International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Sentiment Analysis towards Cryptocurrency and NFT in Bahasa Indonesia for Twitter Large Amount Data Using BERT

Mochammad Haldi Widianto [*1], Yhudi Cornelius [2]

Submitted: 30/10/2022 **Revised:** 18/12/2022 **Accepted:** 02/01/2023

Abstract: Sentiment analysis is one part of Natural Language Processing (NLP), where the system can understand the form of text. Sentiment Analysis itself requires an Artificial Intelligence (AI) algorithm. One of them is the Bidirectional Encoder Representation from Transformers (BERT), which can assess positive or negative sentiment, especially when discussing "Cryptocurrency" and "Non-Fungible Token (NFT)" because in recent years. The discussion on these two topics has been widely discussed. On social media such as Twitter. In this study, a BERT model was created to assess sentiment analysis on "Cryptocurrency" and "NFT", by utilizing data crawling and pre-processing using Rapidminer (student version) with 86% accuracy and 87% precision. The results were obtained by selecting suitable hyperparameters such as learning rate 1e-7, epoch 5, and batch size 32. The results also increased from 2% accuracy to 3% precision. Further research needs to be done to improve the BERT model, not just an increase in the pre-processing part. In further research, the authors suggest combining several models or updating the pre-processing.

Keywords*:* Bidirectional Encoder Representation from Transformers (BERT), Cryptocurrency, Non-Fungible Token (NFT), Sentiment Analysis

1. Introduction

Cryptocurrency is being discussed a lot because it is one of the world's most famous types of finance, which still has several types of money type of risks. Since the beginning of currency development, Cryptocurrencies have affected several businesses, especially those related to finance in some broad cases.[1]. This currency can be used for exchanges and transactions by leveraging algorithms and complex cryptography to secure the network. It is undeniable that Cryptocurrencies are indeed related to the workings of the blockchain framework and take advantage of how the properties of blockchain are decentralized and transparent. However, central interests could not be controlled in a past cryptocurrency system. This way of working actually has a function in validating algorithms on blockchain in solving trust issues in all aspects of interest [2].

After a lot of trouble, the global economy wants to move towards an age where all transactions can be leveraged digitally with cryptocurrencies. Then applied to the entire community in digital form. In the future, the application of financial industry by utilizing digital payments through cryptocurrencies. The application of cryptocurrency as a medium of exchange is formed to exchange digital information. Cryptocurrency is now widely used because of its convenience of use and the security it provides. This happens Because there is no third-party involvement [3], [4]. In addition to cryptography, other technologies have been introduced, such as Non-Fungible Tokens (NFT), because they have recently received great attention in digital technology [5]..

Recently, user interviews have been conducted regarding opinions and suggestions regarding the applicability of cryptocurrencies and NFTs. For researchers to understand users' opinions, attention needs to be paid, especially in Indonesia, because it can help the economy [6]. This needs to be identified whether the community has a negative or positive opinion. A larger user study is needed from social media. Big data collection is a challenge because traditional databases can not process the results and give slow results. Sentiment Analysis (SA) is one of the mathematical methods for categorizing statements (positive or negative) to analyze text-on-language data in several other types of communication, not only tweets [7]. Sentiment analysis is one way to categorize the opinions of users, audiences, services, products or customers, products, etc., which categorize them into negative, neutral, or positive relationships. This is usually the search for suggestions or opinions in identifying the speaker's opinion or attitude [7]. The purpose of the SA to look for common use cases for this technology is to ascertain and identify how people feel about certain topics.

[1,2] Computer Science Department, School of Computer Science, Bina Nusantara University, Bandung Campus, Jakarta, Indonesia, 11480
ORCID ID https://orcid.org/0000-0001-8722-9868

** Corresponding Author Email: mochamad.widianto@binus.ac.id*

Several studies are used in conducting SA, especially those utilizing Bidirectional Encoder Representation from Transformers (BERT) [8], [9] because transformer research is a new approach, so many use it as a means of determining positive or negative sentiment. Of course, by utilizing the training data that has been processed. In this study, the author utilizes the BERT in conducting (SA) assisted by Rapidminer (student version), which is used in cleaning, pre-processing, and initial processing. The application of SA has been studied in detail. It has its foundation in several kinds of literature: such as carrying out the proposed approach to sentiment analysis of data. All of them are collected on social media Twitter. [10] or also use other sources [11]. In previous work [12], his research discusses the opportunity of sentiment analysis, especially detecting sarcasm. The research [13] detects sarcastic undertones in sentiment analysis. In another study using BERT [14], this research has the weakness of a small amount of data and a pre-processing process that is less detailed. Improvements in data processing are needed so that the accuracy used by BERT increases. The ability to use RapidMiner (Student Version) needs to be improved so that the data used can be cleaner and there are no duplicates.

