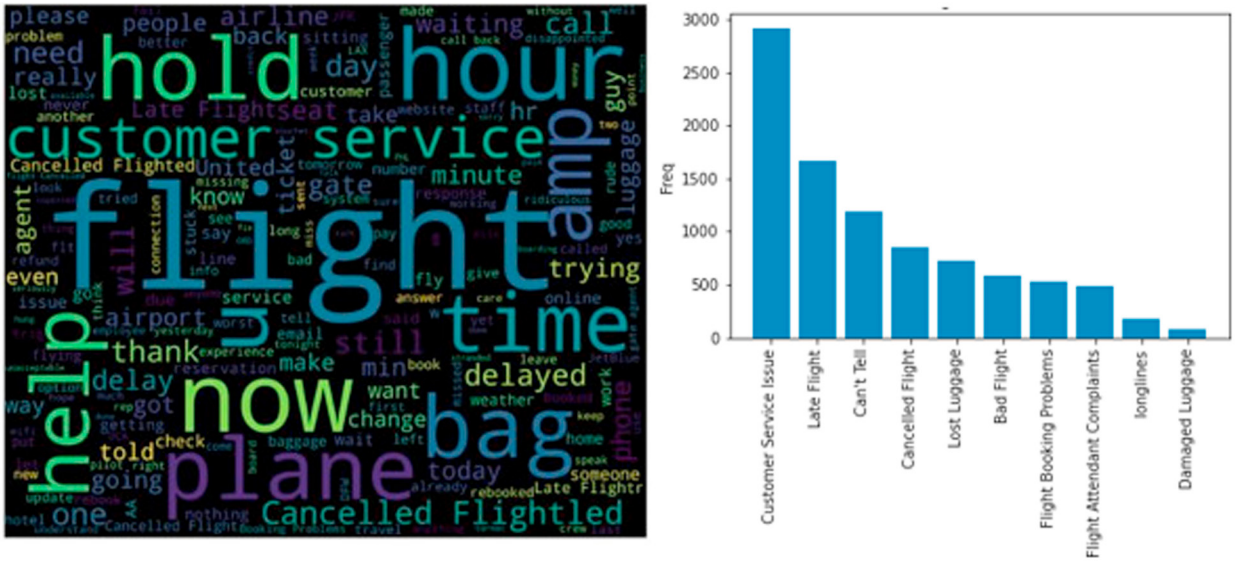

Fig. 4. Left-Word Cloud, Right-Graph showing frequency vs reasons for negative comments

The word cloud in figure 4 (right) depicts a group of terms that often appear in unfavorable or negative remarks. Customer service, baggage, delayed, canceled flights, and many other terms may be found in the word cloud. Numerous stopwords are deleted to create a clean text for a word cloud. As a result, it can be concluded that these are the numerous unfavorable causes for the passengers' bad comments or reviews. Several text data pre-processing were performed, such as eliminating punctuation and stopwords and returning a cleaned text list. With the aid of TF-IDF, the bag of words model is applied. The TF-IDF may be used to identify the data's term frequency and the document's inverse and normalize the vectors to unit length. Various types of classifiers are employed, each employing its own classification algorithms. These are used to classify text and, as a result, for analysis. Decision tree, Adaboost, Support vector machine, K- Nearest Neighbor, and Random Forest are the classifiers that were used. The train and test split uses half of the data for training and the other for testing. Finally, the text is tokenized, and tokens are converted to ids using BERT Tokenizer. Before proceeding, some data modifications must be made, such as adding special tokens [CLS] and [SEP] at the beginning and end of each sentence, padding or truncating sentences for vector length uniformity, and assigning an attention mask to determine which token is to be considered and which is not. The dataset is randomly divided into a training set and a validation set based on the reviews. Giving the model all of the data at once will negatively influence computation. Here also used "Bert-base-uncased" with all lower-case letters, and it is a base model of both BERT base and large, respectively.

