Journal of Theoretical and Applied Information Technology
15th October 2023. Vol.101. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SENTIMENT ANALYSIS BASED ON 7P MARKETING MIX ASPECTS OF THE INDRIVER APPLICATION SERVICE USING THE BERT ALGORITHM, BASED ON USER REVIEWS ON THE GOOGLE PLAY STORE

[1]HANDRIZAL, [2]T. HENNY FEBRIANA HARUMY, [3]HERRIYANCE, [4]M. AMIRUL ILMI
[1,2,3,4]Department of Computer Science, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Jl. University No. 9-A, Medan 20155, Indonesia

E-mail: handrizal@usu.ac.id

ABSTRACT

This research aims to perform sentiment analysis on user reviews of the InDriver application service on the Google Play Store, focusing on the 7P marketing mix aspects. The analysis utilizes the BERT (Bidirectional Encoder Representations from Transformers) algorithm, known for its contextual text understanding and rich representation generation capabilities. The sentiment classification includes positive, negative, and neutral sentiments. The study collects data through scraping with the Google Play Scraper, followed by preprocessing steps such as tokenization and normalization. The collected dataset consists of 3028 user reviews. Three experiments are conducted, varying hyperparameters such as epochs, learning rate, and batch size. The research findings demonstrate the significant accuracy of the sentiment analysis using the BERT algorithm. The first experiment achieves an accuracy of 75%, while the second and third experiments achieve accuracies of 83% each. The results highlight the BERT algorithm's ability to effectively classify user sentiments towards the InDriver application service. This research contributes to understanding user sentiments, providing valuable insights for decision-making and product enhancement. However, the study acknowledges its limitations and suggests areas for further development, including increasing the dataset size or adding additional preprocessing features.

Keywords: *Sentiment Analysis, BERT Algorithm, 7P Marketing Mix, InDriver Application, User Reviews.*

1. INTRODUCTION

In today's digital era, mobile applications have become an integral part of people's lives. The success of an application depends not only on its functionality but also on the overall user experience and satisfaction. To ensure a positive user experience, companies need to understand user sentiments and preferences regarding their applications. One effective way to gain insights into user sentiments is through sentiment analysis [1].

Sentiment analysis, also known as opinion mining, is a technique used to determine the sentiment expressed in a piece of text, such as user reviews or social media posts. By analyzing user sentiments, companies can identify areas of improvement, make informed business decisions, and enhance their products and services accordingly [2].

InDriver is a popular ride-hailing application that provides a unique approach to transportation services. It allows passengers to negotiate fares directly with drivers, giving them more control and potentially lower prices. Understanding user sentiments towards the InDriver application is crucial for its success and further improvement [3].

The 7P marketing mix framework is widely used to analyze and evaluate marketing strategies. It consists of seven elements: product, price, place, promotion, people, process, and physical evidence. By incorporating the 7P marketing mix aspects into sentiment analysis, companies can gain a deeper understanding of how each element influences user sentiments [4].

The BERT (Bidirectional Encoder Representations from Transformers) algorithm has gained significant attention and success in natural language processing tasks, including sentiment analysis. BERT is a state-of-the-art language model that utilizes transformer architectures to capture the context and meaning of text effectively. Its ability to understand the semantic relationships between words and generate rich representations makes it an ideal choice for sentiment analysis tasks [5].

Therefore, this study aims to perform sentiment analysis on user reviews of the InDriver application service available on the Google Play Store. The analysis will focus on the 7P marketing mix aspects to investigate how each element contributes to user sentiments. The BERT algorithm will be employed to classify the sentiments expressed in the reviews accurately. The findings of this research will provide valuable insights for InDriver and other companies in the ride-hailing industry to improve their services based on user preferences and sentiments.

What the Study Includes:

1. Sentiment Analysis of User Reviews: The study will analyze user reviews of the InDriver application available on the Google Play Store to determine their sentiments. This analysis will provide insights into the overall user perception of the application.
2. Examination of 7P Marketing Mix Aspects: The study will investigate how the different elements of the 7P marketing mix (product, price, place, promotion, people, process, and physical evidence) contribute to user sentiments. This examination will provide a comprehensive understanding of the factors affecting user satisfaction.
3. Application of BERT Algorithm: The study will employ the BERT algorithm to accurately classify sentiments expressed in user reviews. BERT's advanced language understanding capabilities will enhance the accuracy of sentiment classification.

