Review Article

BERT: A Review of Applications in Sentiment Analysis

Md Shohel Sayeed [1]*⬤, Varsha Mohan [1], Kalaiarasi Sonai Muthu [1]

[1] *Faculty of Information Science and Technology, Multimedia University, 75450 Melaka, Malaysia.*

Abstract

E-commerce reviews are becoming more valued by both customers and companies. The high demand for sentiment analysis is driven by businesses relying on it as a crucial tool to improve product quality and make informed decisions in a fiercely competitive business environment. The purpose of this review paper is to explore and evaluate the applications of the BERT model, a Natural Language Processing (NLP) technique, in sentiment analysis across various fields. The model has been utilized in certain studies for various languages, restaurant businesses, agriculture, Automated Essay Scoring (AES), Twitter, and Google Play. The BERT model's fine-tuning steps involve using pre-trained BERT to perform various language understanding tasks. Text pre-processing is conducted to clean up the data and convert it to numbers before feeding it into BERT, which generates vectors for each input token. We found that BERT outperformed the norm on a range of general language understanding tasks, including sentiment analysis, paraphrase recognition, question-answering, and linguistic acceptability. The detection of neutral reviews and the presence of false reviews in the dataset are two problems that have an impact on the model's accuracy. Training is also slow because it is huge and there are many weights to update. Additional research could be conducted to improve the BERT model's accuracy by constructing a false review categorization model and providing more training to the model in recognizing neutral reviews.

Keywords: Natural Language Processing; BERT; Fine-Tuning; Machine learning; Sentiment Analysis.

1. Introduction

The word "e-commerce" refers to the exchange of goods and services over the internet. It offers a variety of tools, guidelines, plus resources for both buyers and sellers, including cash on delivery, mobile shopping alternatives, and encryption for online payments [1]. Consumers and businesses alike are valuing reviews more and more. Consumers may use internet reviews to assist their decision about whether to buy a product. Reviews often include text and a rating. The score, which is often a number from 1 to 5, with 1 being the worst and 5 being the best, is the reviewer's reflection on the text.

As illustrated in Figure 1, a survey by the marketing company Fan & Fuel (2023) found that 92% of consumers are swayed by the lack of online evaluations. This group expressed substantial uncertainty about what would happen next; 35% said they were less likely to buy, 32% said they would postpone their purchase until they could do more study, 23% said it would be challenging to make their decision, and 2% said they would simply not buy the product or service [2, 3].
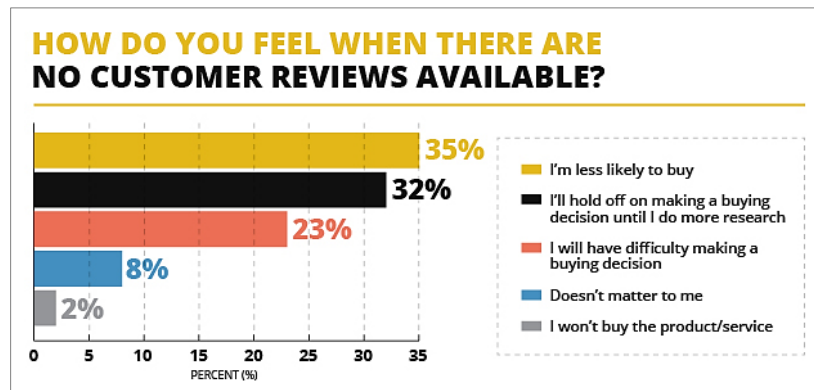
Figure 1. Poll Results on No Customer Reviews on An E-Commerce Shop [2]

Sentiment analysis is the activity of categorizing views and mining emotional words using text mining and natural language processing methods. Sentiment analysis involves a variety of tasks, approaches, and types of analysis. Sentiment analysis is an essential tool for many businesses, particularly e-commerce, to improve the quality of their products and support them in making wise decisions in the increasingly competitive business world of today. The three techniques utilized in sentiment analysis are lexicon-based, hybrid learning, and machine learning (ML). Each of these areas has its own division, as shown in Figure 2.

Supervised learning is the approach to machine learning that is most well-known and regularly utilized [4]. It is common to discuss both supervised and unsupervised machine learning together. In contrast to supervised learning, unsupervised learning uses unlabeled data. The patterns created from these data can be used to solve clustering or association problems. This is quite helpful when subject-matter experts are not aware of common features in a data set. K-means, hierarchical clustering methods, and Gaussian mixture models are commonly employed [5]. However, traditional sentiment analysis methods have encountered limitations in accurately capturing the intricacies of language, especially in the context of nuanced and context-dependent expressions.
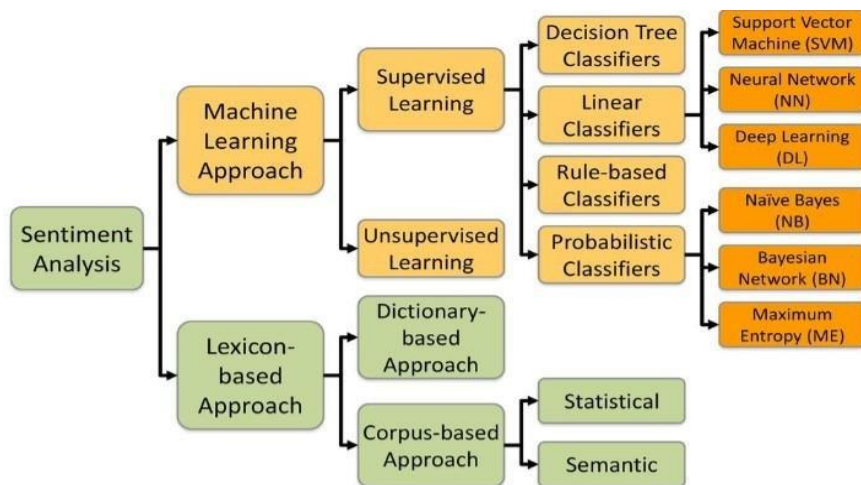

