

Enhancing Sentiment Analysis with BERT: Leveraging Randomized Sentence Order and Stanford Parser Parsed Phrases from Rotten Tomatoes Dataset

SUBMITTED BY:	ADEEL HUSSAIN SHAH	ID: 21031305
---------------	--------------------	--------------

Introduction

In recent years, social media has assimilated into everyday life. People want to post about every single event in their lives on social media. Social media is now used to display pride or esteem by publishing images, text, videos, etc. Users express their thoughts on current events, politics, film reviews, and other subjects in the text, which is a critical component of information sharing. Due to their shortness, these thoughts are often referred to as Short Texts (ST) on social networking platforms [1]. Due to its ease of use and power to influence the public, ST has grown in prominence relative to conventional blogging. Even search engines employ them in the form of queries. Aside from their popularity, ST faces difficulties such as sarcasm detection, emotion, slang use, etc. Sentiment Analysis, sometimes known as SA, is the process of understanding brief messages and drawing insightful conclusions from them [2]. SA was significant in the 2016 US Presidential Elections [3]. On microblogs like Twitter and Facebook, people expressed their likes and dislikes towards a certain political party. Candidates edited their tweets based on these assessments after analyzing those blogs. SA thereby assisted them in gaining more fans and followers. The majority of companies use SA extensively since it can quickly evaluate several documents at once when doing it manually would take longer. Companies utilize SA to develop novel company tactics based on client input [4]. Reviews are short writings that often provide opinions on movies, books, or other things. These reviews are essential to the success of films or the sales of goods [5]. To learn more about the cast, crew, reviews, and ratings of movies, people often consult blogs and movie review sites like IMDb. Reviews thus play a significant part in bringing audiences to the theatres in addition to word-of-mouth advertising. To put it another way, SA on movie reviews facilitates Opinion Summarization [6] by extracting the reviewer's state of mind.

Dataset

The collection comprises tab-separated files with sentences taken from the Rotten Tomatoes dataset. The train/test split has been kept, but the sentence order has been randomly generated to guarantee benchmarking accuracy. The sentences have been parsed using the Stanford parser, producing a number of phrases. Each sentence has a Sentenced allocated to it, while these phrases have individual Phrase IDs. It's important to note that the dataset only contains one instance of recurrent terms, such as frequent short words [7].

Exploratory Investigation of Dataset

As part of our exploratory investigation of the Rotten Tomatoes dataset, we looked at phrases that the Stanford parser had processed. To learn more about the dataset's structural makeup and maybe spot recurrent language trends, the analysis will look at the distribution of phrases, Sentence IDs, and Phrase IDs. Moreover, Fig.1 Shows a sample of the dataset and their sentiment. The distribution of dataset lengths in connection to sentiment categories, especially the "negative" and "positive" labels, are shown visually in Figure 2. With the help of this visualization, it is possible to identify the variations in dataset length within these sentiment classes and identify potential patterns or differences in the phrase structure between negative and positive opinions. Fig.3. Represent visualizes the distribution of sentence lengths in the dataset, providing insights into the typical word count range for the sentences

	text	label
0	I grew up (b. 1965) watching and loving the Th...	0
1	When I put this movie in my DVD player, and sa...	0
2	Why do people who do not know what a particula...	0
3	Even though I have great interest in Biblical ...	0
4	Im a die hard Dads Army fan and nothing will e...	1

Fig.1 Shows the IMDB Movie sentiments.