What the Study Does Not Include:

1. Analysis of Competing Applications: The study will focus solely on user sentiments towards the InDriver application and will not analyze sentiments towards other ride-hailing applications.
2. In-depth Technical Details of BERT: The study will utilize the BERT algorithm for sentiment analysis, but it will not delve into detailed technical explanations of the BERT model.
3. User Demographic Analysis: While the study aims to understand user sentiments, it will not extensively analyze user demographics, socio-economic factors, or other individual characteristics.
4. Business Strategies: The study will not provide specific business strategies or recommendations for InDriver's improvement based on the sentiment analysis. It will focus on insights derived from user sentiments.

The addressed problem in the text is the importance of understanding user sentiments and preferences regarding mobile applications, particularly in the context of the InDriver ride-hailing application. The text highlights that the success of an application goes beyond functionality and extends to user experience and satisfaction. It emphasizes the need for companies to conduct sentiment analysis to gain insights into user sentiments, enabling them to identify areas for improvement and enhance their products and services.

Literature Screening Criteria:

1. Relevance to User Sentiments and Mobile Applications.
2. Focus on User Experience and Satisfaction.
3. Application of Sentiment Analysis Techniques.
4. Incorporation of Marketing Mix Concepts.
5. Utilization of BERT Algorithm.
6. Relevance to Ride-Hailing Industry.
7. Recent and Credible Sources.

Prior investigations have explored the landscape of sentiment analysis, delving into user sentiments in varied domains, including mobile applications. These studies have underscored the significance of understanding user experiences and sentiments to optimize application functionalities.

However, our study differentiates itself through its integration of the 7P marketing mix framework into sentiment analysis, offering a unique perspective on how marketing elements interplay with user sentiments. Furthermore, the utilization of the BERT algorithm sets our research apart, as it capitalizes on a state-of-the-art natural language processing tool that profoundly enhances sentiment analysis accuracy.

By amalgamating sentiment analysis, the 7P marketing mix framework, and BERT's capabilities, our research unveils an innovative approach to comprehending user sentiments and preferences. This synthesis of methodologies has the potential to uncover nuanced insights previously unexplored, positioning our study as a trailblazer in the realm of user sentiment analysis within the context of ride-hailing applications like InDriver.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. LITERATURE REVIEW

2.1 Sentiment Analysis

Sentiment analysis is a technique used to evaluate the feelings or opinions expressed by individuals in their comments or opinions [6]. In sentiment analysis, natural language processing (NLP) methods and techniques are employed to identify and categorize text into different sentiment classifications. This processing can be done using machine learning algorithms such as rule-based learning, deep learning, or hybrid approaches [7].

The results of sentiment analysis can be used for various purposes, such as understanding customer responses to products or services, monitoring public opinions about a brand or company, or identifying trends and patterns in user feedback [8]. There are several characteristics of sentiment analysis, including:

1. Natural Language Processing (NLP): Sentiment analysis utilizes NLP techniques and methods to understand and analyze text containing opinions or sentiments.
2. Sentiment Classification: Sentiment analysis categorizes text into different sentiment classifications, such as positive, negative, or neutral. This helps in understanding the general attitude of people towards a particular topic or entity.
3. Sentiment Scale: In addition to binary sentiment classification (positive/negative), sentiment analysis can provide more detailed sentiment scales, such as strongly positive, moderately positive, strongly negative, moderately negative, or neutral.
4. Entity Identification: Sentiment analysis can identify and label entities mentioned in the text, such as brands, products, or individuals, which helps in understanding specific sentiments towards those entities.
5. Emotions and Opinions: Apart from general sentiment, sentiment analysis can identify emotions related to the text, such as happiness, anger, sadness, and disappointment. Additionally, sentiment analysis can identify specific opinions expressed in the text.
6. Diverse Data Sources: Sentiment analysis can be applied to various data sources, including product reviews, social media comments, news articles, customer surveys, and more.
7. Real-Time Monitoring: Sentiment analysis can be performed in real-time to monitor and track changes in sentiment and opinions over time. This enables quick and responsive decision-making in dynamic business situations.
8. Wide Applications: Sentiment analysis has a wide range of applications, such as brand management, business decision-making, online reputation monitoring, product development, and data-driven marketing.