So in this study, processing a lot of data on Twitter using Rapidminer (student version) with a more specific data cleaning process and selecting a more accurate learning rate. This study also aims to look at the results of sentiment on Twitter as a recommendation on whether crypto and NFT have a positive or negative impact.

2. THEORY

2.1 Sentiment Analysis

SA cannot be separated where one of them is the process of how to see emotional expressions through language, which includes seeing the results of information originating from emotions, processing, and results from data analysis, and then classifying the results of the analysis selected text [15], [16]. Looking at the function, Sentiment analysis is a way to recognize emotions, affective computing, and polarity detection [17]. Text sentiment analysis is an essential technique in learning to perform natural language processing in attracting emotional factors, which is usually applied to crowd opinion monitoring, business intelligence, and AI [18]. There are all kinds of approaches to conducting text sentiment analysis: machine learning, hybrid-based Dictionary [19], [20]

Lately, social networking sites are widely used by the public, so it produces a lot of academic focus, usually used in the analysis of networks and their contents in search of the results of the necessary data filters. Sentiment analysis is related to finding the results of the sentiment communicated by a work that comes from its content. This NLP technology can store long memories in looking at viral public opinion for decision-making, but much preliminary work is needed to address this [21]

SA is very important because it is part of the development of the science of text mining which aims to get results from text that has been processed. This explanation needs to be classified into negative and positive or neutral classes. Several researchers, academics, and industry have used sentiment analysis on several media to obtain datasets from social media [22]. Sentiment analysis adapted from [23]–[25] and presenting information in a holistic approach for sentiment analysis because it is very urgent and does not only talk about classification or categories.

2.2. BERT

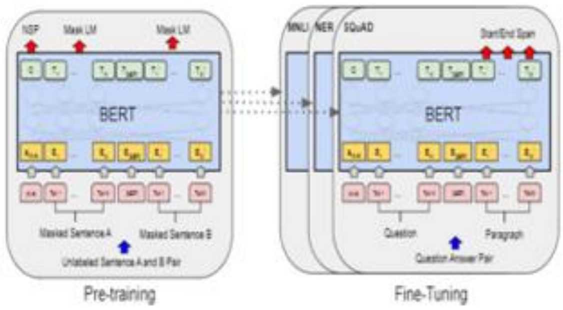

Fig 1. Pre-Training and Fine-Tuning BERT model [26].

BERT is used in training or designing representations in different directions. In several studies that carried out the formation of the BERT model, it was necessary to configure one additional output layer to create the latest model and be able to do many tasks. The BERT model can perform unsupervised and supervised (unlabeled) data classification. Fig. 1 presents the essence of the procedures in fine-tuning and pre-training in the BERT procedure [26].

Figure 1 above describes the stages in BERT. This method usually has two phases: pre-training and refinement. In phase one pre-training, the model was trained on unlabeled data. It is important that the learned part of the model is initially initialized and fine-tuned using the components that were trained in the previous area. After that, set the use of labeled information/data from the beginning of the process. Even though the model starts with the same pre-training parameters, each downstream task has its own customized model.

BERT models typically utilize a multi-layer bidirectional Transformer encoder based on the original application described in [27]. Because the use of Transformers has been implemented almost indicatively and generally, it has become a complete background of the model architecture and refers the reader to [27].
Study [26] This Model is designed based on the encoder architecture of Transformers such as Sentiment Analysis [28], and Text Summarization [29]. Study [26] This study

has produced two types of BERT models, namely: $BERT_{BASE}$ and $BERT_{LARGE}$.

