


Sentiment Analysis Of Indonesia National Team Naturalization Using Bidirectional Encoder Representations From Transformers

Diva Ahmad Maulana¹, Chaerur Rozikin², Aries Surharso³

Computer Science, Singaperbangsa Karawang University, Karawang, Indonesia

Article Info	ABSTRACT
Keywords: Sentiment Analysis, Classification, BERT, Transformers, Indonesian National Team.	In this era of rapid development of information technology, the number of internet users is increasing, supported by the popularity of social media as a medium for sharing information and interacting. The X social media platform is one of the media that is often used to convey public opinion. One of the hot issues discussed on X social media is the Indonesian National Team naturalization program. This program has triggered various public responses, both pro and con. This study aims to analyze public sentiment regarding the program using the Bidirectional Encoder Representation from Transformer (BERT) algorithm with the Knowledge Discovery in Database method. Data was collected using scraping techniques on the X social media platform which were then selected and labeled positive, negative, and neutral. Testing the BERT algorithm using the pre-trained indoBERT model was tested by dividing the training and testing data 80:20, and evaluated with a confusion matrix. With a dropout of 0.3, the evaluation results showed the highest accuracy value of 90%, precision 81%, recall 74%, and f1-score 77%. The results of this study are expected to be useful for evaluation materials and to support decision making by related parties.
This is an open access article under the CC BY-NC license 	Corresponding Author: Diva Ahmad Maulana Computer Science, Singaperbangsa Karawang University Karawang, Indonesia 2110631170134@student.unsika.ac.id

INTRODUCTION

The rapid development of information technology has made the internet an inseparable element in daily activities. The number of users accessing the internet in Indonesia continues to increase every year based on information from We Are Social. The increase in the number of internet users is also supported by the popularity of social media as one of the main platforms for interacting and sharing information. According to data from We Are Social, of the 270 million population of Indonesia, about half of the population actively uses social media, with 139 million users recorded in 2024. Social media has become a discussion space that allows people to freely express their opinions on various social, political, and cultural issues. In this context, social media functions as a digital public space that facilitates the exchange of ideas, debate, and the formation of public discourse (Suhendra & Pratiwi, 2024).

Of the various social media platforms available, X (Twitter) is one of the media that is often used to convey opinions in real time. The X platform has become an effective digital public space to monitor public opinion on various issues (Suhendra & Pratiwi, 2024). These

issues include those related to national sports such as the naturalization of players for the Indonesian National Team.

The issue of naturalization of Indonesian National Team players is a topic that is widely discussed on social media (Lubis et al., 2025). This issue is trending because it often raises pros and cons. Some agree with naturalization, for example, related to improving the quality of the team, while others disagree regarding the lack of opportunities for local players to defend the Indonesian National Team or whether naturalized players are truly part of the Indonesian National Team wholeheartedly. This makes social media user reviews a relevant source of data to understand public sentiment towards the naturalization policy.

Public sentiment regarding the naturalization program of the Indonesian National Team can have an impact on PSSI, the Indonesian National Team, and the players themselves. If the majority opinion is negative, it can raise public trust issues towards PSSI policies and reduce support for the Indonesian National Team. In addition, naturalized players will tend to be the target of criticism and condemnation if they do not play well. Conversely, if the majority opinion is positive, it can build public trust in PSSI and greater support for the team or the players themselves. Reviews on social media X can be used as a reference to see the level of public support or rejection of the Indonesian National Team naturalization policy. This review is important because it can be used as evaluation material by related parties, such as the football federation and the government, in making better decisions and policies in the future.

User reviews of social media platform X are purely unorganized (Setiamukti & Nasvian, 2023). Therefore, in order to be organized, a classification is needed that can divide these reviews into certain categories, such as positive, negative, or neutral, with the help of certain methods. In sentiment analysis, text mining is a technique used to classify text. Text classification can be assumed as the process of determining text documents into certain categories.