These characteristics help in understanding and extracting insights from vast textual data, providing a deeper understanding of people's sentiments and opinions toward a particular topic or entity.

2.2 Marketing Mix

According to [9], the marketing mix is a combination of four marketing elements that are tailored to the goals and business strategies of a company, namely product, price, promotion, and distribution. The marketing mix, also commonly known as the 7P, consists of factors that influence the success of marketing goods or services. The 7P is a combination of Product, Price, Place, Promotion, People, Process, and Physical Evidence.

1. Product: It refers to everything that a company can offer to meet the needs and desires of consumers, including both goods and services.
2. Price: It represents the value exchanged by consumers to obtain the products offered by the company.
3. Place: It involves the distribution process of products and services to consumers, including channel selection, transportation, and sales.
4. Promotion: It encompasses the company's activities to promote the products and services offered, including advertising, sales promotions, public relations, and personal selling.
5. People: It includes the workforce involved in providing products and services, including employees, management, and business partners.
6. Process: It refers to the business processes and services carried out by the company to meet the needs and desires of consumers, including ordering, production, delivery, and after-sales service.
7. Physical Evidence: It represents the physical or non-physical evidence that demonstrates the quality of the products and services offered. This includes product design, packaging, service quality, customer testimonials, and certifications.

The marketing mix elements work together to create a comprehensive marketing strategy that addresses various aspects of the business and ensures the success of marketing efforts.

2.3 BERT Algorithm

BERT (Bidirectional Encoder Representations from Transformers) is a natural language processing (NLP) algorithm introduced by Google in 2018. It is a transformer-based model that revolutionized the field of NLP by significantly improving the understanding of contextual language semantics [10].

BERT is a pre-trained language model that is trained on a large corpus of text data, such as books, articles, and websites. It learns to predict missing words in a sentence by considering the context of the surrounding words. This process enables BERT to capture the meaning and relationships between words more effectively than previous models [11].

One key aspect of BERT is its bidirectional nature, which allows it to consider both the left and right context of a word when encoding its representation. This bidirectional approach helps BERT better understand the meaning of words in a sentence and handle tasks such as sentiment analysis, named entity recognition, and question answering.

After pre-training, BERT can be fine-tuned on specific downstream tasks by adding task-specific layers on top of the pre-trained model. Fine-tuning involves training BERT on a task-specific dataset, such as sentiment analysis or text classification, to adapt it to the specific task and improve its performance.

BERT has achieved state-of-the-art results on various NLP tasks and has become widely adopted in both research and industry for its ability to understand and generate high-quality natural language representations.

3. MATERIAL AND PROPOSED METHOD

In this section, we provide an overview of the benchmark dataset used in the study. Additionally, we introduce our proposed approach to sentiment analysis, which involves integrating semantic information during the fine-tuning of the task.

3.1 Datasets

The dataset we collected consists of 3028 comments from the Google Play Store. The dataset was obtained through the scraping process using Google Play Scrap, and subsequently, the dataset will be labeled. The labeling of the dataset into positive, negative, and neutral categories refers to the process of assigning labels to data that reflect the sentiment or evaluation towards a particular entity or topic. In sentiment analysis, labeled data can be classified into these three categories:

1. Positive: Data labeled as positive indicates a sentiment or evaluation that expresses appreciation, satisfaction, or a positive viewpoint towards the observed entity or topic. Examples of using positive labels include labeling a review as "good" or "satisfied."
2. Negative: Data labeled as negative indicates a sentiment or evaluation that expresses criticism, dissatisfaction, or a negative viewpoint towards the observed entity or topic. Examples of using negative labels include labeling a review as "bad" or "unsatisfied."
3. Neutral: Data labeled as neutral indicates a sentiment or evaluation that is neutral or unclear. This data does not exhibit a clear positive or negative viewpoint toward the observed entity or topic. Examples of using neutral labels include labeling a review as "neutral" or "no specific sentiment."

The process of labeling the dataset into positive, negative, and neutral can be done manually by human annotators who categorize the data based on context and expressed sentiment. Additionally, machine learning algorithms can be used to automatically assign labels by relying on pre-trained models using human-labeled data. Labeling the dataset into positive, negative, and neutral is important in sentiment analysis to classify data and learn about user sentiment towards specific products, services, or topics.

