ICIC Express Letters ICIC International ©2023 ISSN 1881-803X
Volume **17**, Number **1**, January **2023** pp. **123–130**

UTILIZING INDOBERT IN PREDICTING PERSONALITY FROM TWITTER POSTS USING BAHASA INDONESIA

Kelvin[1], Ivan Sebastian Edbert[2] and Derwin Suhartono[2]

[1]Computer Science Department, BINUS Graduate Program – Master of Computer Science
[2]Computer Science Department, School of Computer Science
Bina Nusantara University
JL. K. H. Syahdan No. 9, Kemanggisan, Palmerah, Jakarta 11480, Indonesia
{ kelvin007; ivan.edbert }@binus.ac.id; dsuhartono@binus.edu

Received January 2022; accepted April 2022

Abstract. *Personality prediction is a process of identifying personality from a certain individual in various circumstances. It has been proven to play an important role in the hiring process of an employee. Previous studies have conducted experiments to combine different algorithms or pre-trained models, but no one attempted to conduct research using Indonesian-based pre-trained models. This study aims to contribute to better performance of personality prediction model for Bahasa Indonesia text on social media by utilizing IndoBERT, a pre-trained language model based on Bidirectional Encoder Representations from Transformers (BERT) architecture for Bahasa Indonesia. Through this study, a deep learning model using a combination of BERT and IndoBERT is proposed. The model is trained using Indonesian labeled tweets with Big Five Personality traits and tested by three different scenarios. The final evaluation of the model achieves the performance of 76.06% F1 score and 78.35% average accuracy. It is managed to exceed the performance results of previous studies.*
Keywords: Personality prediction, Social media, Deep learning, Bahasa Indonesia, BERT, IndoBERT

1. **Introduction.** Social media is a popular form of online communication, especially in Indonesia. Studies show that users in Indonesia reach 29 million on Twitter. Jakarta, the capital of Indonesia, accounted for 2.4% of the total tweets made between January and March 2014 [1]. The huge number of social media users gave more opportunities to conduct behavior analysis of a person's post on social, for example, to identify the personalities of individuals by observing their actions and behaviors in different situations [2].

In 2016, research about personality prediction has proven that the Big Five Personality is predictive [3] and has more validity and reliability [4] compared to Dominance, Influence, Steadiness, and Compliance (DISC) assessment, a behavior self-assessment tool made for predicting job performance, and Myers-Briggs Type Indicator (MBTI), an introspective self-report questionnaire indicating differing psychological preferences in how to perceive the world and make decisions. The comparison between these tools is shown in Table 1. Big Five Personality is also proven to be successful in predicting behaviors consistently in the workplace [5].

Conscientiousness contributes 9% in increasing validity, while Openness contributes to the addition of 6% validity [3]. Even though personality prediction plays an important role, the process of personality analysis from prospective employees requires a lot of time and resources. An algorithm can reduce resources needed and should be included in the hiring process. Personality assessment is a powerful tool in eliminating what humans must do repeatedly and are prone to error. Valid and reliable personality assessments provide recruiters with data to make the right decisions.

DOI: 10.24507/icicel.17.01.123

Table 1. Comparison of personality measuring tools

	Big Five Personality	DISC	MBTI
Measured traits	30	4	4
Result uniqueness	∼	12 profiles	16 personalities
Valid	Yes	Yes	No
Reliable	Yes	Yes	No
Predictive	Yes	No	No

Many studies were conducted for predicting personality by using various algorithms or pre-trained models, but no one attempted to conduct research using Indonesian-based pre-trained models. This study aims to contribute to the personality prediction model performance for Indonesian text on social media by utilizing IndoBERT, a pre-trained language model based on BERT architecture for Bahasa Indonesia. We choose the personality traits based on Big Five Personality where the approach represents the structure of human nature. This approach is described with 5 basic factors on most traits in a person's personality: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism [6]. In this study, the researchers will classify using the IndoNLU model that provides a dataset for training, evaluation, and benchmarking for Natural Language Processing (NLP) [7]. IndoNLU provides a model, namely IndoBERT which is based on the BERT model. The proposed model is a combination of two pre-trained models, BERT and IndoBERT, which learn from Indonesian Tweets data. This was inspired by the previous study about concatenation BERT with other pre-trained language models to enhance the performance of the classification model [8]. The model is then compared to several scenarios to evaluate the effectiveness of its predicting capabilities.