Figure 2. The Classification of Sentiment Techniques [5]

In recent years, a breakthrough in Natural Language Processing (NLP), known as "Bidirectional Encoder Representations from Transformers" (BERT), has emerged as a powerful and transformative technique, filling the gap in sentiment analysis by surpassing conventional approaches and enhancing language understanding. BERT is able to enhance context understanding and offers a pre-trained model that is quick and simple to modify for a range of downstream uses. There are also pre-trained models accessible in various languages. BERT can process more text and language despite the model's size (due to the training structure and corpus) and slow training time. It has a high degree of accuracy in the analysis and amplification of human-like languages [6]. In general, one of the most frequently used NLP models currently available is BERT. It takes little data and task-specific adjustment to deliver cutting-edge outcomes for a variety of NLP tasks. As a result, BERT has emerged as a game-changer in sentiment analysis, addressing the gaps left by conventional methods and altering the way businesses harness the power of customer sentiment for strategic decision-making and product enhancement. This review paper aims to explore and evaluate the applications of the BERT model in sentiment analysis across various fields and to highlight the significance of sentiment analysis in the context of e-commerce reviews and its impact on both customers and businesses. Additionally, the paper seeks to identify the strengths and weaknesses of the BERT model, particularly concerning its performance in sentiment analysis tasks.

2. Related Work

Vietnamese sentiment analysis with BERT fine-tuning, Nguyen et al. (2020) demonstrate that BERT could be merged with models built using Recurrent Convolutional Network (RCNN) or other recurrent and convolutional model-combining architectures. According to the experimental findings, the accuracy performance for sentiment analysis on datasets containing Vietnamese reviews is improved using the BERT-RCNN model [7].

Many scientists have been drawn to it and have tried to apply it to a variety of NLP applications. The challenges of text summarization, automated grading, text similarity score prediction, enhanced sentiment categorization, and reranking have all been the subject of several experiments in the past few years [8]. Some works of literature have been seen to work on sentiment analysis using the BERT model for different languages. The model has been used in Indonesian [9, 10] and Bangla customer feedback [11], which yields a high accuracy rate of 94.15% by combining Bangla-BERT and LSTM. BERT has been used in Chinese stock reviews [12], Arabic aspect-based reviews [13], and Urdu user reviews [14]. The model has been experimented with in Bahasa Melayu [15], French [16], and Malayalam-English [17]. BERT has also contributed to Ukrainian and Russian media reports [18], hate speech detection in Hindi-English [19], and Tamil-English mixed text classification [20].

Numerous studies have focused on the analysis of Twitter data sentiment due to easy access to a vast amount of real-time data [21–25]. The outcomes indicated the effectiveness of the BERT model in analyzing the sentiment provided by the users and showed considerable improvements in sentiment classification performance.

In the food service industry, a company may choose to alter the flavor or ingredients in specific locations to match the regional flavor recommended by the reviews, employing the most well-liked terms or phrases from the score ratings. This is accomplished by utilizing the BERT model, where the outcomes can be exploited to produce greater success [26]. In a study for Automated Essay Scoring (AES) [27], the BERT model was employed. It is said to be one of the most difficult issues in NLP. The essay's length, the presence of spelling errors that detract from its quality, and how the essay is represented in terms of the necessary criteria for effective essay grading are among the major problems faced. The BERT model was combined in various ways to assess the effectiveness of AES models. It was determined that deep-encoded features and manually extracted features both improve the functionality of AES models.

Additionally, the BERT model was used to analyze the sentiment of online reviews on Google Play [28]. The results gathered can help with app development. Yelp was also the subject of a sentiment analysis study employing the BERT model [29]. This study tackles Yelp's two main issues at the moment. First, it can be difficult for users to read every text-based review on Yelp due to the site's enormous volume of reviews. Second, Yelp's existing one-to-five star rating system lacks specificity, making it impossible to infer the consumers' motivations if they have given the same rating. The BERT model was able to determine if a review was good or negative based on its content and predict the strength of that positivity or negativity.

In the agriculture sector, consumers are able to assess the quality of agricultural goods, and businesses can improve and upgrade their products by applying sentiment analysis of online customer reviews of agricultural products. The BERT model-based agricultural assessment classification algorithm successfully identified the emotion conveyed in the text, assisting in the subsequent analysis of network evaluation data, the extraction of useful information, and the realization of emotion visualization [30].

A study was conducted by Durairaj & Chinnalagu (2021) to construct a refined BERT model to predict user attitudes using customer reviews from Yelp, Amazon, Twitter, and IMDB Movie Reviews. Hybrid fastText-BiLSTM, BiLSTM, fastText, and Linear Support Vector Machine (LSVM) models were compared. The proposed BERT model performed better in terms of accuracy and model performance, and the model training and data preparation procedures generally required less time. This experiment shows that, compared to other traditional models, the BERT model required greater CPU resources during training. Due to the improved BERT model, sentiment analysis on huge datasets is easier to do [31].