The annotation process for the dataset was performed by a team of two human annotators. Positive labels were converted to the numerical value 1, negative labels were converted to -1, and neutral labels were converted to 0. The conversion of the initial text labels into numerical values is intended to make it easier for the machine to manage the created dataset.

3.2 Proposed Methods

In this study, we propose a sentiment analysis method based on the 7P marketing mix aspects of the Indriver application service using the BERT (Bidirectional Encoder Representations from Transformers) algorithm. The goal is to analyze user reviews of the Indriver application service on the Google Play Store and determine the sentiment associated with each aspect of the marketing mix.

After preprocessing the dataset we do the fine-tuning process, We feed the preprocessed reviews into the BERT model and train it to predict the sentiment label for each aspect of the marketing mix. The model learns to recognize patterns and associations between the textual features and the corresponding sentiment labels. By utilizing the contextual information captured by BERT, our proposed method aims to enhance the accuracy and granularity of sentiment analysis for each aspect of the marketing mix. Figure 1 illustrates the process of the BERT algorithm in classifying a comment and Figure 2 illustrates the layers involved in Sentiment Analysis.

Finally, we evaluate the performance of our proposed method using various metrics such as accuracy, precision, recall, and F1 score. We compare the results with baseline models and existing sentiment analysis approaches to validate the effectiveness of our approach in capturing the nuanced sentiments associated with different aspects of the Indriver application service.

Overall, our proposed method combines the power of the BERT algorithm with the analysis of the 7P marketing mix aspects to perform sentiment analysis on user reviews of the Indriver application service. The aim is to provide valuable insights for the company regarding customer sentiment towards specific aspects of their marketing strategy, enabling them to make data-driven decisions for improving their service and customer satisfaction levels.

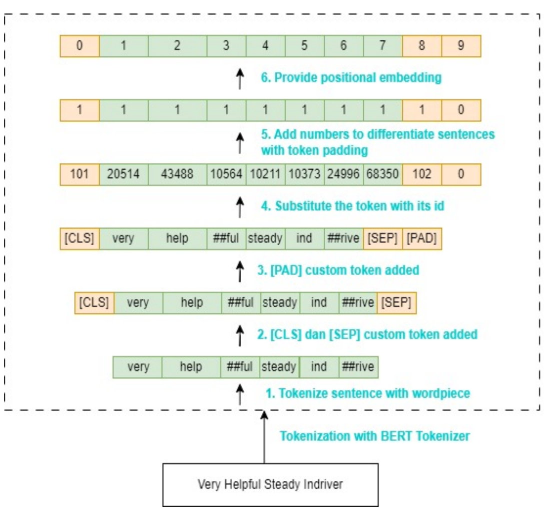

Figure 1: Illustrates The Process Of The BERT Algorithm

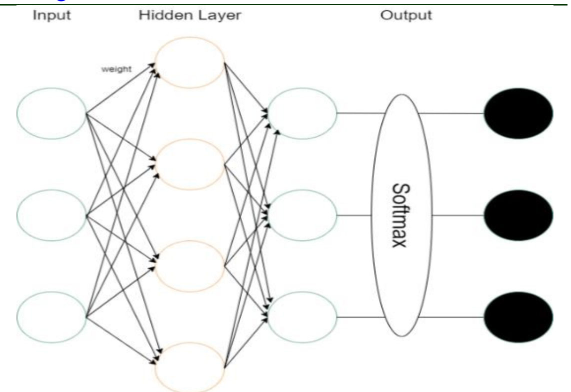

Figure 2: Illustrates The Layers Involved In Sentiment Analysis

4. EXPERIMENTAL SETUP

In the experimental setup of the study titled “Analysis Based on 7P Marketing Mix Aspects of the InDriver Application Service Using the BERT Algorithm, Based on User Reviews on the Google Play Store” several steps were followed.

4.1 Scraping Dataset

The first step in implementing web scraping was to open the InDriver application page on the official Google Play Store website. Several comments were displayed on the page. After accessing the webpage, the author performed data scraping in Google Colab by inputting the webpage URL as the data source. Once the scraping process was completed, the author downloaded the scraped data from Google Colab. A total of 3028 comments were extracted during the scraping process.