2. **Related Works.** Personality prediction has become more advanced along with the development of the natural language processing field, and other fields. For example, facial image recognition uses computer vision, Convolutional Neural Network (CNN) audio, CNN text, and CNN video to analyze the Big Five Personality Traits and resulting 0.0938 Root Mean Square Error (RMSE) [9]. Several machine learning models are developed to solve personality prediction problems, studies attempted to build automated personality prediction in Bahasa Indonesia by using K-Nearest Neighbors (KNN) and Support Vector Machines (SVM) and performed 65% accuracy for the combined model [10]. After that automated personality prediction was developed using eXtreme Gradient Boosting (XGBoost) classifier and achieved 99% precision and accuracy which is higher than the previous KNN and SVM model. However, this textual data is using the English language and done in MBTI [11].

In 2017, personality classification from online text is done by using machine learning technique, XGBoost ensemble model is used to, which has achieved 97% performance by using tweets data in Bahasa Indonesia [12]. This study also provides the Twitter dataset for personality prediction in Bahasa Indonesia.

After machine learning, other techniques are being introduced, such as deep learning that allows multiple processing of layers [13]. Then come study using deep learning models like Long Short-Term Memory (LSTM), CNN, and another deep learning model [14]. Long short-term memory is introduced in 1997 to solve a very long-time *backpropagation* caused by decaying error backflow [15]. LSTM then manages to cover the drawback that the base RNN model has, but it still has drawbacks. It can only see the context from the left to the right perspective. Then come to Bidirectional LSTM, which can see the context from both directions and improve base LSTM model that can only see from left to right. After that, the study proposed BiLSTM, and experiment results show additional training of data and thus BiLSTM-based model offers better predictions than regular LSTM-based models [16].

The topic of this research is also carried out in several languages and sources of data. The proposed model is used to process Facebook user data from the *myPersonality* dataset that has been done before and achieve mean absolute errors of 0.45888 [17].

The field of automatic personality recognition in specific language is also done by using Philippines *Twitter* user datasets and focusing only on Openness and Conscientiousness, this research is done by exploring different multilayer perceptron configuration fed and usage of the pre-trained word from English and Tagalog and resulting root mean square error 0.3344 for Conscientiousness and 0.3419 for Openness [18].

Another study also compares TF vector-based, ontology-based, enriched ontology-based, LSA-based, and deep learning-based methods along with the ensemble modeling method. As the result, ensemble modeling achieves the highest accuracy [19]. However, in the same year, another research was also conducted using word embedding as input for the neural network used, namely LSTM, BiLSTM, and Gated Recurrent Unit (GRU). By using the F-Measure, the resulting number is 0.82812 on average training score using GRU with Rectified Linear Unit (ReLU) as the activation function and achieves a better result [20].

Following with rapid development of attention-based models, BERT is introduced in 2018. Its main innovation is bidirectional representation, where the previous models like OpenAI Generative Pre-Trained Transformer and Embeddings from Language Models (ELMo) only implement one-sided representation [21]. Recently, a new BERT model is trained to specifically handle NLP tasks in Bahasa Indonesia called IndoBERT, and the research also provides a large dataset and corpus [7].

Recent study in personality prediction was done using BERT deep learning architecture with a combination of multiple pre-trained models such as BERT, RoBERTa, and XLNet as features extraction method from social media data, and the system used model averaging to make its prediction [22]. Most of the time, the model in the study is a multilingual model, type of machine learning model that can understand different languages, that is the main reason IndoBERT model still has a lot of potential that might be obtained in the development of personality prediction tools.

3. **Research Methodology.** This study focuses on the application of IndoBERT on the Transformer architectural design to predict each trait. The data used are tweet data contained on the profile page of Indonesian Twitter users.

3.1. **Data preparation and processing.** The dataset used was obtained from the previous study [12] which was gathered from 508 Twitter users, containing their tweets and profile data, then annotated with five personality models by a psychology expert. The annotation consists of two labels, which are "high" and "low".

Several methods are used to remove noise from the data which are

1) Mentions were replaced to [UNAME] and hashtags were replaced to [HASHTAG].
2) Remove URL.
3) Retweets omitting.
4) Remove Emoji.