For the Transformers architecture, the number of layers in the encoder is 6 layers [30]. Each layer, as in the encoder, consists of two sub-layers, multi-head attention and feed-forward, which have a normalization layer [30]. For the structure of the multi-head attention layer. For each word that is entered into the layer, it will be converted into Q, K, and V in the form of a vector where Q (query) is the word entered, K (key) is the keyword or meaning of the word that has been entered, and V (value) is the purpose or intent of the word that has been entered [30] The equation for multi-head attention can be seen in equation (1) and equation (2) [30]

$$MultiHead\ (Q, K, V) = Concat\ (head_i, \dots, head_n)W^O \quad (1)$$

$$head_i = Attention\ (QW_i^Q, KW_i^K, VW_i^V) \quad (2)$$

Where W (weighted) is the parameter matrices(3)

$$W_i^Q \in \mathbb{R}^{d_{model} \times d_k}, W_i^K \in \mathbb{R}^{d_{model} \times d_k}, W_i^K \in \mathbb{R}^{d_{model} \times d_k}, W_i^V \in \mathbb{R}^{d_{model} \times d_v}, and\ W^O \in \mathbb{R}^{hd_v \times d_{model}} \quad (3)$$

Variable h sum is the parallel attention layer, and variable d is the dimension value of each parameter Q, K, V, and model.

The output of each sub-layer is the normalization layer [30]. The equations are contained in the normalization layer in equation (4) [30].

$$LayerNorm\ \left(x + Sublayer\ (x)\right) \quad (4)$$

The variable x is the word embedding, and the Sublayer (x) is the output of the multi-head attention layer [30].

2.3. Comparison Other BERT Model

This study aims to improve the sentiment analysis model as shown in table 1 below:

Table 1. Comparison Method

ref	Sentiment Analysis Model	Sentiment in Indonesia Language	Deep Learning	Large Data
[31]	Sastrawi	Yes	No	No
[32]	Sastrawi	Yes	No	No
[14]	BERT	Yes	Yes	No
Proposed Method	BERT	Yes	Yes	Yes

Table 1 attached what differences will be made in the study compared to other research types so that in this study's improvement, large data will affect accuracy, confidence matrix, and precision. So the application of pre-processing and selection of hyperparameters is essential. The design model will be explained in further Research Units

3. System Design

In this section, the author will explain the design section used in conducting sentiment analysis on NFT and Crypto. So in this section, the author will divide it into several stages, first Crawling and Pre-Processing, second BERT Model, third hyperparameter, and fourth Evaluation

3.1. Crawling and Pre-Processing

large data, in this study, pre-processing was made using Rapidminer (Student Version). So the author can see the schematic in figure 2:

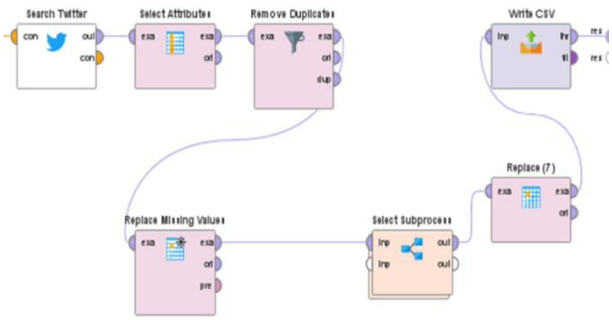

Fig 2. Pre-Processing Proposed Method

As shown in Figure 2, the pre-processing stage is carried out starting from several stages, especially from:

1. Twitter Search: at this stage a tweet search (crwaling Twitter Data) is carried out using the Query Language "Crypto" and "NFT" especially using Indonesian
2. Select Attributes: looking for attributes to be continued, at this stage only data in the form of "text" and "From-User" are passed to enter the next stage
3. Remove Duplicates: At this stage, deletion is carried out when there is data that is identified as having the same contents. Its function is so that the data is data that does not have the same content
4. Select Subprocess: At this stage, the letters that are commonly referred to as normalization contain meanings such as (!@#$%^&*()_+{}|”:>?<)
5. Replace: At this stage, the words (http and after / https and after) are replaced with spaces, thus eliminating existing links
6. Write CSV: At the end of this stage, the data results are saved in CSV format

The difference in previous studies [14] is that there is no Replace stage, so the resulting data is not completely clean. Expected to contribute if using large amounts of data. After the above steps are carried out, the BERT model fine-tuning is processed.

