Jurnal Teknologi Informasi dan Pendidikan
Volume 19, No. 1, March 2026
https://doi.org/10.24036/jtip.v19i1.1104

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Implementation of IndoBERT in Sentiment Analysis of Free Nutritious Meal Programs on the X Social Media Platform

Muhammad Mursyid[1*✉], Arif Setiawan[1], Muhammad Arifin[1]
[1]Faculty of Engineering, Universitas Muria Kudus, Kudus, Indonesia
Corresponding Author: 202153144@std.umk.ac.id

Article Information

Article history:

No. 1104
Rec. January 22, 2026
Rev. February 26, 2026
Acc. February 28, 2026
Pub. March 07, 2026
Page. 1228 – 1242

Keywords:

- *Sentiment Analysis*
- *Free Nutritious Meals*
- *indoBERT*
- *Social Media X*
- *Natural Language Processing*

ABSTRACT

The Free Nutritious Meal Program (MBG) is a strategic initiative by the Indonesian government to improve child nutrition and prevent stunting. However, its implementation has sparked diverse public opinions on social media, which are difficult to analyze manually due to the large volume of data. This study aims to identify public sentiment toward the MBG program through social media X by implementing the IndoBERT model. A total of 4,380 tweets were collected using web scraping techniques with relevant keywords between March and May 2025. The research process included preprocessing (data cleaning, stopword removal, stemming, and tokenization), semi-automatic data labeling, and data division into a 71.97% training set, 8.02% validation set, and 20.01% test set. The model used was the Indonesian RoBERTa Base Sentiment Classifier architecture, which underwent a fine-tuning process for 20 epochs. The results showed that the IndoBERT model achieved an accuracy rate of 80.11% and a weighted average F1-score of 0.8000. Negative sentiment was detected most accurately with an F1-score of 0.8301. Although effective, the model still faces challenges in handling linguistic ambiguity in neutral sentiment and the risk of overfitting. Further research is recommended to expand slang language normalization and apply stricter model regulation techniques.

How to Cite:

Mursyid, M., & et al. (2026). Implementation of IndoBERT in Sentiment Analysis of Free Nutritious Meal Programs on the X Social Media Platform. Jurnal Teknologi Informasi Dan Pendidikan, 19(1), 1228-1242. https://doi.org/10.24036/jtip.v19i1.1104

1. INTRODUCTION

The Free Nutritious Meals Program (MBG) is a strategic initiative by the Indonesian government to improve the nutrition of children and adolescents, address food access inequality, and prevent stunting in school-age children [1] . However, its implementation has sparked a range of public opinions, from support to criticism, regarding logistics, budget transparency, and effectiveness [2]. Social media, as a digital public space, has become the main channel for the public to express their opinions [3]. The challenge today is how to capture public opinion quickly, broadly, and accurately amid an abundance of data that is disproportionate to manual analysis capacity [4].

Previous studies have demonstrated the application of various methods in sentiment analysis regarding public policy issues on social media. The first study used Support Vector Machine (SVM) combined with Recursive Feature Elimination (RFE) to analyze sentiment toward the entertainment tax hike policy and achieved 95% accuracy. This demonstrates the effectiveness of this method in classifying policy issues [5]. The second study utilized the Naïve Bayes Classifier to evaluate sentiment towards the second COVID-19 booster vaccination and produced an accuracy of 86.84%, with the majority of public sentiment being positive towards the policy [6]. Furthermore, the third study used a combination of SVM, TF-IDF, and Random Forest approaches to classify public opinion on the Jakarta Provincial Government's policies during the pandemic. The SVM method proved to be the most accurate with an accuracy value of 77.58%, superior to the other two methods [7]. Furthermore, the fourth study applied the Naïve Bayes method with NLTK preprocessing support in analyzing sentiment about PPKM, with an accuracy ranging from 75% to 77% depending on the filtering used [8]. Furthermore, the fifth study used the indoBERT method, a transformer-based language model specifically for Indonesian, which achieved a fairly high accuracy of 97% in classifying e-commerce reviews, much higher than the LSTM model, which only reached 85.08% [9].