A study conducted by Sousa et al. (2019) in the stock market industry attempts to address the issues of news quantity and news analysis response times. In order to perform stock market sentiment analysis, they set out to examine BERT. The outcomes show that BERT outperforms word embeddings and convolutional neural networks in terms of performance. The results demonstrated that one can extract certain news from particular companies and conduct data processing and analysis on the value of their stock as future work. Additionally, one can observe news about a company, gather its accounting information, and develop a more accurate prediction. The outcomes could enhance the quality of financial agents' decisions [32].

A study conducted by Lee et al. (2022) used Word2Vec, Term Frequency Inverse Document Frequency (TF-IDF), BERT, and word embeddings to examine the effects of objectivity and subjectivity on sentiment analysis. There were two datasets used in this study: data from Wikipedia and Shopee user reviews. Results from their research indicate that BERT embedding, with an accuracy score of 99.77%, provided the best result for subjectivity classification [33].

3. Methods

In order to assess the effectiveness of BERT in comparison to more established techniques, the results of BERT are compared with a number of models, including Naive Bayes, Support Vector Machines (SVM), Random Forest, Long Short-Term Memory (LSTM), Bi-directional LSTMs, Decision Trees (DT), Valence-Aware Dictionarys, and Sentiment Reasoners (VADER). These methods are each described briefly.

The BERT model produces cutting-edge outcomes independent of the specific NLP problem. High-quality models can be produced quickly and effectively with little effort and training. It focuses on employing the novel Masked Language Model (MLM) as opposed to the conventional one-way language model or the technique of shallow splicing two one-way language models for pre-training in order to construct an intricate bidirectional language representation [34]. Despite the model being large and slow to train due to the training framework and corpus, BERT can process more text and language. It has a high degree of accuracy in its ability to analyze and fine-tune human-like languages. In essence, one of the most widely used NLP models at the moment is BERT. It provides cutting-edge outcomes for a number of NLP tasks with less data and task-specific adaptability [7, 9].

Naive Bayes is a simple yet effective statistics-based method for predictive modeling. The Bayesian theorem, based on likelihood, calculates the probability for each event. The highest probability output is expected since this method assumes that each characteristic is independent. The NB classifier has the benefit of using little training data while still producing effective results. The issue with this technique is that it makes poor estimates since it assumes that each attribute is independent [35–38].

SVM analyzes data and searches for patterns using a variety of directed learning approaches for regression classification and analysis. SVM has the benefits of performing exceedingly well when groups of data items are clearly distinct from one another, being able to be applied to both regression and classification problems, and being highly effective even with high-dimensional data. The difficult work of choosing the best kernel had to be completed, and it took more time to train SVM on a big data set when classes in the data were not well divided by points, as this indicates the presence of overlapping classes [39–43].

RF is an ensemble method that uses many different decision trees. By averaging the results of various decision trees, the RF algorithm's output is determined. It automatically fills up any data that has missing values. Consequently, as the number of trees grows, the RF's accuracy grows as well. Additionally, the RF technique solves the overfitting issue that the DT algorithm encountered [39, 41, 43–45].

Long-term dependency is a problem with RNNs that is solved by LSTM. LSTM also addresses the issues with vanishing gradients and extending gradients that emerge throughout the training phase. In contrast to the majority of ML models, LSTM has a long-term memory. This is made possible by its architecture's cell-named explicit memory unit. LSTMs are built as a series of repeated neural network modules. One of its shortcomings is that, because of its complexity, it consumes more resources than standard RNNs. Compared to standard RNNs, it takes longer to train. The interpretation of LSTM can be difficult [44, 46–49].

Bi-directional LSTMs are used to educate both the forward and backward time dependencies. It resolves the fixed sequence-to-sequence prediction issue. Each unit is divided into two independent ones in a bidirectional LSTM, each of which is linked to the same output and has the same input. The forward time sequence employs one unit, whereas the reverse time sequence employs the other. As a result, while learning from time-series data with a long history, it shows improved results without lengthening the training period. Bi-directional LSTM is expensive since it uses two LSTM cells [44, 47, 49].

The DT is a logic-based method that divides a single complex decision into numerous straightforward, easier judgments. It is a mathematical model that is used to depict the process of making decisions. This method allows for the construction of a logical tree with numerous tiers of logical conditions and possibilities to get the desired outcome [50–53].

The primary tool that VADER uses to analyze emotions and sentiments is its diction, which developers must download in order to execute the tool. The dictionary records whether a term is good, neutral, or negative. Additionally, a compound score is kept for each word. VADER calculates the compound score of the sentence after compiling the compound scores for each word contained in the sentiment. The sentiment is positive if the score is higher than the cutoff point; otherwise, it is negative. Due to its simplicity in implementation and adjustment, VADER has an advantage. Despite the benefit, VADER has a drawback when interacting with terms it is unfamiliar with. If a word is found outside of VADER's diction, it merely receives a neutral score of 1 and a compound score of 0. Furthermore, developing the diction for VADER is both costly and time-consuming [54, 55].

3.1. BERT Model Fine-Tuning

BERT performed better than average on a range of general language understanding tasks, including sentiment analysis, question-answering, paraphrase identification, and linguistic acceptability. Think about the case where we are

creating a question-and-answer application. When a question is given as input, the application's goal is to choose a suitable response from a corpus. This is fundamentally a prediction problem. The model then uses a question and a context paragraph to predict a start token and an end token from the paragraph that most likely answers the query. Therefore, using BERT, a model for our application may be created by learning two more vectors that signify the start and end of the response [21].