4.2 Datasets Preprocessing

To build our proposed models and make the evaluation comparisons with other techniques, the first step should involve preprocessing the utilized dataset to remove unnecessary characters from the raw text and normalizing letters that are usually written interchangeably, which can lead to data sparsity [12]. The preprocessing steps we applied include:

1. Removing URLs, mentions, retweets, and hashtag symbols,
2. Replacing underscores in hashtag texts with spaces,
3. Removing all diacritical marks and punctuation,
4. Removing repeated letters in the words,
5. Removing letter elongation in Indonesia,
6. Normalizing different forms of Indonesia letters

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.3 Split Dataset

Before performing the classification, the data was divided into three parts: training data, validation data, and testing data. The training data was used to train the model, while the validation data was used to prevent overfitting in the neural network. The testing data was used as the final evaluation to assess the accuracy of the trained network using the training data. The illustration of the dataset-splitting process can be seen in Figure 3.

```
# split dataset into training, testing, and validation
df_train, df_test = train_test_split(df, test_size=0.1)
df_val, df_test = train_test_split(df_test, test_size=0.5)
```

Figure 3: Split Dataset

4.4 Pre-trained Model

To train the model, the usage of data loaders is required, allowing iteration over each dataset. This is done to manage memory usage during the training process, avoiding the need to load the entire dataset into memory simultaneously. Additionally, blocks are created to form Data Loaders that generate tokenized comments. Both comments and sentiments have a maximum length of 168 words. For sentiment analysis, an additional layer is employed, incorporating Dropout with a probability of 0.1, following the recommendations in BERT research [13]. The authors also conducted fine-tuning using several hyperparameters, which were selected based on BERT recommendations:

1. Epoch: 10
2. Batch Size: 32
3. Learning rate: 2e-5

The authors made the selection of the above hyperparameters based on several considerations. A batch size of 32 was chosen because larger batch sizes result in a longer time to complete a single batch. This is in line with previous research [14]. Furthermore, a learning rate of 2e-5 was utilized to help BERT address the issue of catastrophic forgetting. Catastrophic forgetting refers to the problem that occurs in machine learning when a model, trained to learn one task, forgets or erases its knowledge of a previous task while learning a new task. This problem commonly arises in the context of transfer learning, where a model trained on an initial dataset is intended for use on a different task.

The authors conducted experiments using three different numbers of epochs to determine the optimal number of epochs. Based on the results of these three experiments, it can be concluded that using 10 epochs yields better performance compared to 3 epochs and 7 epochs on the dataset used in this study. Figure 4 displays the accuracy achieved with 10 epochs and Figure 5 illustrates the training process of BERT.

	precision	recall	f1-score	support
0	0.47	0.30	0.37	70
1	0.67	0.64	0.66	76
2	0.89	0.94	0.92	460
accuracy			0.83	606
macro avg	0.68	0.63	0.65	606
weighted avg	0.81	0.83	0.82	606

Figure 4: Accuracy Achieved With 10 Epochs

```
======= Epoch 1 / 3 =======
Training...
Batch    40 of     65.     Elapsed: 0:00:37
   Average training loss: 0.71
   Training epoch took: 0:01:00
Running Validation...
   Accuracy: 0.79
   Validation took: 0:00:04
======= Epoch 2 / 3 =======
Training...
Batch    40 of     65.     Elapsed: 0:00:38
   Average training loss: 0.62
   Training epoch took: 0:01:02
Running Validation...
   Accuracy: 0.79
   Validation took: 0:00:04
======= Epoch 3 / 3 =======
Training...
Batch    40 of     65.     Elapsed: 0:00:39
   Average training loss: 0.54
   Training epoch took: 0:01:03
Running Validation...
   Accuracy: 0.79
   Validation took: 0:00:04
Training complete!
```

Figure 5: Training Process Of BERT

The classification of the 7P marketing mix aspects in this research was performed by manually labeling each comment. Each aspect has its dataset, which is then used for training the BERT model. Since there are 7 aspects in this study, the BERT model is trained 7 times, resulting in 7 different models for each aspect. After training each aspect model and saving them to the drive, they are called into a main code to perform multiclass classification on the 7P marketing mix aspects. In Figure 6 an example sentence "The application is good and the fare is cheap" is classified into positive, negative, or neutral sentiment, as well as based on its aspect.