6. RESULTS AND DISCUSSION

Comparison of the performance of ML approaches and the BERT framework to perform sentiment analysis on the airline review dataset is done. The performance metrics used are, accuracy, precision, recall and F1 score. Table 1 shows the comparison amongst all the ML models as well as BERT framework based on chosen performance measures. The ”Random Forest” is used as a baseline against which the results of the ”BERT Model” are compared

Table 1. Performance Comparision of various models

Models	Accuracy	Precision			Recall			F-1 score	
		Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative
Logistic Regression	65%	0	0.65	0	0	1	0	0	0.78
KNN	67%	0.56	0.8	0.48	0.54	0.77	0.47	0.55	0.78
Support Vectors	65%	0	0.65	0	0	1	0	0	0.78
Decision Tree	67%	0.53	0.79	0.4	0.48	0.79	0.43	0.51	0.79
Random Forest	77%	0.78	0.78	0.63	0.5	0.95	0.37	0.61	0.86
Adaboost	72%	0.69	0.73	0.56	0.56	0.95	0.08	0.62	0.82
BERT	83%	0.78	0.85	0.79	0.77	0.96	0.46	0.78	0.9

because its performance is the best among the machine learning models. Figure 5 shows a comparison of all the models concerning model accuracy. In addition, the random forest and BERT architectures are compared(See table 2).

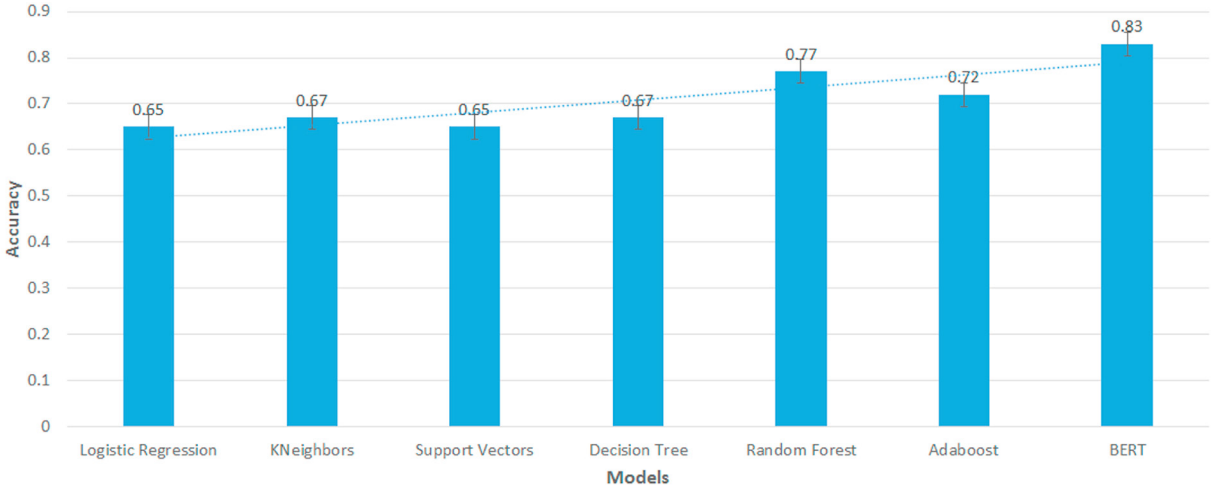

Fig. 5. Model performance comparison

A macro average is a calculation that calculates all of the potential metrics for a given class independently and averages the results. On the other hand, the weighted average is a machine learning strategy that combines all of the predictions generated thus far from several models. The accuracy score of the BERT model is 83%, which is much higher than the random forest model (77%). BERT outperforms Random Forest not just in terms of accuracy but also in terms of precision, recall, and even F1-score values. As a result, it may infer that the BERT architecture outperforms alternative machine learning algorithms for sentiment analysis in the application domain that was picked. This is due to several intrinsic benefits of BERT, such as its speed of growth, the fact that it requires less training

Table 2. Comparision of models based on macro and weighted average

Models	Macro Average			Weighted Average		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Random Forest	0.73	0.61	0.65	0.75	0.77	0.74
BERT	0.81	0.73	0.75	0.82	0.83	0.82

data and produces superior results. However, BERT also has several restrictions, such as the inability to function with language modeling, text generalization, and translation. We also assess how well our work stacks up against other techniques already in use. The results demonstrate that BERT performs better than models like LSTM, Roberta, and Electra (see table 3).