Previous research conducted by (Lubis et al., 2025) used Support Vector Machine (SVM) as an algorithm in analyzing public sentiment on Twitter regarding the naturalization of the Indonesian National Team. Data was obtained by scraping and then preprocessing. SVM was applied as a model using accuracy, precision, recall, and f1-score as evaluation metrics. The evaluation results showed the performance of the model with data division of 80% training and 20% testing resulting in an accuracy of 71%. Another study by (Franko et al., 2024) analyzed the sentiment of player naturalization on Youtube by comparing the performance of the Decision Tree algorithm with Naive Bayes. The best performance was shown by the Naive Bayes algorithm with an accuracy of 85.4%, higher than the accuracy of the Decision Tree algorithm of 70.9%.

The development of natural language processing (NLP) technology has introduced deep learning-based models such as BERT (Bidirectional Encoder Representations from Transformers). (Alawi & Bozkurt, 2024) research compared several algorithms to analyze the sentiment of Twitter user reviews about universities in Turkey. Conventional machine learning algorithms used in this study such as Naive Bayes, Logistic Regression, SVM, Decision Tree, Multi-layer Perceptron (MLP), Random Forest, Categorical Boosting Algorithm, Extreme Gradient Boosting, and Light Gradient Boosting Machine. Deep learning algorithms such as

LSTM, BiLSTM, CNN, and hybrid CNN-BiLSTM are also used as a comparison in this study. Finally, the BERT algorithm is implemented as a model, by applying several pre-trained models. Based on the comparison of each algorithm, the highest accuracy value of 91% was produced from the BERT-BiLSTM-CNN hybrid model with precision, recall, and f1-score values of 89%, 87%, and 88%. Another study by (Pradana, 2024) conducted a sentiment analysis of the performance of DKI Jakarta Governor Heru Budiarto Hartono on the Twitter platform using the IndoBERT algorithm. The evaluation results showed good results with an accuracy of 90.5%.

From the research background that has been explained previously, conventional machine learning algorithms such as SVM and Naive Bayes have been used to analyze sentiment regarding the naturalization of the Indonesian National Team, but the accuracy is not too high. Meanwhile, the BERT (Bidirectional Encoder Representations from Transformers) algorithm can provide quite good performance in analyzing the sentiment of X (Twitter) reviews on certain topics. BERT is able to capture the context of language more deeply and contextually (Ibrahim et al., 2025). Research on sentiment analysis of the naturalization of the Indonesian National Team using the BERT algorithm has never been done. So this study will solve the problem by applying sentiment analysis using the BERT model to reviews of X users regarding the naturalization program of the Indonesian National Team. The classification process will be divided into three categories of sentiment, namely negative, positive, and neutral, and will be carried out in Indonesian. This study will evaluate the performance of the BERT model using IndoBERT as a pre-trained model. It is hoped that this study can provide useful information for football federations, governments, and the public in evaluating and formulating policies related to player naturalization.

METHODS

In this study, the object is the review of X users regarding the naturalization of the Indonesian National Team. The data used in 2024 uses Indonesian language data. The algorithm used in this study is Bidirectional Encoder Representations from Transformer (BERT) with the Knowledge Discovery In Database (KDD) method which is a series of steps designed to find patterns or knowledge from datasets. The stages of this process include data selection, data preprocessing, data transformation, data mining, and evaluation (Rahayu et al., 2024).

Data Selection

In this initial stage, descriptive statistics are conducted to obtain an overview of the review data regarding the naturalization of the Indonesian national football team. Data is taken using a scraping technique by entering an authentication token. Web scraping is a technique for obtaining information or content contained on a website (Adila et al., 2022). The data was taken from social media X which allows its users to read and send reviews which are usually called "tweet" (Toy et al., 2021). Search parameters are set by entering the keyword "naturalization of the Indonesian national team" and setting the time span of tweets during 2024 in Indonesian. The data obtained will be classified into positive, negative and neutral labels by researchers manually and then validated and reviewed by an Indonesian language teacher.

Data Preprocessing

Data cleaning is done at the preprocessing stage to eliminate and overcome data noise. This preprocessing is done so that the data obtained is clean and easy to process for the next process. This process aims to make the calculation results optimal. Here are the preprocessing stages.