The reason why we propose this method is that attributes such as mentions and hashtags do not provide significant information about the words contained in them, but rather on their usage so we change it to a token that is simpler and can be recognized during the training stage, while other attributes such as URL and emojis are difficult for tokenizers to understand so they are removed from the text. Along with that, we also filter out tweets that were not generated by the user himself, using retweets by simple not including retweeted text. In this pre-processing, traditional processes such as removing stop words are not performed. Because BERT can learn by itself that stop words do not have much effect and reduce the weight of attention on these words [23]. The data used to test the performance of the classification is obtained by dividing the data by a ratio of 9 : 1 between the training data and the test data using stratified random sampling. This ratio is proposed to avoid overfitting or underfitting.

3.2. **Proposed model architecture.** This study was conducted by using 2 kinds of pre-trained models, BERT and IndoBERT. In the previous study, combination is done by using stacking ensemble method between different multilingual model and algorithms, but in this proposed model, we use a combination of two different BERT models, and we call it dual BERT, which consists of BERT and IndoBERT pre-trained models. They are trained using the same inputs. Having a combination of both models will preserve the ability of IndoBERT to classify text in Bahasa Indonesia. While implementing BERT's ability to recognize English text. The BERT model used in this study is a pre-trained multilingual uncased BERT Base. For IndoBERT, IndoBERT Base phase1 is used, which is also an uncased model. The proposed architecture is illustrated in Figure 1.

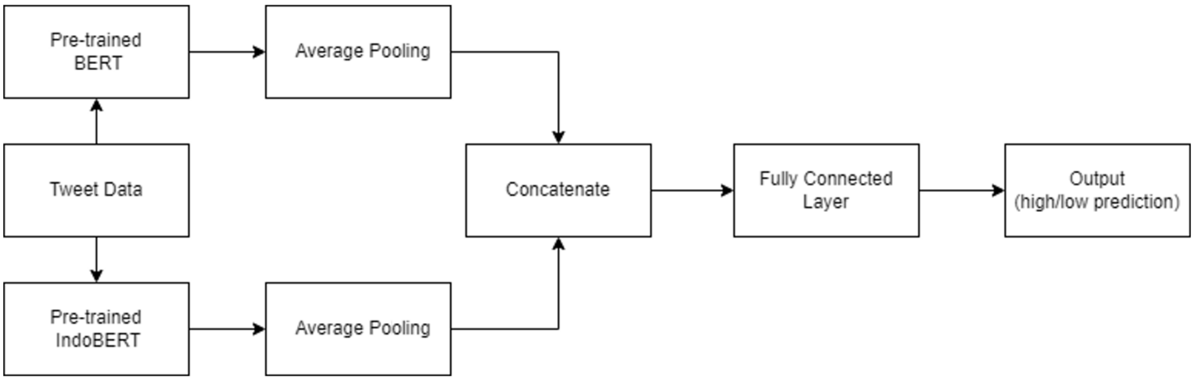

FIGURE 1. Architecture of dual BERT model

After passing two different pre-trained models, the final hidden state for our transformer is pooled an average operation to reduce the previous hidden state. While the pooling layer is often used in CNN, the previous study proves the usefulness of pooling operation in NLP. The goal is to extract the most representative features of the sentence, this is an important part because it compacts the vector representation to capture the salient parts of the text and can be used as input of the classifier layer [24]. Suppose the pooling region of inputs is C, the average pooling is defined as Equation (1).

$$\text{Average Pooling} = \frac{\sum C}{|C|} \tag{1}$$

where C = pooling region of inputs.

The next layer is concatenated and passed into a fully connected layer that combines them and creates the probability output.

There are three different scenarios created. The model scenario is as follows:

1) BERT model
2) IndoBERT model
3) BERT + IndoBERT models (proposed model)

While having dual BERT for the proposed model, Figure 2 shows scenarios 1 and 2 are also prepared to evaluate the differences in performance generated by the dual BERT model, where the scenario 1 using BERT model, while scenario 2 using IndoBERT model. The scenarios are chosen because of how well the performance of the BERT model, and we also implement the scenario using IndoBERT model for classifying personality.