The IndoBERT model was chosen because it is a transformer-based model that has been specifically trained for the Indonesian language and has been proven to provide high accuracy in Indonesian sentiment classification tasks, both in the product and public policy domains [10]. The five previous studies show that choosing the right analysis method greatly affects the accuracy and effectiveness of understanding public opinion through social media. However, this study was conducted in relation to the challenges of discussing public opinion quickly, broadly, and accurately. One of the main challenges faced is the mismatch between the large amount of data on social media and the manual analysis capacity needed to identify complex sentiment patterns. The need for automated and structured opinion analysis can be achieved through the field of Natural Language Processing (NLP). This study focuses on identifying public sentiment towards the Free Nutritious Meals program on the X platform by implementing the IndoBERT model for automatic and contextual classification. Through automated sentiment analysis, public opinion can be assessed objectively and comprehensively to provide a more accurate picture of the program currently underway. The results of this research are expected to provide useful insights for the government in evaluating policies, improving public communication, and promoting transparency in the future.

2. RESEARCH METHOD

The methodology in this study is designed in a structured manner so that the problem-solving process can be carried out effectively. The detailed research flow can be seen in Figure 1.

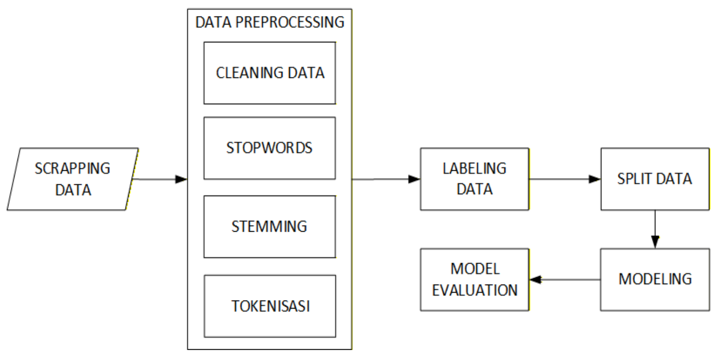

Figure 1. Research Methodology

2.1. Data Scraping

The initial stage of this research began with data collection using web scraping techniques from social media X. Data scraping is a method of utilizing Search API for raw data collection. The tool used was Tweet Harvest based on Node.js, which allows automatic data retrieval based on specific keywords this data was then used as the basis for sentiment analysis in this research [11].

2.2 Data Preprocessing

The second stage is data preprocessing, which transforms raw text data through a process to eliminate noise and improve feature quality [12]. The preprocessing stage in this study consists of three main steps as follows

2.2.1 Data Cleaning

The initial stage involves cleaning the text of non-informative elements, including removing URLs, hashtags, mentions, emojis, punctuation marks, and other special characters [13]. This aims to minimize the decline in model accuracy.

2.2.2 Stopword Removal

The stopword removal process involves removing common words in Indonesian that appear frequently but do not carry significant meaning in sentiment analysis, such as conjunctions or prepositions (e.g., "yang," "dan," "itu") [14]. This process aims to reduce the data dimension without losing essential information.

2.2.3 Stemming

The stemming process is the conversion of inflected words in the data into their root forms. The purpose of the stemming process is to facilitate data analysis and grouping. This study uses the sastrawi library, which was developed specifically to handle Indonesian text [15].

2.2.4 Tokenization

The purpose of the tokenization process is to break down text into smaller units (tokens) [16], so that it can be processed efficiently by the IndoBERT model. At this stage, non-standard words (slang) are also normalized into standard forms according to PUEBI to ensure that the model understands the semantic context accurately.

2.3 Data Labeling

The labeling process is carried out with the aim of classifying each tweet into a sentiment category. In this study, the categories used are positive, negative, and neutral. Labeling is carried out semi-automatically, which means that the system provides an initial sentiment prediction automatically using the previously trained IndoBERT model. After the initial prediction is obtained, a manual verification process is carried out by researchers to avoid classification errors and improve label accuracy. This verification involves several annotators or human evaluators who understand the context of sentiment in Indonesian. This manual review is important to handle ambiguous cases, irony, or sarcasm that are difficult for automatic models to understand.

2.4 Data Split

After the labeling process is complete, the data is then divided into two main subsets, namely the training set and the testing set. The division ratio follows the general standard in machine learning, which is 80% for training and 20% for testing. The training data (80%) is used to train the model to recognize patterns in the data and learn to distinguish sentiment based on existing features. The test data (20%) is used to measure the model's performance on data it has never seen before, thus providing an objective picture of the model's generalization capabilities. The division is done in a stratified manner, meaning that the proportion of sentiments (positive, negative, neutral) in each subset is maintained to be representative.