1. Cleaning
Reducing noise in tweet data. In this process, duplicates and unnecessary attributes such as hashtags, punctuation characters, and other unique symbols are cleaned.
2. Case Folding
Changing all letters in the data to lowercase. This is done to facilitate the data processing process and ensure consistency of the data to be processed.
3. Normalization
Making changes to non-standard words into words that are in accordance with language rules.
4. Filtering
Removing words that are considered unimportant and meaningless, such as the words "bisa", "hanya", "dari" and others using stopwords.
5. Stemming
Making changes to affixed words into their basic form.

Data Transformation

In data analysis or modeling, transformation is used as a process of changing data from its original structure or format into a format that is more appropriate and informative for use. In this study, data transformation includes the following processes.

1. Token Embeddings
A numeric representation of each token (word or sub-word) in the input text. Token embeddings are obtained from the BERT vocabulary, which has been pre-trained to produce rich word representations.
2. Segment Embeddings
Additional embeddings used to distinguish parts of the text in the input, for example in the case of sentence pair classification. This allows the model to understand the input structure consisting of two different sentences in one representation.
3. Position Embeddings
Embeddings that provide information about the position or order of tokens in an input. Since Transformers like BERT do not have inherent sequential properties, position embeddings ensure that the model understands the order of words in a sentence.
4. Filtering
The process of combining token embeddings, segment embeddings, and position embeddings into one representation that can be used as input to the model. This representation serves as the main numeric input processed by the Transformer layer in BERT.

Data Mining

In the Knowledge Discovery in Database (KDD) process, the data mining stage is the main part, where the use of machine learning models to extract insights and information from previously processed data. Data mining aims to identify patterns or relationships that can provide useful insights (Haryatmi & Hervianti, 2021). In this study, the Bidirectional Encoder Representations from Transformers (BERT) algorithm is used to perform classification. This model can be applied to a variety of Natural Language Processing (NLP) tasks by making only minor changes to the model training (Chakkarwar et al., 2023). Classification using the BERT algorithm will be tried using different amounts of data to determine the effect of the amount of data on the accuracy results, to determine the best accuracy results on BERT.

The data mining stage includes data division, where the review data that has been labeled with sentiment will be divided into two parts, namely training data and testing data. This data is divided with a ratio of 80% for training and also 20% for testing. This ratio is chosen based on the Pareto principle, which is commonly used in machine learning modeling.

For the BERT model training process, training data will be used. Numerical input from the transformed data is processed through the Transformer layer in BERT to understand the context of each word in the sentence, both from left to right and right to left. To test the performance of the trained model, the next testing data is used. The model will predict the sentiment label (negative, positive, or neutral) for the testing data to measure the model's performance.

Evaluation

In this stage, an evaluation is carried out to measure the accuracy of the results of the BERT algorithm application that has been carried out previously. In this evaluation stage, the Confusion Matrix method is used with parameters including accuracy, precision, recall, and f1-score (Singh et al., 2021).

RESULTS AND DISCUSSION

Data Mining

The data mining process is carried out using Bidirectional Encoder Representation from Transformer (BERT). IndoBERT is used as a pre-trained model specifically for Indonesian. The split data operator is applied to divide the training data and testing data with a ratio of 80:20. The hyperparameters in this study use the Adam optimizer, learning rate $5e-5$, dropout 0.3, and 8 epochs.