Before the text data is sent to the BERT model, it will be cleaned up using text pre-processing. In the text pre-processing and numerical conversion workflow, raw text data undergoes cleaning steps to remove noise, punctuation, and variable capitalization. After lowercasing and tokenization, common stopwords are eliminated to streamline the data. The final stage involves converting the tokenized text into a numerical format, enabling machines to process the information effectively. This processed and numerical representation of the text is then ready to be fed into the BERT model for sentiment analysis or other natural language processing tasks. Figure 3 shows the data pre-processing steps of the BERT model.
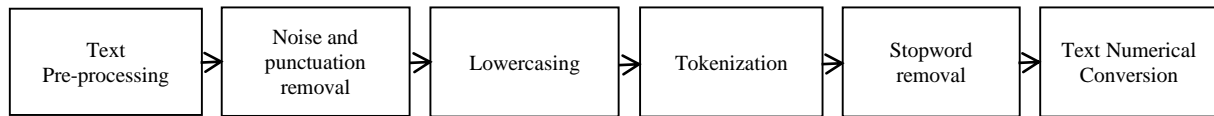

Figure 3. Data Pre-Processing steps

Numerous downstream activities, such as categorization and question-answering, are made possible by the BERT architecture [26]. A pre-trained BERT will produce H = 768-shaped vectors, which are intended to be a black box, for each input token (word) in a sequence. Here, the sequence may begin with a token [CLS] and may contain either a single sentence or two sentences divided by the separator [SEP]. Figure 4 shows the overall process of the BERT model.
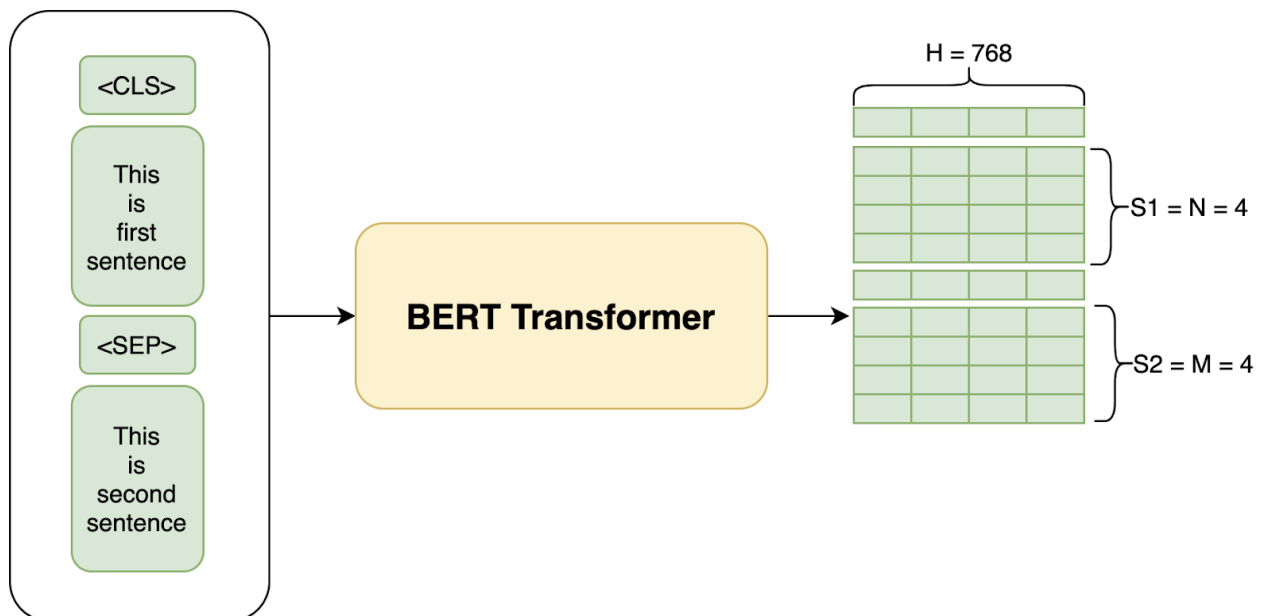

Figure 4. Overall Process of BERT Model [56]

4. Results and Discussion

In the study conducted by Kang et al. (2021), sentiment analysis was performed to gauge the public sentiment towards Malaysian Airlines using six different models, namely the Linear Support Vector Classifier, BERT Model, Ensemble Method, Multinomial Naive Bayes, Bi-LSTM, and Random Forest [43]. The research findings revealed that deep learning techniques exhibited superior performance compared to traditional machine learning approaches. Specifically, the Bidirectional LSTM achieved an accuracy of 77%, while the BERT model outperformed all other models with an impressive accuracy of 86%. The experiments further demonstrated that BERT's performance surpassed that of common pre-processing methods, such as decapitalization, punctuation removal, stopword removal, and emoji conversion to text. Additionally, the BERT model also surpassed the results of unsupervised text categorization, indicating its ability to effectively capture and analyze sentiment patterns. Overall, the research highlighted the remarkable effectiveness of the BERT model, establishing it as a powerful tool for sentiment analysis and affirming its superiority over the other models tested in the study [44].