3.2. BERT Design

To handle large data, at this stage, the use of the BERT Model is carried out by referring to Figure 3 below:

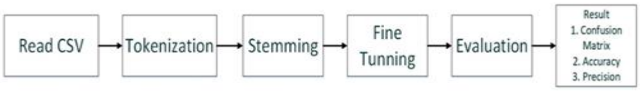

Fig 3. BERT Design

Figure 3 shows the BERT Design where after getting the data that has been processed, several stages are carried out in the BERT Design, such as:

1. Read CSV: This is a process where data is read and comes from pre-processing
2. Tokenization: At this stage, sentence solving is looking for word pieces
3. Stemming: It is the process of finding basic words
4. Fine Tuning: The process is carried out in BERT, and there is a selection of other hyperparameters
5. Evaluation: This stage is an evaluation stage that compares the results of testing and training
6. Result: At this stage, the evaluation results are mapped in the form of a confusion matrix, accuracy, and precision

3.3 Hyperparameter

To overcome a large amount of data, it is also necessary to set the correct hyperparameters, such as One of the parameters, namely, the learning rate, which is widely utilized in AI, especially (deep learning systems) and is also used in this BERT algorithm. The learning rate value usually ranges from one (1) to zero (0) [33]. Another thing is also essential in finding this value because the speed of learning is very much aligned with the amount of input data [34]

Other parameters, such as Epoch, are used to repeat the AI learning process. The biggest the repetition, the longer the processing time. The accuracy and precision are better [35]. If you look at the batch size, this parameter is used in deep learning, which is usually used for the amount of data / images that can be used, especially in network or model training [36]

3.4 Evaluation

For the results and measure performance, a confusion matrix is used to evaluate the trials. In this study, the function of the confusion matrix is to see the binding classification performance by looking at the test data. An example is attached to table 2 [37]..

Table 2. Confusion Matrix [37].

Actual class	***Assigned class***	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Furthermore, this study used several formulas to look at precision, F1 scores and accuracy, and other indicators in conducting assessments to ensure the model [38]. The formula in (5)-(7):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{F1 Score} = 2x \frac{Precision\ x\ Recall}{Precision + Recall} \quad (6)$$

4. Result

In this section, some results of crawling and testing are presented. Below is the section that will be shown:

a. Crawling data results
b. Proposed Method for BERT Accuracy
c. Proposed Method for BERT Precision
d. Comparison Accuracy and Precision

4.1 Crawling Data Results

In the session the author will be on the results of "Cryptocurrency" and "NFT" through RapidMiner (Student Version). The author also does pre-process with some conditions to be removed, such as commas (@,(,), http, RT, etc.). Then if there is a punctuation mark it will be removed (only removing the punctuation mark (,) does not change the word or sentence). (Disclaimer = statements used in this study (see Table 3) obtained from publicly posted tweets and related news).

Table 3. Tweet Result

No	***Account***	***Tweet***	***Sentiment***
1	Monke[redacted]	*mata uang kripto adalah 306sset digital*	Positive
2	Lip[redacted]	*seberapa amankah mata uang kripto dan bagaimana nasib 306sset pengguna jika dicuri*	Negative

Table 3 is one of the results of the tweets obtained and shows the results of tweets with tweets originating from twitter (not essays), derived from the words "Cryptocurrency" and "NFT" data-preprocessing shows tweets that are very clean because the process, after This is the next most important stage is the stage of entering the results into the BERT Model.

The data in SA processing that is utilized amounts to 4000 Twitters, which results will be presented to show if the opinion of each tweet contains negative or positive sentiments. (90% Training 10% Test data).

4.2 BERT Accuracy

In this study, several experiments were carried out to measure the results of sentiment on the words "Cryptocurrency" and "NFT" using hyperparameters in the form of learning rates in the range of 1e-5, 1e-6, and 1e-7, and utilizing various epochs along with Batch size of 32 as shown in Table 4.