2.5 Modeling

The modeling stage is the main process in sentiment classification, where the IndoBERT model is used and adjusted to the research dataset through a fine-tuning process. The system is implemented using the PyTorch framework and the Transformers library from Hugging Face as the backend for model training and testing. This stage includes data tokenization, loading the pre-trained IndoBERT model, and adjusting the model parameters so that it can optimally recognize sentiment patterns in the text.

2.5.1 IndoBERT Model Architecture

IndoBERT is a language model based on the Transformer architecture, which consists of a multi-layer encoder. IndoBERT is based on the multilingual BERT base model, but has been retrained using a large Indonesian-language corpus, such as Wikipedia, Kompas, Detik, and other news sources. This model has 12 layers, 12 attention heads, and approximately 110 million parameters. The Transformer architecture works with a self-attention mechanism, which allows the model to understand the context between words more deeply and contextually, unlike the traditional approach based on ordinary word embedding [17].

2.5.2 Fine-Tuning the Model

The pre-trained IndoBERT model is then fine-tuned to specifically handle sentiment classification tasks on tweets. Fine tuning is an advanced training process in which the model weights are adjusted to a new dataset so that the model can perform classification with higher accuracy. In this process, a classification architecture is added to the top of IndoBERT in the form of a fully connected layer that functions as the final decision maker for the sentiment class. The training process uses AdamW optimization, learning rate decay, and early stopping to prevent overfitting. The model is trained for several epochs (usually 3–5 iterations of the entire dataset), with an adjusted batch size (e.g., 16 or 32), depending on the capacity of the GPU memory used. The pre-trained IndoBERT model is then fine-tuned to specifically handle the task of sentiment classification for tweets [18].

2.5.3 Pseudocode Algoritma Sentiment Classification

Based on Figure 1, it can be concluded that the pseudocode flow is as follows:

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Table 1 Pseudocode Algoritma Sentiment Classification

Pseudocode Algoritma Sentiment Classification
BEGIN
1. Load dataset D
2. Preprocess text: Clean text (URL, emoji, punctuation), Remove stopwords, Normalize slang, Stemming (Sastrawi)
3. Label tweets (semi-automatic + manual validation)
4. Encode labels and split data into Train, Validation, Test
5. Tokenize text using IndoBERT tokenizer (max_length = 128)
6. Load IndoBERT model (num_labels = 3)
7. Fine-tune model on Train set for defined epochs
8. Evaluate model using Test set: Compute Accuracy, Precision, Recall, F1
9. Predict sentiment and generate confusion matrix
END

Based on the Table, the sentiment classification algorithm consists of several main stages, namely data collection (scraping), preprocessing, labeling, data separation, tokenization, model training, and evaluation. The raw scraping data is cleaned through a process of removing URLs, mentions, hashtags, stopwords, slang normalization, stemming, and tokenization using the IndoBERT WordPiece tokenizer. The data is then semi-automatically labeled, manually verified, and divided into training, validation, and test data using stratified split. The IndoBERT model is then fine-tuned using AdamW for 20 epochs with a batch configuration of 32 and max_length of 128. At the evaluation stage, the model is measured using accuracy, precision, recall, and F1-score to determine the comprehensive performance of sentiment classification.

2.6 Model Evaluation

The IndoBERT model testing utilizes Scikit-learn to calculate evaluation metrics such as accuracy, precision, recall, and F1-score. The purpose of model evaluation is to measure how well the model's predictions perform in classifying sentiment [19]. Accuracy is the ratio between the number of correct predictions for both positive and negative classes to the total number of data analyzed. Precision is used to show how many of all positive predictions match the actual conditions. Recall is used to measure how much data that should be included in the positive class is correctly recognized by the model. F1-score is the average value of precision and recall [20].

3. RESULTS AND DISCUSSION

This section presents the results of implementing the IndoBERT model to analyze public sentiment towards the Free Nutritious Meals program on the X platform. The discussion focuses on presenting the test results, evaluating the model's performance through accuracy metrics, and interpreting findings related to public opinion trends. All results presented aim to provide an objective picture of the model's effectiveness in automatically classifying sentiment.