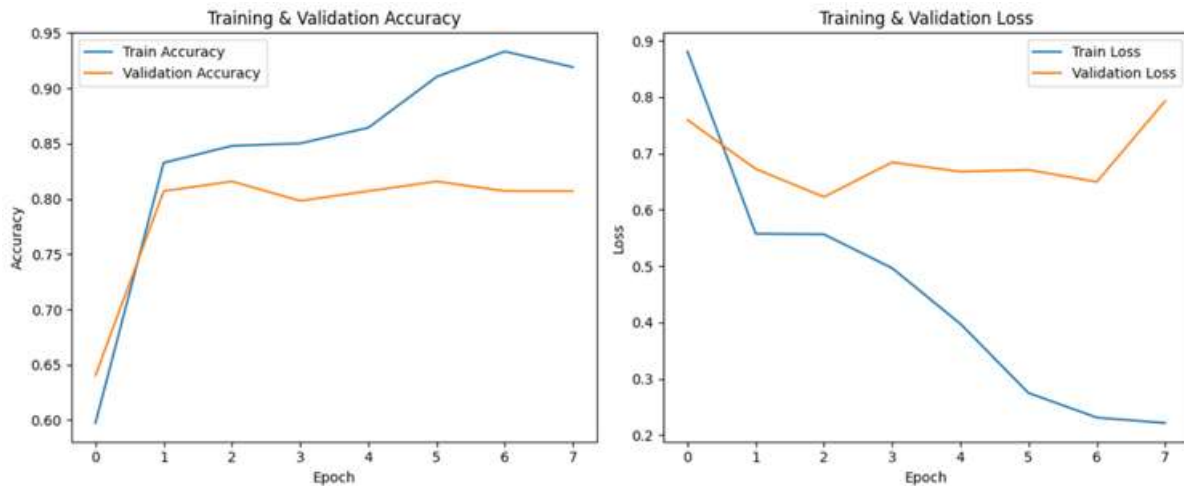


Figure 1. Accuracy dan Loss Graphic

After fine-tuning using BERT, it can be seen in Figure 1, the training accuracy graph has increased significantly, while validation tends to start to stabilize at around 80%. Conversely, the loss graph in the training process decreases as the number of epochs increases. The loss in validation also decreases, but increases at the end of training.

Evaluation

In this stage, an evaluation is carried out to measure the accuracy of the results of the BERT algorithm application that has been carried out previously. In this evaluation stage, the confusion matrix method is used to support the research objectives, with parameters including accuracy, precision, recall, and f1-score.

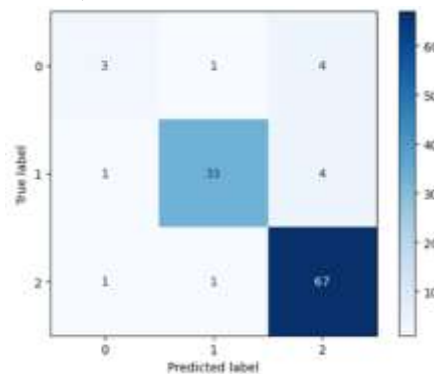


Figure 2 Confusion Matrix Results

Based on the confusion matrix results shown in Figure 2, it can be seen that the positive class (label 2) is classified very well, with 67 out of 69 data correctly classified. The neutral class (label 1) also shows good performance, although there are still 5 data misclassifications (1 to negative and 4 to positive). The negative class (label 0) shows low performance, 3 of which are correctly classified, while the rest are incorrectly classified as neutral and positive. This indicates that the model has difficulty distinguishing negative sentiment from other classes.

	precision	recall	f1-score	support
0	0.60	0.38	0.46	8
1	0.94	0.87	0.90	38
2	0.89	0.97	0.93	69
accuracy			0.90	115
macro avg	0.81	0.74	0.77	115

Figure 3. Classification Report

As can be seen from Figure 3, from the three sentiment categories, the positive sentiment class shows good performance with a precision value of 0.89, recall 0.97, and f1-score 0.93. The neutral category also shows quite good values with a precision value of 0.94, recall 0.87, and f1-score 0.90. However, the model performance for the negative category still tends to be less good, with a precision of 0.60, recall 0.38, and f1-score 0.46. The macro average is the average value of precision, recall, and f1-score which show values of 0.81, 0.74, and 0.77 respectively. The final accuracy result of the model is 90%.

CONCLUSION

Based on the results of the research that has been done, it can be concluded that this study conducted a sentiment analysis of the Indonesian National Team naturalization program from social media X. Using the Knowledge Discovery in Database (KDD) method, the dataset was selected, data preprocessing was carried out, then transformed into numeric, which was then fine-tuned using the Bidirectional Encoder Representation from Transformer model and evaluated using a confusion matrix. Based on the results of the performance test evaluation, with a data division of 80:20 the model showed quite good performance with an accuracy value of 90%, precision 81%, recall 74%, and f1-score 77%. Suggestions for further research, namely increasing and balancing the amount of data in the dataset can be implemented to improve the accuracy results of the classification. Implementation of data division with various scenarios can also be done to obtain the most optimal results from the model. Implementing different algorithms or hyperparameters to find out which algorithm is better in performing sentiment analysis.