Table 4. Result Accuracy from the Study

Using 32 Batch Size			
	Epoch 2	Epoch 3	Epoch 5
Learning Rate 1e-5	51%	53%	53%
Learning Rate 1e-6	80%	83%	85%
Learning Rate 1e-7	80%	85%	86%

Table 4 shows that the large change in epoch affects the accuracy percentage. The most important thing is that it explain the effect of learning rate on accuracy, which the best result of accuracy is 86%, with Learning Rate at 1e-7. This is because large data requires a smaller learning rate to get a suitable learning rate.

4.3 BERT Precision

By considering accuracy alone is not enough, this study conducted a trial to measure precision on words such as "Cryptocurrency" and "NFT" as in table 5.

Table 5. Result from Average Precision from Study

Using 32 Batch Size			
	Epoch 2	*Epoch 3*	*Epoch 5*
Learning Rate 1e-5	51%	53%	53%
Learning Rate 1e-6	80%	83%	85%
Learning Rate 1e-7	80%	86%	87%

The table shows that the best precision is at 87% by utilizing a epoch 5 and a learning rate of 1e-7 . This shows that the smaller the learning rate affects the accuracy of the larger data. Then, the larger number of epochs helps in the process of increasing precision. Here the researcher finds data processing is very difficult using BERT and needs to be considered in pre-processing and modeling.

4.4 Comparison with Other BERT Models

In this section the author will explain the comparison between the previous model and the one used by the researcher, shown in the following figure 5

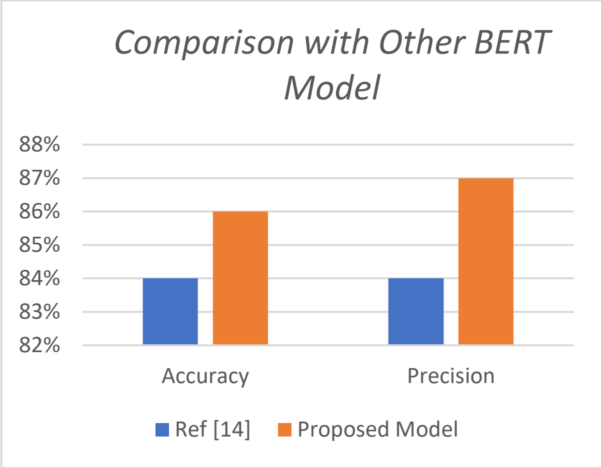

Fig 5. Result Comparison

Figure 5 shows the comparison results of previous studies by increasing the accuracy and precision parts, this happens because the pre-processing model used in this study can improve accuracy and precision. On the other hand, the right amount of hyperparameter selection can improve accuracy and precision. Accuracy is increased by about 2%, and precision is increased by 3%.

5. Conclusion

This study focuses on sentiment analysis of words such as "Crypto" and "NFT" with a very large amount of 4000 data. This study focuses on modeling using BERT and pre-processing data using Rapidminer (student version). The results show that the accuracy is 86%, and the precision is 87%. The results utilize a epoch 5 and learning rate of 1e-7. Compared with previous studies, the authors managed to make a 2% increase in accuracy and 3% in precision. According to the next researcher, it will be continued with the BERT model, which needs to be modified. Back in the face of Big Data, not just a lot of data.

References

[1] Z. Shahbazi and Y. C. Byun, "Machine Learning-Based Analysis of Cryptocurrency Market Financial Risk Management," *IEEE Access*, vol. 10, pp. 37848–37856, 2022, doi: 10.1109/ACCESS.2022.3162858.

[2] Z. Shahbazi and Y. C. Byun, "Improving the cryptocurrency price prediction performance based on reinforcement learning," *IEEE Access*, vol. 9, pp. 162651–162659, 2021, doi: 10.1109/ACCESS.2021.3133937.

[3] A. Ghosh, S. Gupta, A. Dua, and N. Kumar, "Security of Cryptocurrencies in blockchain technology: State-of-art, challenges and future prospects," *Journal of Network and Computer Applications*, vol. 163, p. 102635, 2020, doi: https://doi.org/10.1016/j.jnca.2020.102635.

[4] A. Narayanan, J. Bonneau, E. Felten, A. Miller, and S. Goldfeder, *Bitcoin and cryptocurrency technologies: a comprehensive introduction*. Princeton University Press, 2016.

[5] H. R. Hasan *et al.*, "Incorporating Registration, Reputation, and Incentivization Into the NFT Ecosystem," *IEEE Access*, vol. 10, pp. 76416–76433, 2022, doi: 10.1109/ACCESS.2022.3192388.