In the Naver (2021) study, the objective was to classify Swedish sentences based on their tenses using LSTM, Naive Bayes, and BERT models. The results demonstrated that BERT outperformed the other models, achieving an impressive

accuracy of 96.3% [37]. This accuracy level aligns with the findings of a study by Holmer [56], who also employed the same pre-trained BERT model to classify Swedish text. The superior performance of the BERT model can be attributed to its extensive pre-training on Swedish language data, which allowed it to grasp the intricacies of the language better. As a result, fine-tuning BERT for specific tasks might not require as much additional training data. Moreover, the researcher found out that BERT's bidirectional nature enables it to excel in distinguishing between words with similar spellings but different meanings, as well as being more contextually aware, which provides a significant advantage in language understanding and classification tasks [37].

In another study done by Geetha & Karthika Renuka (2021), consumer review data was categorized into positive and negative emotions using sentiment analysis. LSTM, BERT, Naive Bayes Classification, and SVM were used to classify reviews using the various classification models. BERT outperforms other predictive models in terms of accuracy, according to performance evaluation criteria and comparison. Tests that combined the results of the BERT model with the performance of other machine learning algorithms showed that the BERT model outperformed other machine learning algorithms in terms of performance measures. BERT produces better accuracy, which was 88.48%. Many of the SA methods currently in use for this text data from online customer product reviews are erroneous and frequently require more training time. This study demonstrated that the sentiment analysis problem might be resolved using the BERT model, a potent Deep Learning model. The BERT model outperformed the other machine learning techniques in the experimental evaluation with high prediction and good accuracy [38].

A different study revealed that the BERT model's accuracy was 79% in estimating reviewer satisfaction from the text description of Amazon Fine Food [26]. The researchers employed three epochs and a learning rate of 1e-5. They used so few epochs in their model analysis because the model was well-trained and only a small number of epochs were required for fine-tuning. The approach helps food service businesses forecast their overall performance locally or nationwide. One of the challenges encountered during the research was the possibility that false reviews could affect the model's accuracy. The researchers suggested that the accuracy of the BERT model would be increased if a false review categorization model was already created and included in the model analysis. This would help remove the fake reviews. A word cloud was done to analyze the sentiment of customers, as shown in Figure 5. The words with the highest frequency were "well", "gluten-free", and "taste like". It can be seen that the vast majority of customers have worries about products that contain gluten. The comparison of foods is another way that individuals use the phrase "taste like."

Figure 5. Word Cloud of the Sentiment among Amazon Fine Food Customers [16]

In an experiment conducted by Azhar & Khodra (2021) to assess Indonesian aspect-based sentiment analysis using the BERT model, the researchers used a batch size of 32 with a learning rate of 2e-5 and 25 epochs. They have encountered difficulties primarily related to misclassifications in the data. These misclassifications were attributed to incorrectly labeled data and the presence of keywords for aspect categories or other emotions that were mistakenly considered neutral in sentiment polarity. Consequently, such instances couldn't be definitively associated with either positive or negative feelings. Additionally, a significant challenge arose when a single statement contained conflicting attitudes toward one or more aspect categories. In such cases, the occurrences were marked as having negative sentiments, even when the data also contained keywords with positive sentiments. This conflicting information made it challenging for the BERT Model to make accurate predictions, leading to sub-optimal performance in sentiment analysis [10].

In research conducted by Kusnadi et al. (2021), the BERT model was applied to analyze sentiment in a dataset from Google Play's Genshin Impact mobile game [28]. The researchers used a fine-tuning hyper-parameter of 32 batch size with a learning rate (Adam) of 2e-5 and 10 epochs. The model demonstrated impressive performance in predicting positive sentiment, achieving a precision score of 0.86%, an f1-score of 0.82%, and a recall score of 0.78%. These results are promising and offer valuable insights for game improvement, as positive sentiment plays a crucial role in user satisfaction and engagement. However, the research also highlighted a challenge faced in sentiment analysis, which is

the detection of neutral sentiment, which proved to be more difficult for the model [28]. The accuracy score for neutral sentiment was relatively lower compared to the scores for positive and negative sentiments. This observation aligns with the findings of other studies [10], underscoring the complexity of accurately classifying neutral sentiments in text data.

5. Conclusion

This paper provides a comprehensive overview of the versatile applications of the BERT Model across diverse fields and languages, showcasing its capability in resolving a range of challenging NLP tasks, including sentiment analysis for discerning positive and negative reviews. The contextualized, pre-trained language representations of BERT have proven to be highly effective in capturing the nuanced meanings and semantic relationships within text data, elevating sentiment analysis to new heights of accuracy and understanding. However, the BERT model does encounter certain challenges that warrant further attention. Notably, the model faces difficulties in accurately detecting neutral sentiment, which can impact the overall performance of sentiment analysis tasks. Moreover, the presence of false reviews within the dataset poses a significant hurdle to the model's accuracy, potentially leading to misclassifications. In order to address this issue, researchers can look into the integration of a dedicated fake review classification model into the BERT analysis pipeline, enabling the identification and elimination of false reviews. Another aspect that demands exploration is the effective handling of conflicting sentiments present within the data. As some reviews may contain mixed emotions or ambiguous expressions, devising robust strategies to address such misclassification scenarios becomes crucial for further improving the BERT model's accuracy. Researchers should endeavor to devise innovative techniques that allow the model to better weigh and contextualize multiple sentiments within a single text, enabling more accurate sentiment predictions and reducing misclassifications.

In conclusion, the BERT Model's significant performance in sentiment analysis and other NLP tasks has revolutionized the field, offering a powerful tool for gaining deeper insights into customer sentiments and preferences. In order to harness its full potential, addressing the challenges of detecting neutral sentiment and handling false reviews is essential. This allows BERT's effectiveness in sentiment analysis to be further optimized, increasing its position as an important instrument for businesses seeking to make informed decisions and cater to customer needs in an increasingly competitive landscape.