3.1 Data Scraping

The data collection process was carried out automatically using Tweet Harvest based on Node.js, using the keywords "makan siang gratis lang:id", "makan siang gratis", "makan bergizi gratis". A total of 4,380 rows of raw data were collected, with a time limit from March to May 2025, in a structured format containing the main text (full_text), user identity, Tweet URL. As shown in Table 1

Table 2. Data Scraping Results

Full Text	Tweet URL	Username
Dukung Program Makan Bergizi Gratis (MBG) untuk kesehatan anak dan pertumbuhan ekonomi lokal! Dengan melibatkan berbagai pihak MBG membawa manfaat yang luas bagi masyarakat. #DukungProgramMBG #DukungMBG #MakanBergiziGratis https://t.co/Z0SCJ5VCiA	https://x.com/olieverra/status/1925152254790926787	olieverra
Program Makan Bergizi Gratis (MBG) bukan hanya tentang kesehatan anak tapi juga tentang mendorong ekonomi lokal! Dengan melibatkan petani nelayan dan UMKM MBG membawa manfaat nyata bagi perekonomian. #DukungProgramMBG #DukungMBG #MakanBergiziGratis https://t.co/oyWxblPJlo	https://x.com/olieverra/status/1925152250093306172	olieverra
Program Makan Bergizi Gratis (MBG) tidak hanya mendukung kesehatan anak tetapi juga mendorong perputaran ekonomi lokal. Keterlibatan petani nelayan dan UMKM dalam penyediaan bahan pangan menjadikan MBG berdampak langsung pada sektor perekonomian. #DukungProgramMBG#DukungMBG https://t.co/MA3tL6sGyP	https://x.com/MadePuspita3/status/1925149805329973734	MadePuspita3

As shown in the figure, this raw data still contains many non-informative elements such as URLs, hashtags, and special characters that require further processing at the pre-processing stage to ensure the quality of input for the IndoBERT model.

3.2. Data Preprocessing

The preprocessing stage produced 3,643 clean data from the initial 4,380 tweets. A total of 737 data were deleted due to duplication, containing only URLs or media without text, or becoming empty after the cleaning process. The preprocessing stage in this study included data cleaning, stopword removal, and tokenization. This section describes the results of each step, not the technical procedures described in the Methods chapter.

3.2.1. Data Cleaning

The data cleaning process produces more concise text that is free of non-infomartive elements such as links, mentions, hashtags, emojis, or special characters. There results of the cleaning can be seen in the Table 3, which shows a comparison between the original text and the cleaned text that is ready to be processed to the next stage.

Table 3 Data Cleaning Results

Full Text	Clean Text
@reaareaa__ Program Makan Bergizi Gratis di Papua yang digagas Presiden Prabowo bikin masa depan anak-anak lebih cerah	program makan gizi gratis di papua digagas presiden prabowo bikin masa depan anakanak lebih cerah
@KangManto123 Pantes Ngotot bikin Proyek Makan Bergizi Gratis..... Ternyata untuk ngasih makan Influencer & pendukungnya	pantes ngotot bikin proyek makan bergizi gratis ternyata ngasih makan influencer pendukungnya
Makan siang gratis langkah kecil demi kemajuan besar bangsa Indonesia #PenuhiGiziIndonesia https://t.co/jIRn1gBLZp	makan siang gratis langkah kecil kemajuan besar bangsa indonesia

3.2.2. Stopwords

After text cleaning, stopwords are removed to retain only words that are relevant to sentiment determination. This reduction of common words makes the text representation more focused and concise. An example of the results of stopword removal is shown in Table 4.

Table 4 Stopword Results

Full Text	Stopwords
program makan gizi gratis di papua digagas presiden prabowo bikin masa depan anakanak lebih cerah	program makan gizi gratis papua gagas presiden prabowo bikin masa depan anakanak lebih cerah
pantes ngotot bikin proyek makan bergizi gratis ternyata ngasih makan influencer pendukungnya	pantes ngotot bikin proyek makan gizi gratis nyata ngasih makan influencer dukung
makan siang gratis langkah kecil kemajuan besar bangsa indonesia	makan siang gratis langkah kecil maju besar bangsa indonesia

3.2.3. Tokenization

During the tokenization stage, text that has undergone stopword removal is converted into tokens according to the IndoBERT input format. Tokenization is performed using a WordPiece-based tokenizer so that complex words can be broken down into sub-words that are easier for the model to understand. Examples of tokenization are shown in Table 5.