Table 3. Performance Comparison with existing methods

Models	Accuracy	Precision			Recall			F-1 score	
		Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative
Random Forest	77%	0.78	0.78	0.63	0.5	0.95	0.37	0.61	0.86
BERT	83%	0.78	0.85	0.79	0.77	0.96	0.46	0.78	0.9
[23] LSTM	76%	0.72	0.83	0.74	0.72	0.77	0.78	0.72	0.8
[2] Roberta	80.8%	-	-	-	-	-	-	-	-
[2] Electra	79.8%	-	-	-	-	-	-	-	-

7. CONCLUSION

BERT is a novel algorithm that Google presented recently. It requires extremely little data to train for specific tasks because it has been pre-trained on a huge dataset. It performs substantially better in categorization tasks. It's also utilized for the next sentence prediction test, which makes sense because people use search engines to get information, implying what the following words may be. It outperforms other algorithms in terms of total performance. BERT architecture is used to do sentiment analysis on a dataset of airline customer reviews. BERT was later on used to compare the outcomes of multiple machine learning algorithms for the same area. The random forest has been discovered to outperform all other machine learning algorithms. However, when compared to the BERT design, the random forest was proven to be unable to produce generalized outcomes. In our work, BERT produced the highest results in terms of accuracy, precision, recall, and F1-score.

The success of various BERT-based models aids the airline sector in implementing solutions to a variety of consumer issues. If BERT large is implemented instead of BERT base, the model performance can be improved and make it more accurate and exact, but at the same moment it is very time consuming during the training of data. However, it will take some investigation before it can be considered the most efficient algorithm. Many other BERT variants such as ALBERT, RoBERTa, ELECTRA and many more can be used for future development, but more knowledge and some investigation are required to implement further.

References

[1] Khan Md. Hasib, Md. Ahsan Habib, Nurul Akter Towhid, Md. Imran Hossain Showrov.(2021),*"A Novel Deep Learning based Sentiment Analysis of Twitter Data for US Airline Service"*,International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD), IEEE

[2] Spraha Kumawat, Inna Yadav, Nisha Pahal, Deepti Goel.(2021),*"Sentiment Analysis Using Language Models: A Study"*,International Conference on Cloud Computing, Data Science and Engineering (Confluence 2021) 11th International Conference on Cloud Computing, Data Science and Engineering,IEEE

[3] Twitter US Airline Sentiment,*https://www.kaggle.com/crowdflower/twitter-airline-sentiment*

[4] U. Naseem, S. K. Khan, I. Razzak and I. A. Hameed.(2019),"Hybrid words representation for airlines sentiment analysis",*In Australasian Joint Conference on Artificial Intelligence, pp. 381-392.*

[5] Pillai R., Oza P., Sharma P. (2020) , "Review of Machine Learning Techniques in Health Care." *In: Singh P., Kar A., Singh Y., Kolekar M., Tanwar S. (eds) Proceedings of ICRIC 2019. Lecture Notes in Electrical Engineering, vol 597. Springer, Cham.*

[6] Oza P.,Dhamasia A., Prajapati K.(2014), "Wearable live streaming gadget using Raspberry pi". *IJCSC vol. 7, no. 1, p. 67*

[7] Oza P., Sharma P.(2014) "Automation using Data Aggregation in Wireless Sensor Networks", *IJCSC, vol. 5, no. 1, p. 47*

[8] Levent Guner, Emilie Coyne, and Jim Smit.(2019),"Sentiment analysis for Amazon.com reviews", *KTH Royal Institute of Technology, Stockholm*

[9] Tanjim Ul Haque, Nudrat Nawal Saber, and Faisal Muhammad Shah.(2018), "Sentiment Analysis on Large Scale Amazon Product Reviews", *Ahsanullah University of Science and Technology, IEEE International Conference on Innovative Research and Development(ICIRD)*

[10] Tahura Shaikh, Dr. Deepa Deshpande.(2016),"Feature Selection Methods in Sentiment Analysis and Sentiment Classification of Amazon Product Reviews",*International Journal of Computer Trends and Technology (IJCTT), V36(4) pp. 225-230*

[11] E. Prabhakar, M. Santhosh, A. Hari Krishnan, T. Kumar, and R. Sudhakar.(2019),"Sentiment analysis of US airline twitter data using new Adaboost approach",*International Journal of Engineering Research and Technology (IJERT) V7, issue1, pp. 1-6*