15th October 2023. Vol.101. No 19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

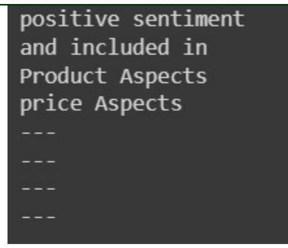

Figure 6: Example Sentence Classification

4.5 Evaluation Metrics

After testing the dataset, it was found that the overall accuracy using BERT in the first, second, and third experiments was 75%, 83%, and 83% respectively. Thus, the average accuracy of sentiment analysis with BERT is 80.3%. The results from BERT were higher than the baseline accuracy of 82%. The difference in system accuracy is influenced by the randomization of the dataset during the splitting into training, testing, and evaluation datasets in each experiment. The confusion matrix diagrams for each experiment can be seen in Figure 7, Figure 8, and Figure 9.

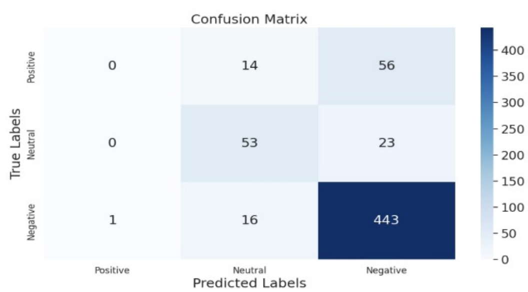

Figure 7: Confusion Matrix For The First Experiment

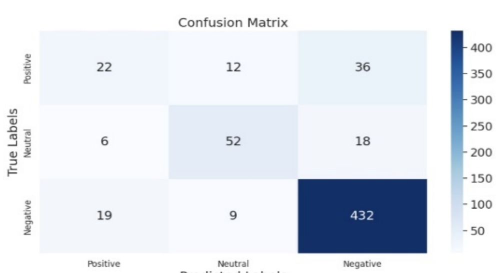

Figure 8: Confusion Matrix For The Second Experiment

Figure 9: Confusion Matrix For The Third Experiment

Here are some critically evaluated points from the research

1. Accuracy Variability in BERT Experiments: The study conducted three experiments with varying hyperparameters and observed accuracy rates of 75%, 83%, and 83%. This variability raises questions about the consistency and stability of the BERT algorithm. While hyperparameter tuning is essential, such significant fluctuations warrant a deeper investigation into the reasons behind this variation.
2. Negative Sentiment Dominance: The finding that BERT excels in classifying negative sentiment comments implies a potential bias towards this sentiment category. This prompts a concern about the generalizability of the algorithm to other sentiment types. A more balanced distribution of sentiments in the dataset could provide a clearer understanding of BERT's overall performance.
3. Dataset Quality and Accuracy Correlation: The correlation between dataset quality and accuracy is well-established, but the study doesn't delve into the specifics of dataset quality improvement. It's crucial to elaborate on the strategies employed to enhance dataset quality and how these improvements impacted accuracy.
4. Comparative Analysis with SVM: While the study asserts BERT's superiority over SVM, it would be insightful to present a detailed comparison of both algorithms. This could involve a discussion of SVM's strengths and limitations, especially in the context of sentiment analysis, to provide a more holistic perspective.
5. Hyperparameter Selection Rationale: The study provides hyperparameters used in the experiments but lacks a comprehensive rationale for these selections. Explaining the motivations behind these choices, along with the potential implications of altering them,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

could enhance the credibility of the experimentation process.

6. Dataset Representation: The study doesn't elaborate on the dataset's size, diversity, or sources. These aspects can significantly impact the generalizability of findings and the study's external validity. Providing insights into dataset composition would offer readers a clearer context.
7. Limitations and Future Directions: The text doesn't explicitly outline the limitations of the study and avenues for future research. Identifying limitations, such as potential biases or scope restrictions, would underscore the study's integrity and guide future researchers.
8. Algorithm Transparency: While BERT's effectiveness is established, the study could elaborate on how the algorithm reaches its decisions. Providing examples of misclassifications or challenging cases would offer transparency into its decision-making process.