Table 5 Tokenization Results

StopWords	Tokenization
program makan gizi gratis papua gagas presiden prabowo bikin masa depan anakanak lebih cerah	["program", "makan", "gizi", "gratis", "papua", "gagas", "presiden", "Prabowo", "bikin", "masa", "depan", "anak", "lebih", "cerah"]
pantes ngotot bikin proyek makan gizi gratis nyata ngasih makan influencer dukung	["ngotot", "bikin", "proyek", "makan", "gizi", "gratis", "nyata", "ngasih","makan", "influencer", "dukung"]
makan siang gratis langkah kecil maju besar bangsa indonesia	["makan", "siang", "gratis", "langkah", "kecil", "maju", "besar", "bangsa", "Indonesia"]

3.3. Data Labeling

After data preprocessing, the next step is data labeling, which is to assign target values or sentiment categories to each line of cleaned text so that the machine can learn the patterns of differences between positive, negative, and neutral opinions. From the 3,643 data sets that underwent preprocessing, the data labeling process resulted in 1,556 negative labels (42.71%), 1,124 neutral labels (30.85%), and 963 positive labels (26.43%), as shown in Figure 2 below.

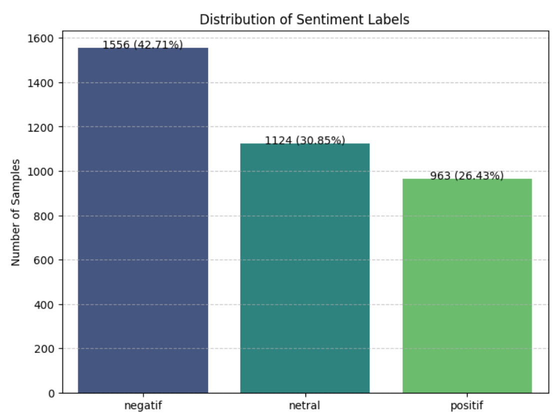

Figure 2. Data Labeling Result

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

3.4. Data Split

The data division stage was carried out systematically to ensure model integrity in the learning and testing processes, involving a total dataset of 3,643 entries. Before the separation, the researchers transformed the category labels into a numerical format using the LabelEncoder instrument so that they could be processed by the deep learning algorithm. The data was then divided into three main subsets using the stratified splitting technique with the parameter random_state=42 to maintain a balanced label proportion in each part. Based on the execution results represented in Figure 3.

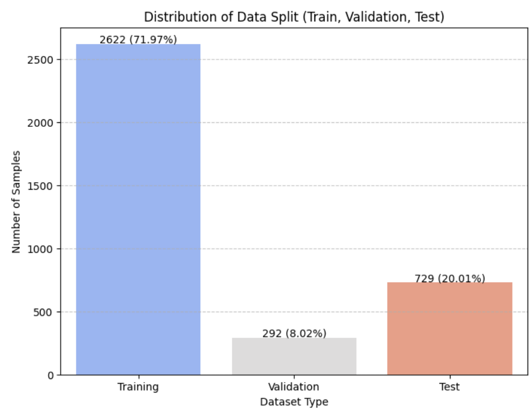

Figure 3. Data Split Result

Based on the figure, the data distribution allocated 2,622 samples, equivalent to 71.97%, as the Training Set, which was used as the main learning basis for the model to extract linguistic features. In addition, 292 samples or 8.02% were set aside as the Validation Set, which serves as a control parameter for hyperparameter optimization during the training process. As a final testing stage, 729 samples or 20.01% were designated as the Test Set, which is independent to test the model's generalization ability against new data that has never been encountered before.

3.5. Modeling

At this stage, the modeling process is carried out by applying the IndoBERT model described in the Methods chapter. The model is trained using training data that has undergone preprocessing and tokenization, with parameter configurations according to the research design. The fine-tuning process runs for several epochs until the model reaches learning stability, then continues with testing on test data to obtain the final performance. This stage ensures that the model has adapted to the characteristics of the data and is ready to be evaluated in the next stage.