REFERENCE

- Adila, N., Falentino, S., & Jatmiko, W. (2022). Implementation of Web Scraping for Journal Data Collection on the SINTA Website. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, 7(4), 2478–2485. <https://doi.org/10.33395/sinkron.v7i4.11576>
- Alawi, A. B., & Bozkurt, F. (2024). A hybrid machine learning model for sentiment analysis and satisfaction assessment with Turkish universities using Twitter data. *Decision Analytics Journal*, 11. <https://doi.org/10.1016/j.dajour.2024.100473>
- 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. https://doi.org/10.2991/978-94-6463-136-4_12
- Franko, B., Wilyanto, N., & Irsyad, H. (2024). Analisis Sentimen Terhadap Naturalisasi Pemain

- pada Youtube Menggunakan Decision Tree dan Naive Bayes. *SESSION: Software Development, Digital Business Intelligence, and Computer Engineering*, 3(1), 8–16. <https://doi.org/10.57203/session.v3i1.2024.8-16>
- Haryatmi, E., & Hervianti, S. P. (2021). Penerapan Algoritma Support Vector Machine Untuk Model Prediksi Kelulusan Mahasiswa Tepat Waktu. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(2), 386–392. <https://doi.org/10.29207/resti.v5i2.3007>
- Ibrahim, A. K., Mustikasari, M., & Bastian, I. (2025). Analisis Sentimen Pada Ulasan Aplikasi ASTRO – GROCERIES IN MINUTES Di Google Play Store Menggunakan Metode Bidirectional Encoder Representations From Transformers (BERT). *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, 13(1), 1–48. <https://doi.org/10.23960/jitet.v13i1.6033>
- Lubis, M. G. R., Sitompul, D. S., Giovanni, T. M., & Dewi, S. (2025). Evaluasi Kinerja Algoritma Support Vector Machine (SVM) Dalam Analisis Sentimen Publik Terhadap Naturalisasi Timnas Indonesia di Twitter. *JALAKOTEK: Journal of Accounting Law Communication and Technology*, 2(1), 81–89. <https://doi.org/10.57235/jalakotek.v2i1.4180>
- Pradana, L. S. (2024). *Analisis Sentimen Masyarakat Media Sosial twitter Terhadap Kinerja Pejabat Gubernur Dki Jakarta Menggunakan Model IndoBERT* [UIN Syarif Hidayatullah Jakarta]. <https://repository.uinjkt.ac.id/dspace/handle/123456789/77071>
- Rahayu, P. W., Sudipa, I. G. I., Suryani, Surachman, A., Ridwan, A., Darmawiguna, I. G. M., Sutoyo, M. N., Slamet, I., Harlina, S., & Maysanjaya, I. M. D. (2024). *Buku Ajar Data Mining* (Issue January 2024). PT. Sonpedia Publishing Indonesia.
- Setiamukti, M. F., & Nasvian, M. F. (2023). Social Network Analysis #Usuttuntas Pada Media Sosial Twitter (Data Twitter 11 November 2022). *Ekspresi Dan Persepsi: Jurnal Ilmu Komunikasi*, 6(1), 124–137. <https://doi.org/10.33822/jep.v6i1.5427>
- Singh, K. K., Singh, A., Elhoseny, M., & Elngar, A. A. (2021). Machine Learning and the Internet of Medical Things in Healthcare. In *Machine Learning and the Internet of Medical Things in Healthcare*. Mara Conner. <https://doi.org/10.1016/C2019-0-03077-4>
- Suhendra, & Pratiwi, F. S. (2024). Peran Komunikasi Digital dalam Pembentukan Opini Publik : Studi Kasus Media Sosial. *Resiliensi Indonesia Dalam Pusaran Disrupsi Global*, 293–315. <https://doi.org/10.30589/proceedings.2024.1059>
- Toy, K. V. S., Sari, Y. A., & Cholissodin, I. (2021). Analisis Sentimen Twitter menggunakan Metode Naive Bayes dengan Relevance Frequency Feature Selection (Studi Kasus: Opini Masyarakat mengenai Kebijakan New Normal). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(11), 5068–5074. <http://j-ptiik.ub.ac.id>