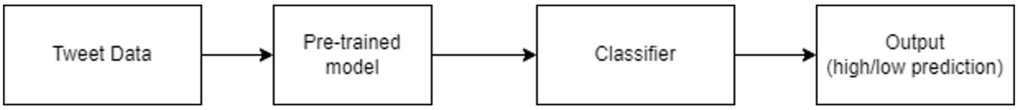

Figure 2. Scenario of pre-trained model

3.3. **Model training and evaluation.** The number the batch size used in training is 16 for both the BERT and IndoBERT baseline models. Meanwhile, the batch size for the dual BERT model is 10 due to memory limitation. For the epoch, as the BERT authors recommended [21], between 2 and 4 epochs and only pick the best results. In this case, 3 epochs are for all traits in proposed model, and 4 epochs are for all traits in baseline models. For BERT authors also recommend the use of Adam with a learning rate of 2e-5, but due to better performance results, we use a learning rate of 1e-5 for our models. Finally, in our proposed model, sparse categorical cross-entropy is used as a loss function.

The models were then evaluated based on classifying performance for each trait, which consists of Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. The first model is IndoBERT, the second model is BERT, and a combination of both IndoBERT and BERT models.

4. **Result and Discussion.** The experiment is held by using Pytorch and Keras libraries on the Google Colab platform. The result is based on the accuracy and F1 of the high and low labels for each trait. All scenario results are shown in Table 2. As shown in the table, our proposed model successfully outperforms both the BERT and IndoBERT model in every trait, except the Extraversion trait, where the BERT model performs better in terms of accuracy. The IndoBERT model performs better F1 Low score. This happened due to the nature of Extraversion not only based on the content of the tweet but also the number of replies made by users.

Table 2. Performance comparison between models

Trait	**Proposed model**			**BERT**			**IndoBERT**		
	Accuracy	**F1 Low**	**F1 High**	**Accuracy**	**F1 Low**	**F1 High**	**Accuracy**	**F1 Low**	**F1 High**
O	**73.25%**	**66.90%**	**77.56%**	69.59%	57.31%	74.67%	71.51%	63.35%	76.70%
C	**82.49%**	**87.41%**	**71.26%**	81.61%	84.50%	68.79%	81.01%	86.32%	68.32%
E	78.15%	56.67%	**85.39%**	**79.86%**	53.71%	80.32%	77.20%	**61.11%**	83.87%
A	**77.31%**	**77.37%**	**77.25%**	74.91%	64.79%	75.66%	75.49%	73.43%	**77.25%**
N	**80.53%**	**82.02%**	**78.78%**	80.10%	77.29%	76.44%	77.86%	79.54%	75.89%
Average	**78.35%**	**74.07%**	**78.05%**	77.21%	67.52%	75.18%	76.61%	72.75%	76.41%

Table 2 shows the comparison of our proposed model's performance to the baseline model for each trait. It shows that our proposed model produces the highest performance in all traits, except for the Extraversion trait where the BERT model is got a better result in accuracy score and the IndoBERT model in F1 Low score. For our baseline model evaluation result, BERT is better on accuracy. While IndoBERT model gets a better result in F1 scores, both low and high. We are optimistic that IndoBERT should be able to predict more accurately because it focused on Bahasa Indonesia. The difference between our BERT and IndoBERT model performance is not too much, this happened because BERT is a multilingual model which also can recognize even more languages, including Bahasa Indonesia and English.

Compared to the average of all indicators, the proposed model also shows a better performance compared to our baseline, 78.35% accuracy, 74.07% F1 Low score, and 78.05% F1 High score. From this result, it can be concluded that the proposed model performs better than our baseline model.

Table 3. Comparison model performance

Research work	Average accuracy	Average F1
Pratama and Sarno [10]	65%	–
Ong et al. [25]	70.50%	–
Ong et al. [12]	74.23%	–
Ergu et al. [26]	75.7%	–
Christian et al. [22]	76.98%	75.7%
Our Study		
BERT	77.21%	72.96%
IndoBERT	76.61%	74.58%
Proposed model	**78.35%**	**76.06%**

Finally, in Table 3 we compare our model's performance with the previous study that uses Big Five Personality models and our model has the best performance on the Twitter dataset.

The result experiments show that even though our proposed model is still capped by memory limitation to get the ideal batch size, the dual BERT method successfully outperforms the baseline model, achieving 76.06% average F1 score and 78.35% average accuracy. When compared to the previous study [22], our proposed model produces better results for F1 score by 0.36% and accuracy by 1.37%. The single BERT and IndoBERT performance fail to surpass the previous result, in which the cause is most likely because most Indonesian users on Twitter sometimes also use mixed language in their tweets, and on the other hand, the proposed model, which is a combination of BERT and IndoBERT, succeeds scoring a higher result in terms of accuracy and F1 score. Based on this result, the ability of the IndoBERT model in processing Indonesian text, combined with BERT's ability in other language texts coupled with the given combination method, has proven to be able to provide more accurate predictions when compared to the methods in previous studies.