6. Declarations

6.1. Author Contributions

Conceptualization, M.S.S. and V.M.; methodology, M.S.S.; validation, V.M.; formal analysis, V.M.; investigation, K.S.M.; writing—original draft preparation, V.M. and M.S.S.; writing—review and editing, M.S.S. and K.S.M.; visualization, V.M.; supervision, M.S.S. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

No new data were created or analyzed in this study. Data sharing is not applicable to this article.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. 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.

7. References

[1] Haque, T. U., Saber, N. N., & Shah, F. M. (2018). Sentiment analysis on large scale Amazon product reviews. 2018 IEEE International Conference on Innovative Research and Development (ICIRD). doi:10.1109/icird.2018.8376299.

[2] Fan & Fuel. No online customer reviews mean BIG problems in 2017. Available online: https://fanandfuel.com/no-online-customer-reviews-means-big-problems-2017/ (accessed on April 2023).

[3] Gumelar, A. B., Yogatama, A., Adi, D. P., Frismanda, & Sugiarto, I. (2022). Forward feature selection for toxic speech classification using support vector machine and random forest. IAES International Journal of Artificial Intelligence, 11(2), 717–726. doi:10.11591/ijai.v11.i2.pp717-726.

[4] Noor, F., Bakhtyar, M., & Baber, J. (2019). Sentiment Analysis in E-commerce Using SVM on Roman Urdu Text. Emerging Technologies in Computing. iCETiC 2019, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 285, Springer, Cham, Switzerland. doi:10.1007/978-3-030-23943-5_16.

[5] Ying, O. J., Zabidi, M. M. A., Ramli, N., & Sheikh, U. U. (2020). Sentiment analysis of informal Malay tweets with deep learning. IAES International Journal of Artificial Intelligence, 9(2), 212–220. doi:10.11591/ijai.v9.i2.pp212-220.

[6] Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to Fine-Tune BERT for Text Classification?. Chinese Computational Linguistics. CCL 2019. Lecture Notes in Computer Science, 11856, Springer, Cham, Switzerland. doi:10.1007/978-3-030-32381-3_16.

[7] Nguyen, Q. T., Nguyen, T. L., Luong, N. H., & Ngo, Q. H. (2020). Fine-Tuning BERT for Sentiment Analysis of Vietnamese Reviews. 2020 7th NAFOSTED Conference on Information and Computer Science (NICS). doi:10.1109/nics51282.2020.9335899.

[8] Chakkarwar, V., Tamane, S., & Thombre, A. (2023). A Review on BERT and Its Implementation in Various NLP Tasks. Proceedings of the International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022), 112–121. doi:10.2991/978-94-6463-136-4_12.

[9] Nugroho, K. S., Sukmadewa, A. Y., Wuswilahaken DW, H., Bachtiar, F. A., & Yudistira, N. (2021). BERT Fine-Tuning for Sentiment Analysis on Indonesian Mobile Apps Reviews. 6th International Conference on Sustainable Information Engineering and Technology 2021. doi:10.1145/3479645.3479679.

[10] Azhar, A. N., & Khodra, M. L. (2020). Fine-tuning Pretrained Multilingual BERT Model for Indonesian Aspect-based Sentiment Analysis. 2020 7th International Conference on Advance Informatics: Concepts, Theory and Applications (ICAICTA). doi:10.1109/icaicta49861.2020.9428882.

[11] Prottasha, N. J., Sami, A. A., Kowsher, M., Murad, S. A., Bairagi, A. K., Masud, M., & Baz, M. (2022). Transfer Learning for Sentiment Analysis Using BERT Based Supervised Fine-Tuning. Sensors, 22(11), 4157. doi:10.3390/s22114157.

[12] Li, M., Chen, L., Zhao, J., & Li, Q. (2021). Sentiment analysis of Chinese stock reviews based on BERT model. Applied Intelligence, 51(7), 5016–5024. doi:10.1007/s10489-020-02101-8.

[13] M.Abdelgwad, M., A Soliman, T. H., I.Taloba, A., & Farghaly, M. F. (2022). Arabic aspect based sentiment analysis using bidirectional GRU based models. Journal of King Saud University - Computer and Information Sciences, 34(9), 6652–6662. doi:10.1016/j.jksuci.2021.08.030.

[14] Khan, L., Amjad, A., Ashraf, N., & Chang, H. T. (2022). Multi-class sentiment analysis of Urdu text using multilingual BERT. Scientific Reports, 12(1). doi:10.1038/s41598-022-09381-9.

[15] Chong, W. K., Ng, H., Yap, T. T. V., Soo, W. K., Goh, V. T., & Cher, D. T. (2022). Objectivity and Subjectivity Classification with BERT for Bahasa Melayu. Proceedings of the International Conference on Computer, Information Technology and Intelligent Computing (CITIC 2022), 246–257. doi:10.2991/978-94-6463-094-7_20.

[16] Martin, L., Muller, B., Ortiz Suárez, P. J., Dupont, Y., Romary, L., de la Clergerie, É., Seddah, D., & Sagot, B. (2020). CamemBERT: a Tasty French Language Model. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. doi:10.18653/v1/2020.acl-main.645.

[17] Chakravarthi, B. R., Jose, N., Suryawanshi, S., Sherly, E., & McCrae, J. P. (2020). A sentiment analysis dataset for code-mixed Malayalam-English. arXiv preprint arXiv:2006.00210. doi:10.48550/arXiv.2006.00210.