5. CONCLUSION

Here are the conclusions based on the analysis, design, implementation, and testing results of the system:

1. In the culmination of the comprehensive analysis, systematic design, meticulous implementation, and rigorous testing conducted throughout this study, several key insights emerge regarding the effectiveness and applicability of the sentiment analysis system utilizing the Bidirectional Encoder Representations from Transformers (BERT) algorithm.
2. The experiments undertaken in sentiment analysis using the BERT algorithm have furnished significant revelations. The variation in hyperparameters, namely the batch size of 32, the learning rate of 2e-5, and 10 epochs resulted in notable differences in accuracy rates across three experiments, namely 75%, 83%, and 83%. These fluctuations underline the influence of hyperparameters on the algorithm's performance, indicative of its sensitivity to parameter fine-tuning.
3. The granular analysis of the confusion matrix corroborates the robustness of the BERT algorithm. Particularly proficient in categorizing comments imbued with negative sentiment, BERT's adeptness reaffirms its prowess in identifying and classifying prevalent sentiments within the dataset.
4. The salient observation gleaned from this study pertains to the dataset quality's pivotal role in shaping accuracy outcomes. It stands as an incontrovertible truth that an elevated dataset quality corresponds with heightened accuracy levels. This correlation underscores the significance of meticulous curation and refinement of training datasets for optimal performance of BERT.
5. Drawing a comparative assessment between BERT and Support Vector Machine (SVM), it becomes evident that BERT reigns supreme. The accuracy score of 82% achieved by SVM pales in comparison to the accuracy derived from BERT, specifically employing the multilingual-case model. This contrast substantiates BERT's ascendancy over traditional machine learning paradigms, accentuating its utility in sentiment analysis.

In conclusion, the multi-faceted findings encapsulated in this study unequivocally substantiate the BERT algorithm's efficacy and supremacy within the realm of sentiment analysis. By adeptly navigating the intricacies of hyperparameters, demonstrating prowess in sentiment classification, and showcasing its adaptability and dominance over conventional models, BERT solidifies its status as a transformative tool. As the digital landscape advances and user sentiments continue to play an instrumental role in shaping business decisions, the resounding message is that BERT stands poised as a cornerstone technology for discerning sentiment nuances and enhancing user experience.

REFERENCES

[1]. H. Chen, M. Sun, C. Tu, Y. Lin, & Z. Liu. (2016, November). Neural sentiment classification with user and product attention. In Proceedings of the 2016 conference on empirical methods in natural language processing (pp. 1650-1659).
[2]. W. Medhat, A. Hassan, & H. Korashy. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093-1113.
[3]. D. Tang, B. Qin, & T. Liu. (2016). Aspect-level sentiment classification with deep memory network. arXiv preprint arXiv:1605.08900.

[4]. J. D. M. W. C. Kenton, & L. K. Toutanova. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of naacL-HLT (Vol. 1, p. 2).

[5]. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, A. N, & I. Polosukhin. (2017). Attention is all you need. Advances in neural information processing systems, 30.

[6]. N. J. Prottasha, A.A. Sami, M. Kowsher, S. A. Murad, A. K. Bairagi, M. Masud, & M. Baz. (2022). Transfer learning for sentiment analysis using BERT-based supervised fine-tuning. Sensors, 22(11), 4157.

[7]. A. Ghallab, A. Mohsen, & Y. Ali. (2020). Arabic sentiment analysis: A systematic literature review. Applied Computational Intelligence and Soft Computing, 2020, 1-21.

[8]. K. Kanclerz, P. Miłkowski, & J. Kocoń. (2020). Cross-lingual deep neural transfer learning in sentiment analysis. Procedia Computer Science, 176, 128-137.

[9]. J. Y. Lee, & C. H. Jin. (2019). The role of ethical marketing issues in consumer-brand relationship. Sustainability, 11(23), 6536.

[10]. J. Vig, A. Madani, L.R Varshney, C. Xiong, R. Socher, & N. F. Rajani, (2020). Bertology meets biology: Interpreting attention in protein language models. arXiv preprint arXiv:2006.15222.

[11]. K. Clark, M. T. Luong, Q. V. Le, & C. D. Manning. (2020). Electra: Pre-training text encoders as discriminators rather than generators. arXiv preprint arXiv:2003.10555.

[12]. A. R. Lubis, & M. K. Nasution, (2023). Twitter Data Analysis and Text Normalization in Collecting Standard Word. Journal of Applied Engineering and Technological Science (JAETS), 4(2), 855-863.

[13]. T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, & A. M. Rush. (2020, October). Transformers: State-of-the-art natural language processing. In Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations (pp. 38-45).

[14]. J. Wu, X. Y. Chen, H. Zhang, L. D. Xiong, H. Lei, & S. H. Deng. (2019). Hyperparameter optimization for machine learning models based on Bayesian optimization. Journal of Electronic Science and Technology, 17(1), 26-40.