[6] G. Y. Pratomo, "Kripto Lokal Berpotensi Dongrak Ekonomi Digital Indonesia," Feb. 2022. https://www.liputan6.com/crypto/read/4897884/kripto-lokal-berpotensi-dongrak-ekonomi-digital-indonesia (accessed Jun. 11, 2022).

[7] A. Dwivedi, R. P. Pant, S. Pandey, and K. Kumar, "Internet of Things' (IoT's) Impact on Decision Oriented Applications of Big Data Sentiment Analysis," in *2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU)*, 2018, pp. 1–10. doi: 10.1109/IOT-SIU.2018.8519922.

[8] Y. Zhuang and J. Kim, "A bert-based multi-criteria recommender system for hotel promotion management," *Sustainability (Switzerland)*, vol. 13, no. 14, Jul. 2021, doi: 10.3390/su13148039.

[9] B. Muller, B. Sagot, D. Seddah, and B. Muller Benoˆıt Benoˆıt Sagot Djamé Seddah, "Enhancing BERT for Lexical Normalization," in *The 5th Workshop on Noisy User-generated Text (WNUT)*, 2019. [Online]. Available: https://gitlab.inria.fr/bemuller/bert

[10] E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Ureña-López, and A. R. Montejo-Ráez, "Sentiment analysis in Twitter," *Nat Lang Eng*, vol. 20, no. 1, pp. 1–28, Jan. 2014, doi: 10.1017/S1351324912000332.

[11] H. M. Zin, N. Mustapha, M. A. A. Murad, and N. M. Sharef, "Term Weighting Scheme Effect in Sentiment Analysis of Online Movie Reviews," *Adv Sci Lett*, vol. 24, no. 2, pp. 933–937, Mar. 2018, doi: 10.1166/asl.2018.10661.

[12] M. Bouazizi and T. Otsuki, "A Study on Text Mining on Twitter: Identifying Opinion and Detecting Different Forms of Speech Using Writing Patterns," Tokyo, 2019.

[13] L. H. Son, A. Kumar, S. R. Sangwan, A. Arora, A. Nayyar, and M. Abdel-Basset, "Sarcasm detection using soft attention-based bidirectional long short-term memory model with convolution network," *IEEE Access*, vol. 7, pp. 23319–23328, 2019, doi: 10.1109/ACCESS.2019.2899260.

[14] M. H. Widianto, Y. Heryadi, W. Suparta, and A. Wibowo, "Sentiment Analysis of Cooking Oil using Bidirectional Encoder Representations from Transformers," in *2022 5th International Conference on Information and Communications Technology (ICOIACT)*, 2022, pp. 110–115. doi: 10.1109/ICOIACT55506.2022.9971861.

[15] M. v Mäntylä, D. Graziotin, and M. Kuutila, "The evolution of sentiment analysis—A review of research topics, venues, and top cited papers," *Comput Sci Rev*, vol. 27, pp. 16–32, 2018, doi: https://doi.org/10.1016/j.cosrev.2017.10.002.

[16] Z. Li, R. Li, and G. Jin, "Sentiment Analysis of Danmaku Videos Based on Naïve Bayes and Sentiment Dictionary," *IEEE Access*, vol. 8, pp. 75073–75084, 2020, doi: 10.1109/ACCESS.2020.2986582.

[17] K. Bloom, N. Garg, and S. Argamon, "Extracting Appraisal Expressions," in *HLT/NAACL*, Aug. 2007, pp. 308–315.

[18] E. Cambria, "Affective Computing and Sentiment Analysis," *IEEE Intell Syst*, vol. 31, no. 2, pp. 102–107, Mar. 2016, doi: 10.1109/MIS.2016.31.

[19] S. M. Liu and J.-H. Chen, "A multi-label classification based approach for sentiment classification," *Expert Syst Appl*, vol. 42, no. 3, pp. 1083–1093, 2015, doi: https://doi.org/10.1016/j.eswa.2014.08.036.

[20] S. Sun, C. Luo, and J. Chen, "A review of natural language processing techniques for opinion mining systems," *Information Fusion*, vol. 36, pp. 10–25, 2017, doi: https://doi.org/10.1016/j.inffus.2016.10.004.