5. **Conclusion.** Within this study, it can be shown that the field of personality prediction still has room to develop. The results provided through this research are the architecture of the classification model of personality that can be applied to applications to support and perform personality analysis on social media texts for several purposes, such as the interview process when applying for new employees. The proposed model is the combination of a pre-trained model based on Bahasa Indonesia, and multilingual deep learning models, which are IndoBERT and BERT. To serve as a comparison, three scenarios are used in this study. The dual BERT model successfully having increased both the average of accuracy and average of F1 score in comparison to the previous study and our baseline model, with better accuracy means that the application made will be better and more reliable in determining decisions. However, the impact of doing a combination of training two models simultaneously causes problems such as memory limitations and training times that take longer than usual.

In the future, personality prediction may also consider exploring a combination of another pre-trained model like RoBERTa or XLNet, which has proven to have an improvement over BERT. Furthermore, another pre-trained model like ALBERT, lite version of BERT focused on solving the problem of memory limitations and training speed of BERT, as it is the main problem in this research.

REFERENCES

[1] K. M. Carley, M. Malik, M. Kowalchuck, J. Pfeffer and P. Landwehr, Twitter usage in Indonesia, *SSRN Electronic Journal*, DOI: 10.2139/ssrn.2720332, 2018.

[2] B. K. Singh, M. Katiyar, S. Gupta and N. G. Ganpatrao, A survey on: Personality prediction from multimedia through machine learning, *Proc. of the 5th International Conference on Computing Methodologies and Communication (ICCMC2021)*, pp.1687-1694, DOI: 10.1109/ICCMC51019.2021.9418384, 2021.

[3] F. L. Schmidt and J. E. Hunter, The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings, *Psychological Bulletin*, vol.124, no.2, pp.262-274, DOI: 10.1037/0033-2909.124.2.262, 1998.

[4] J. A. Johnson, Measuring thirty facets of the five factor model with a 120-item public domain inventory: Development of the IPIP-NEO-120, *Journal of Research in Personality*, vol.51, pp.78-89, DOI: 10.1016/j.jrp.2014.05.003, 2014.

[5] S. V. Paunonen and M. C. Ashton, Big five factors and facets and the prediction of behavior, *Journal of Personality and Social Psychology*, vol.81, no.3, pp.524-539, DOI: 10.1037/0022-3514.81.3.524, 2001.

[6] S. Roccas, L. Sagiv, S. H. Schwartz and A. Knafo, The Big Five Personality factors and personal values, *Personality and Social Psychology Bulletin*, vol.28, no.6, pp.789-801, DOI: 10.1177/0146167202289008, 2002.

[7] B. Wilie, K. Vincentio, G. I. Winata, S. Cahyawijaya, X. Li, Z. Y. Lim, S. Soleman, R. Mahendra, P. Fung, S. Bahar and A. Purwarianti, IndoNLU: Benchmark and resources for evaluating Indonesian natural language understanding, *Proc. of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 2020.

[8] N. Azzouza, K. Akli-Astouati and R. Ibrahim, TwitterBERT: Framework for Twitter sentiment analysis based on pre-trained language model representations, in *Emerging Trends in Intelligent Computing and Informatics. IRICT 2019. Advances in Intelligent Systems and Computing*, F. Saeed, F. Mohammed and N. Gazem (eds.), Cham, Springer, 2020.

[9] O. Kampman, E. J. Barezi, D. Bertero and P. Fung, Investigating audio, video, and text fusion methods for end-to-end automatic personality prediction, *ACL 2018 – 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)*, vol.2, pp.606-611, DOI: 10.18653/v1/p18-2096, 2018.

[10] B. Y. Pratama and R. Sarno, Personality classification based on Twitter text using Naïve Bayes, KNN and SVM, *Proc. of 2015 International Conference on Data and Software Engineering (ICoDSE2015)*, pp.170-174, DOI: 10.1109/ICODSE.2015.7436992, 2015.