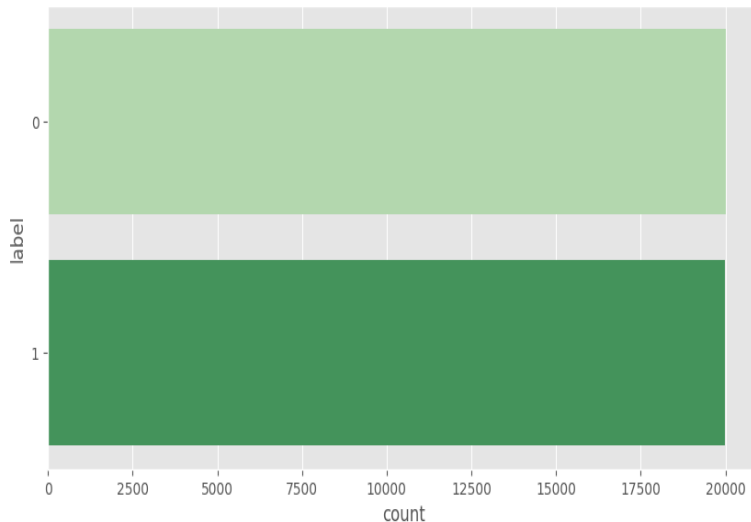


Fig.2 Shows the dataset length and sentiments

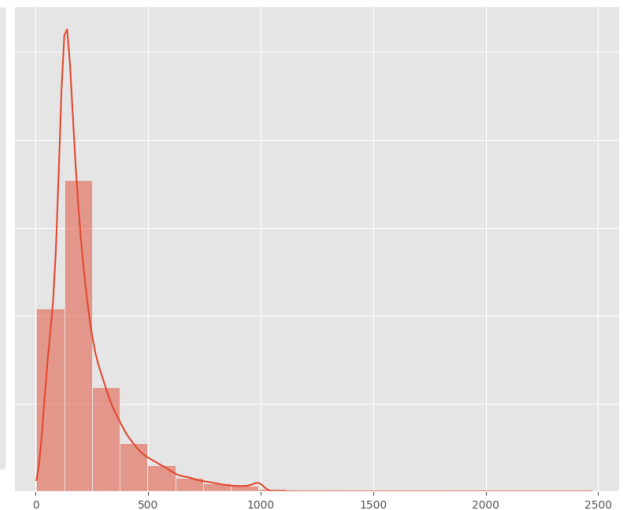


Fig.3 Shows Sentence Lengths in the dataset

Model Used in This Assignment BERT

In the area of natural language processing, the BERT (Bidirectional Encoder Representations from Transformers) model is a powerful pre-trained language representation model. By collecting contextual information from both the left and right contexts of each word in a phrase, revolutionized several NLP tasks and provided an in-depth understanding of language semantics. The transformer design used by BERT allows it to manage long-range dependencies and record complex word associations. The two tasks that make up its pre-training, masked language modeling and next sentence prediction, provide a model that is adaptable and powerful and can be modified for a variety of NLP tasks, delivering state-of-the-art performance across many benchmarks [8]. Fig.4 Shows the block diagram of the model.

Sparse Categorical Cross-Entropy:

In classification models where the objective labels are integers denoting classes, a mathematical measure known as sparse categorical cross-entropy is often used. In order to help the model successfully learn to assign the proper classes, it measures how different the projected probability distribution and the actual label distribution are from one another. In situations like image classification assignments, where classes are integer-encoded and mutually exclusive, this loss function is very helpful [9].

Adam Optimizer

Adam, which stands for "Adaptive Moment Estimation," is a well-known algorithm for optimization used in the training of machine learning models, particularly neural networks. The learning rates of specific parameters during training are adaptively adjusted by combining the concepts of momentum and the RMSProp optimizer. This promotes quicker convergence and better performance in a variety of optimization problems by overcoming difficulties like vanished or increasing gradients [10].

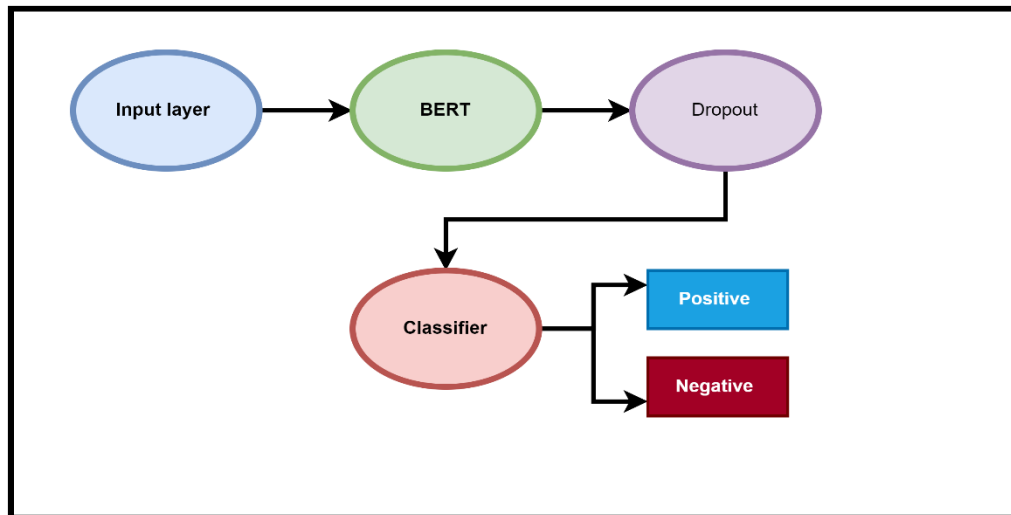


Fig. 4. Block Diagram of Model

Result of The BERT Model

The BERT model has a loss of 0.0931 and an accuracy of 0.9685 on the training data after three epochs of training on the Movie Review dataset. The model's validation findings revealed a validation accuracy of 0.9190 and a validation loss of 0.2159. Moreover, the results are shown in Table 1. These results show that the model did well on the training set, obtaining a high accuracy and a reasonably low loss. Fig. 5 shows training loss and validation loss. The model's performance generalizes to unidentified data quite well, according to the validation findings, while there is a slight decrease in accuracy and an increase in loss when compared to the training set. Fig. 6 shows the confusion matrix of our model.