[21] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif Intell Rev*, 2022, doi: 10.1007/s10462-022-10144-1.

[22] A. A. Lutfi, A. E. Permanasari, and S. Fauziati, "Sentiment Analysis in the Sales Review of Indonesian Marketplace by Utilizing Support Vector Machine," *Journal of Information Systems Engineering and Business Intelligence*, vol. 4, no. 1, pp. 57–64, Apr. 2018, doi: 10.20473/jisebi.4.1.57-64.

[23] A. Kumar and A. Jaiswal, "Systematic literature review of sentiment analysis on Twitter using soft computing techniques," *Concurr Comput*, vol. 32, no. 1, p. e5107, 2020, doi: https://doi.org/10.1002/cpe.5107.

[24] E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, "Sentiment Analysis Is a Big Suitcase," *IEEE Intell Syst*, vol. 32, no. 6, pp. 74–80, Nov. 2017, doi: 10.1109/MIS.2017.4531228.

[25] A. Ligthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: a tertiary study," *Artif Intell Rev*, vol. 54, no. 7, pp. 4997–5053, 2021, doi: 10.1007/s10462-021-09973-3.

[26] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

[27] A. Vaswani *et al.*, "Attention is All you Need," in *Advances in Neural Information Processing Systems*, 2017, vol. 30. [Online]. Available: https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[28] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning," *IEEE Access*, vol. 8, pp. 23522–23530, 2020, doi: 10.1109/ACCESS.2020.2969854.

[29] A. P. Widyassari, A. Affandy, E. Noersasongko, A. Z. Fanani, A. Syukur, and R. S. Basuki, "Literature Review of Automatic Text Summarization: Research

Trend, Dataset and Method," in *2019 International Conference on Information and Communications Technology (ICOIACT)*, 2019, pp. 491–496. doi: 10.1109/ICOIACT46704.2019.8938454.

[30] A. Vaswani *et al.*, "Attention Is All You Need," *31st Conference on Neural Information Processing Systems*, pp. 1–11, 2017, doi: 10.1109/2943.974352.

[31] B. Siswanto, "Sentiment Analysis in Indonesian on Jakarta Culinary as A Recommender System," in *2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Dec. 2021, pp. 46–50. doi: 10.1109/ISRITI54043.2021.9702772.

[32] B. Siswanto, F. L. Gaol, B. Soewito, and H. L. H. S. Warnars, "Sentiment Analysis of Big Cities on The Island of Java in Indonesia from Twitter Data as A Recommender System," in *2021 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS*, Oct. 2021, pp. 124–128. doi: 10.1109/ICIMCIS53775.2021.9699147.

[33] Y. Lu, "Deep neural networks and fraud detection," 2017.

[34] A. S. A. Sukor, A. Zakaria, N. A. Rahim, L. M. Kamarudin, and H. Nishizaki, "Abnormality detection approach using deep learning models in smart home environments," in *ACM International Conference Proceeding Series*, Apr. 2019, pp. 22–27. doi: 10.1145/3330180.3330185.

[35] W. Hastomo, A. S. B. Karno, N. Kalbuana, A. Meiriki, and Sutarno, "Characteristic Parameters of Epoch Deep Learning to Predict Covid-19 Data in Indonesia," *J Phys Conf Ser*, vol. 1933, no. 1, p. 12050, Jun. 2021, doi: 10.1088/1742-6596/1933/1/012050.

[36] I. Kandel and M. Castelli, "The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset," *ICT Express*, vol. 6, no. 4, pp. 312–315, 2020, doi: 10.1016/j.icte.2020.04.010.

[37] K. M. Ting, "Confusion Matrix," in *Encyclopedia of Machine Learning and Data Mining*, G. I. Sammut Claude and Webb, Ed. Boston, MA: Springer US, 2017, p. 260. doi: 10.1007/978-1-4899-7687-1_50.

[38] M. Kang, K. H. Lee, and Y. Lee, "Filtered bert: Similarity filter-based augmentation with bidirectional transfer learning for protected health information prediction in clinical documents," *Applied Sciences*, vol. 11, no. 8, 2021, doi: 10.3390/app11083668.