[11] A. S. Khan, H. Ahmad, M. Z. Asghar, F. K. Saddozai, A. Arif and H. A. Khalid, Personality classification from online text using machine learning approach, *International Journal of Advanced Computer Science and Applications*, vol.11, no.3, pp.460-476, DOI: 10.14569/ijacsa.2020.0110358, 2020.

[12] V. Ong, A. D. S. Rahmanto, W. Williem, D. Suhartono, A. E. Nugroho, E. W. Andangsari and M. N. Suprayogi, Personality prediction based on Twitter information in Bahasa Indonesia, *Proc. of the 2017 Federated Conference on Computer Science and Information Systems (FedCSIS2017)*, no.10, pp.367-372, DOI: 10.15439/2017F359, 2017.

[13] Y. Lecun, Y. Bengio and G. Hinton, Deep learning, *Nature*, vol.521, no.7553, pp.436-444, DOI: 10.1038/nature14539, 2015.

[14] T. Tandera, Hendro, D. Suhartono, R. Wongso and Y. L. Prasetio, Personality prediction system from Facebook users, *Procedia Computer Science*, vol.116, pp.604-611, DOI: 10.1016/j.procs.2017.10.016, 2017.

[15] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Pakistan Journal of Zoology*, vol.50, no.6, pp.2199-2207, DOI: 10.17582/journal.pjz/2018.50.6.2199.2207, 1997.

[16] S. Siami-Namini, N. Tavakoli and A. S. Namin, The performance of LSTM and BiLSTM in forecasting time series, *Proc. of 2019 IEEE International Conference on Big Data (Big Data 2019)*, pp.3285-3292, DOI: 10.1109/BigData47090.2019.9005997, 2019.

[17] S. Leonardi, D. Monti, G. Rizzo and M. Morisio, Multilingual transformer-based personality traits estimation, *Information (Switzerland)*, vol.11, no.4, pp.1-21, DOI: 10.3390/info11040179, 2020.

[18] E. Tighe, O. Aran and C. Cheng, Exploring neural network approaches in automatic personality recognition of Filipino Twitter users, *The 20th Philippine Computing Science Congress*, University of the Cordilleras, Baguio City, Philippines, 2020.

[19] M. Ramezani, M.-R. Feizi-Derakhshi, M.-A. Balafar, M. Asgari-Chenaghlu, A.-R. Feizi-Derakhshi, N. Nikzad-Khasmakhi, M. Ranjbar-Khadivi, Z. Jahanbakhsh-Nagadeh, E. Zafarani-Moattar and T.

Rahkar-Farshi, Automatic personality prediction; An enhanced method using ensemble modeling, 2020, *Neural Computing and Applications*, vol.34, pp.18369-18389, DOI: 10.1007/s00521-022-07444-6, 2022.

[20] N. H. Jeremy and D. Suhartono, Automatic personality prediction from Indonesian user on Twitter using word embedding and neural networks, *Procedia Computer Science*, vol.179, pp.416-422, DOI: 10.1016/j.procs.2021.01.024, 2021.

[21] J. Devlin, M. W. Chang, K. Lee and K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, *NAACL HLT 2019 – 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies – Proceedings of the Conference*, vol.1, pp.4171-4186, 2019.

[22] H. Christian, D. Suhartono, A. Chowanda and K. Z. Zamli, Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging, *Journal of Big Data*, vol.8, no.1, pp.1-20, DOI: 10.1186/s40537-021-00459-1, 2021.

[23] Y. Qiao, C. Xiong, Z. Liu and Z. Liu, Understanding the behaviors of BERT in ranking, *arXiv.org*, arXiv: 1904.07531, 2019.

[24] V. Suárez-Paniagua and I. Segura-Bedmar, Evaluation of pooling operations in convolutional architectures for drug-drug interaction extraction, *BMC Bioinformatics*, vol.19, no.8, DOI: 10.1186/s12859-018-2195-1, 2018.

[25] V. Ong, W. Williem and D. Suhartono, *Exploring Personality Prediction from Text on Social Media: A Literature Review Personality Recognition View Project Argumentation Mining View Project*, https://www.researchgate.net/publication/316254176, 2017.

[26] I. Ergu, Z. Isik and I. Yankayis, Predicting personality with Twitter data and machine learning models, *2019 Innovations in Intelligent Systems and Applications Conference (ASYU)*, pp.1-5, DOI: 10.1109/ASYU48272.2019.8946355, 2019.