[18] Salnikova, S., & Kyrychenko, R. (2021). Sentiment Analysis Based on the BERT Model: Attitudes Towards Politicians Using Media Data. Advances in Social Science, Education and Humanities Research. doi:10.2991/assehr.k.211218.007.

[19] Rani, P., Suryawanshi, S., Goswami, K., Chakravarthi, B. R., Fransen, T., & McCrae, J. P. (2020). A comparative study of different state-of-the-art hate speech detection methods in Hindi-English code-mixed data. Proceedings of the second workshop on trolling, aggression and cyberbullying, May 2020, Marseille, France.

[20] Bhargava, N., & Johari, A. (2023). Enhancing Deep Learning Approach for Tamil English Mixed Text Classification. Proceedings of the International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022), 829–837. doi:10.2991/978-94-6463-136-4_73.

[21] Pota, M., Ventura, M., Catelli, R., & Esposito, M. (2021). An effective BERT-based pipeline for twitter sentiment analysis: A case study in Italian. Sensors (Switzerland), 21(1), 1–21. doi:10.3390/s21010133.

[22] Singh, M., Jakhar, A. K., & Pandey, S. (2021). Sentiment analysis on the impact of coronavirus in social life using the BERT model. Social Network Analysis and Mining, 11(1). doi:10.1007/s13278-021-00737-z.

[23] Gao, Z., Feng, A., Song, X., & Wu, X. (2019). Target-dependent sentiment classification with BERT. IEEE Access, 7, 154290–154299. doi:10.1109/ACCESS.2019.2946594.

[24] Nguyen, D. Q., Vu, T., & Tuan Nguyen, A. (2020). BERTweet: A pre-trained language model for English Tweets. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 9-14. doi:10.18653/v1/2020.emnlp-demos.2.

[25] Benitez-Andrades, J. A., Alija-Perez, J. M., Garcia-Rodriguez, I., Benavides, C., Alaiz-Moreton, H., Vargas, R. P., & Garcia-Ordas, M. T. (2021). BERT Model-Based Approach for Detecting Categories of Tweets in the Field of Eating Disorders (ED). 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS). doi:10.1109/cbms52027.2021.00105.

[26] Zhao, X., & Sun, Y. (2022). Amazon Fine Food Reviews with BERT Model. Procedia Computer Science, 208, 401–406. doi:10.1016/j.procs.2022.10.056.

[27] Beseiso, M., & Alzahrani, S. (2020). An empirical analysis of BERT embedding for automated essay scoring. International Journal of Advanced Computer Science and Applications, 11(10), 204–210. doi:10.14569/IJACSA.2020.0111027.

[28] Kusnadi, R., Yusuf, Y., Andriantony, A., Ardian Yaputra, R., & Caintan, M. (2021). Sentiment Analysis of the Game Genshin Impact Using Bert. Rabit: Jurnal Teknologi Dan Sistem Informasi Univrab, 6(2), 122–129. doi:10.36341/rabit.v6i2.1765. (In Indonesian).

[29] Lee, S. (2022). Sentiment Analysis Using Bert on Yelp Restaurant Reviews. Ph.D. Thesis, Purdue University Graduate School, West Lafayette, United States.

[30] Cao, Y., Sun, Z., Li, L., & Mo, W. (2022). A Study of Sentiment Analysis Algorithms for Agricultural Product Reviews Based on Improved BERT Model. Symmetry, 14(8). doi:10.3390/sym14081604.

[31] Durairaj, A. K., & Chinnalagu, A. (2021). Transformer based Contextual Model for Sentiment Analysis of Customer Reviews: A Fine-tuned BERT. International Journal of Advanced Computer Science and Applications, 12(11), 474-480. doi:10.14569/ijacsa.2021.0121153.

[32] Sousa, M. G., Sakiyama, K., Rodrigues, L. de S., Moraes, P. H., Fernandes, E. R., & Matsubara, E. T. (2019). BERT for Stock Market Sentiment Analysis. 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). doi:10.1109/ictai.2019.00231.

[33] Lee, X. J., Yap, T. T. V., Ng, H., & Goh, V. T. (2022). Comparison of Word Embeddings for Sentiment Classification with Preconceived Subjectivity. Proceedings of the International Conference on Computer, Information Technology and Intelligent Computing (CITIC 2022), 488–502. doi:10.2991/978-94-6463-094-7_39.

[34] Su, Y. (2022). Emotion Analysis of Microblog Epidemic Coexistence Based on BERT. Proceedings of the 2022 2nd International Conference on Modern Educational Technology and Social Sciences (ICMETSS 2022), 275–280. doi:10.2991/978-2-494069-45-9_33.

[35] Abbas, M., Ali Memon, K., & Aleem Jamali, A. (2019). Multinomial Naive Bayes Classification Model for Sentiment Analysis. IJCSNS International Journal of Computer Science and Network Security, 19(3), 62.

[36] Shamrat, F. M. J. M., Chakraborty, S., Imran, M. M., Muna, J. N., Billah, M. M., Das, P., & Rahman, M. O. (2021). Sentiment analysis on twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm. Indonesian Journal of Electrical Engineering and Computer Science, 23(1), 463–470. doi:10.11591/ijeecs.v23.i1.pp463-470.