Fig. 5 Shows the training and validation loss of model

Table 1. Shown Results of The BERT Model

loss:	accuracy:	val_loss:	val_accuracy:
0.0931	0.9685	0.2159	0.9190

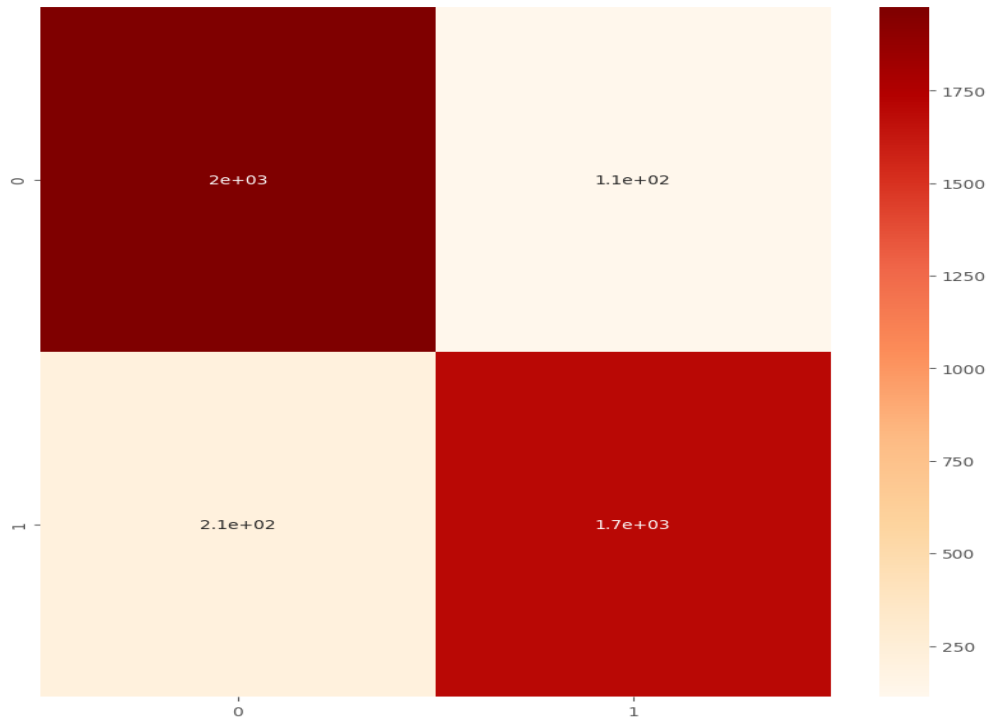


Fig.6. Represent the Confusion Matrix

Results prediction from text

BERT-based model to perform text classification. The model analyses the provided text and concludes that it has a very high likelihood of being classified as "Positive" sentiment (with a score of roughly 0.996) and a very low probability of being classified as "Negative" sentiment (with a score of roughly 0.004). This implies that the model perceives the text as being primarily helpful in sentiments. The results are shown in Fig.7.



Google Colab:

<https://colab.research.google.com/drive/1Spss4OyXVw0oQU4SXjylC9zMCfMgn80B?usp=sharing#scrollTo=2j0E4dOZseGE>

GitHub Link: <https://github.com/AdeelHusaain/BERT-Assignmet>

Reference:

- [1] Dos Santos, C., & Gatti, M. (2014, August). Deep convolutional neural networks for sentiment analysis of short texts. In *Proceedings of COLING 2014, the 25th international conference on computational linguistics: technical papers* (pp. 69-78).
- [2] Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in human behavior*, 31, 527-541.
- [3] Ahmad, R., Pervaiz, A., Mannan, H., & Zaffar, F. (2017). Aspect Based Sentiment Analysis for Large Documents with Applications to US Presidential Elections 2016. In *AMCIS*.
- [4] Xu, K., Liao, S. S., Li, J., & Song, Y. (2011). Mining comparative opinions from customer reviews for competitive intelligence. *Decision support systems*, 50(4), 743-754.
- [5] Tripathy, A., Agrawal, A., & Rath, S. K. (2016). Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications*, 57, 117-126.
- [6] Moussa, M. E., Mohamed, E. H., & Haggag, M. H. (2018). A survey on opinion summarization techniques for social media. *Future Computing and Informatics Journal*, 3(1), 82-109.
- [7] Bouaziz, D. E., & Khantoul, B. (2020). Sentiment analysis with deep learning.
- [8] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- [9] Kakarla, J., Isunuri, B. V., Doppalapudi, K. S., & Bylapudi, K. S. R. (2021). Three-class classification of brain magnetic resonance images using average-pooling convolutional neural network. *International Journal of Imaging Systems and Technology*, 31(3), 1731-1740.
- [10] Bock, S., Goppold, J., & Weiß, M. (2018). An improvement of the convergence proof of the ADAM-Optimizer. *arXiv preprint arXiv:1804.10587*.