[12] Ankita Rane, Anand Kumar.(2018),"Sentiment Classification System of Twitter Data for US Airline Service Analysis",*IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo: IEEE*

[13] Abdelrahman I. Saad.(2020),"Opinion Mining on US Airline Twitter Data Using Machine Learning Techniques",*16th International Computer Engineering Conference (ICENCO), Cairo: IEEE*

[14] F. Rustam, I. Ashraf, A. Mehmood, S. Ullah, and G. Choi.(2019),"Tweets Classification on the Base of Sentiments for US Airline Companies",*Entropy, vol. 21, no. 11, p. 1078*

[15] Raheesa Safrin, K.R.Sharmila, T.S.Shri Subangi, E.A.Vimal.(2017),"Sentiment Analysis on online PRODUCT review",*International Research Journal of Engineering and Technology (IRJET) Volume: 04 Issue*

[16] Abdul Mohaimin Rahat, Abdul Kahir , Abu Kaisar Mohammad.(2019),"Comparison of Naive Bayes and SVM Algorithm based on Sentiment Analysis Using Review Dataset",*8th International Conference on System Modeling and Advancement in Research Trends*

[17] Devlin, Jacob, et al.(2018) "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805*

[18] Mohammed, Athar Hussein, and Ali H. Ali.(2021) "Survey of BERT (Bidirectional Encoder Representation Transformer) types." *Journal of Physics: Conference Series. Vol. 1963. No. 1. IOP Publishing*

[19] Ashutosh Adhikari, Achyudh Ram, Raphael Tang, and Jimmy Lin.(2019), "DocBERT: BERT for Document Classification", *David R. Cheriton School of Computer Science University of Waterloo, arXiv:1904.08398v3 [cs.CL]*

[20] McCormick Chris, Ryan, Nickl and S. Levy, "Bert fine tuning tutorial with PyTorch", *Retrieved from: colab.research.google.com/drive/1pTuQhug6Dhl9XalKB0zUGf4FIdYFlpcX#scrollTo=EKOTlwcmxmej*

[21] C. Kariya and P. Khodke.(2020) ,"Twitter Sentiment Analysis",*2020 International Conference for Emerging Technology (INCET), Belgaum: IEEE*

[22] Shah N., Agrawal S., Oza P. (2021), "Data Ingestion and Analysis Framework for Geoscience Data." *In: Singh P.K., Singh Y., Kolekar M.H., Kar A.K., Chhabra J.K., Sen A. (eds) Recent Innovations in Computing. ICRIC 2020. Lecture Notes in Electrical Engineering, vol 701. Springer, Singapore.*

[23] Khan Md. Hasib, Nurul Akter Towhid, Md. Golam Rabiul Alam.(2021), "Online Review based Sentiment Classification on Bangladesh Airline Service using Supervised Learning", *5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), IEEE*

[24] P. Oza, P. Sharma, S. Patel, and P. Kumar, "Deep convolutional neural networks for computer-aided breast cancer diagnostic: a survey", *Neural Computing & Application 34, 18151836 (2022). https://doi.org/10.1007/s00521-021-06804-y*

[25] Thomas, J., Sharma, N. C., Kumar, P., Chauhan, A., & Chauhan, P. (2022), "Effect of biostimulant and biofertilizers on soil biochemical properties and plant growth of apple (Malus x domestica Borkh.) nursery", *Journal of Environmental Biology, 43(2), 276-283.*

[26] Oza, P., Sharma, P., Patel, S., Adedoyin, F., & Bruno, A. (2022), "Image Augmentation Techniques for Mammogram Analysis", *Journal of Imaging, 8(5), 141.*

[27] Oza, Parita, Paawan Sharma, and Samir Patel. "A Drive Through Computer-Aided Diagnosis of Breast Cancer: A Comprehensive Study of Clinical and Technical Aspects." *Recent Innovations in Computing (2022): 233-249.*

[28] Oza, Parita, Yash Shah, and Marsha Vegda, "A Comprehensive Study of Mammogram Classification Techniques", *Tracking and Preventing Diseases with Artificial Intelligence. Springer, Cham, 2022. 217-238.*