[37] Navér, N. (2021). The past, present or future?: A comparative NLP study of Naive Bayes, LSTM and BERT for classifying Swedish sentences based on their tense. Master of Science Programme in Information Technology Engineering, Uppsala Universitet, Uppsala, Sweden.

[38] Geetha, M. P., & Karthika Renuka, D. (2021). Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model. International Journal of Intelligent Networks, 2, 64–69. doi:10.1016/j.ijin.2021.06.005.

[39] Hussain, M. A., Chen, Z., Wang, R., Shah, S. U., Shoaib, M., Ali, N., ... & Ma, C. (2022). Landslide Susceptibility Mapping using Machine Learning Algorithm. Civil Engineering Journal, 8(2), 209-224. doi:10.28991/CEJ-2022-08-02-02.

[40] Nurhidayat, I., Pimpunchat, B., Noeiaghdam, S., & Fernández-Gámiz, U. (2022). Comparisons of SVM kernels for insurance data clustering. Emerging Science Journal, 6(4), 866-880. doi:10.28991/ESJ-2022-06-04-014.

[41] Singh, J., & Tripathi, P. (2021). Sentiment analysis of Twitter data by making use of SVM, Random Forest and Decision Tree algorithm. 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT). doi:10.1109/csnt51715.2021.9509679.

[42] Akhtar, M. B. (2022). The use of a convolutional neural network in detecting soldering faults from a printed circuit board assembly. HighTech and Innovation Journal, 3(1), 1-14. doi:10.28991/HIJ-2022-03-01-01.

[43] Kang, H. W., Chye, K. K., Ong, Z. Y., & Tan, C. W. (2021). The Science of Emotion: Malaysian Airlines Sentiment Analysis using BERT Approach. Conference Proceedings: International Conference on Digital Transformation and Applications (ICDXA 2021). doi:10.56453/icdxa.2021.1013.

[44] Huang, X., Zhang, W., Tang, X., Zhang, M., Surbiryala, J., Iosifidis, V., Liu, Z., & Zhang, J. (2021). LSTM Based Sentiment Analysis for Cryptocurrency Prediction. Database Systems for Advanced Applications. DASFAA 2021, Lecture Notes in Computer Science, 12683, Springer, Cham, Switzerland. doi:10.1007/978-3-030-73200-4_47.

[45] Kumar, P., & Singh, A. K. (2022). A comparison between MLR, MARS, SVR and RF techniques: hydrological time-series modeling. Journal of Human, Earth, and Future, 3(1), 90-98. doi:10.28991/HEF-2022-03-01-07.

[46] Minaee, S., Azimi, E., & Abdolrashidi, A. (2019). Deep-sentiment: Sentiment analysis using ensemble of CNN and Bi-LSTM models. arXiv preprint arXiv:1904.04206. doi:10.48550/ARXIV.1904.04206.

[47] Wen, S., Wei, H., Yang, Y., Guo, Z., Zeng, Z., Huang, T., & Chen, Y. (2021). Memristive LSTM Network for Sentiment Analysis. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 51(3), 1794–1804. doi:10.1109/TSMC.2019.2906098.

[48] Zhou, J., Lu, Y., Dai, H. N., Wang, H., & Xiao, H. (2019). Sentiment analysis of Chinese microblog based on stacked bidirectional LSTM. IEEE Access, 7, 38856–38866. doi:10.1109/ACCESS.2019.2905048.

[49] Yang, G., & Guo, X. (2023). Application of Decision Tree Mining Algorithm in Data Model Collection of Hydraulic Engineering Equipment. Proceedings of the 4th International Conference on Big Data Analytics for Cyber-Physical System in Smart City - Volume 1. BDCPS 2022, Lecture Notes on Data Engineering and Communications Technologies, 167, Springer, Singapore. doi:10.1007/978-981-99-0880-6_76.

[50] Mookdarsanit, P., & Mookdarsanit, L. (2021). The covid-19 fake news detection in Thai social texts. Bulletin of Electrical Engineering and Informatics, 10(2), 988–998. doi:10.11591/eei.v10i2.2745.

[51] Phan, H. T., Tran, V. C., Nguyen, N. T., & Hwang, D. (2019). Decision-Making Support Method Based on Sentiment Analysis of Objects and Binary Decision Tree Mining. Advances and Trends in Artificial Intelligence. From Theory to Practice. IEA/AIE 2019. Lecture Notes in Computer Science, 11606, Springer, Cham, Switzerland. doi:10.1007/978-3-030-22999-3_64.

[52] Rathi, M., Malik, A., Varshney, D., Sharma, R., & Mendiratta, S. (2018). Sentiment Analysis of Tweets Using Machine Learning Approach. 2018 Eleventh International Conference on Contemporary Computing (IC3). doi:10.1109/ic3.2018.8530517.

[53] Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proceedings of the International AAAI Conference on Web and Social Media, 8(1), 216–225. doi:10.1609/icwsm.v8i1.14550.

[54] Bajaj, A. Can Python understand human feelings through words? – A brief intro to NLP and VADER Sentiment Analysis. Analytics Vidhya. Available online: https://www.analyticsvidhya.com/blog/2021/06/vader-for-sentiment-analysis/ (accessed on May 2023).

[55] Holmer, D. (2020). Context matters: Classifying Swedish texts using BERT's deep bidirectional word embeddings. Department of Computer and Information Science, Linköping University, Linköping, Sweden.

[56] Wu, K. (2023). BERT Transformers: How Do They Work?. Available online: https://dzone.com/articles/bert-transformers-how-do-they-work (accessed on May 2023).