Jurnal Teknologi Informasi dan Pendidikan
Volume 19, No. 1, March 2026
https://doi.org/10.24036/jtip.v19i1.1104

3.6. Model Evaluation

The evaluation was conducted to measure the generalization ability of the IndoBERT model against 729 test data. Based on the classification report results shown in Figure 4

Classification Report:

	precision	recall	f1-score	support
0	0.7977	0.8652	0.8301	319
1	0.7910	0.7227	0.7553	220
2	0.8187	0.7842	0.8011	190
accuracy			0.8011	729
macro avg	0.8025	0.7907	0.7955	729
weighted avg	0.8012	0.8011	0.8000	729

Figure 4. Classification Report Results

Based on the image, The model achieved 80% accuracy, with an average precision, recall, and F1-score of 0.80. These values indicate that the model is capable of consistently classifying sentiment across all classes. A review by class shows that negative sentiment performed best with an F1-score of 0.8301 and a recall of 0.8652. Positive sentiment also shows good performance with an F1-score of 0.8011, while neutral sentiment obtains an F1-score of 0.7553. Although it is the class with the lowest score, its performance is still quite stable considering the neutral language characteristics on social media, which tend to be ambiguous.

The distribution of predictions made by the model on the test data can be observed through the Confusion Matrix visualization. This is shown in Figure 5 below.

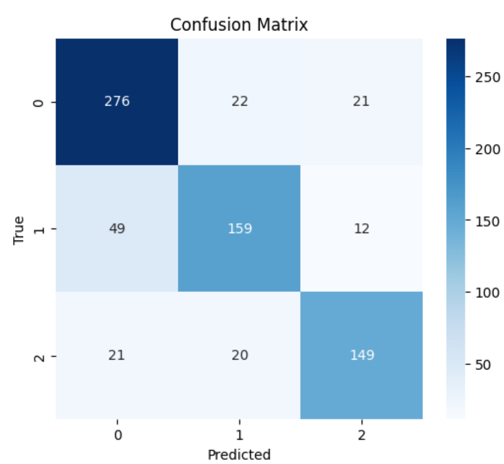

Figure 5. Confusion Matrix Results

P.ISSN: 2086 – 4981
E.ISSN: 2620 – 6390
tip.ppj.unp.ac.id

Jurnal Teknologi Informasi dan Pendidikan
Volume 19, No. 1, March 2026
https://doi.org/10.24036/jtip.v19i1.1104

Based on the figure, the model was able to accurately predict 276 negative data, 159 neutral data, and 149 positive data according to their actual labels. However, there are linguistic challenges evident from the 49 neutral data that were misclassified as negative sentiment. This reflects the characteristics of social media texts, where informative or objective news sentences often use strong diction, causing the machine to mistakenly categorize them into the negative sentiment group.

To provide a more comprehensive picture of the stability and convergence process of the model during the fine-tuning phase, the researchers conducted an in-depth analysis of the learning patterns recorded through metric development graphs, as shown in Figure 6.

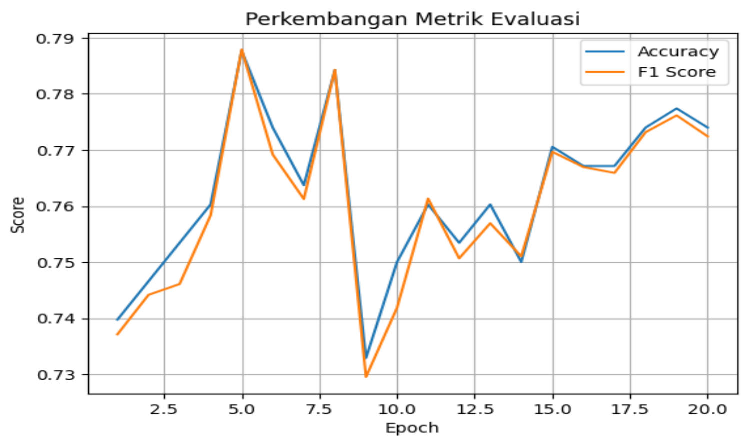

Figure 6. Evaluation Matrix Development Results

Based on this figure, the model shows a dynamic performance improvement trend from the beginning of the training stage to the end of the iteration in the 20th epoch. Accuracy and F1-score started at a value of 0.74 and experienced significant fluctuations as the number of epochs increased. The highest performance peak on the validation data was seen to have reached a figure close to 0.79 in the 5th epoch, before finally stabilizing in the range of 0.77 to 0.78 until the last epoch. The consistency and stability of these metrics at the end of the iteration provide a strong indication that the model has successfully identified the optimal semantic patterns for performing sentiment classification tasks.

In line with the development of these metrics, the dynamics of loss in the Training vs. Validation Loss curve show contrasting learning characteristics between the training and validation datasets, as shown in Figure 7.

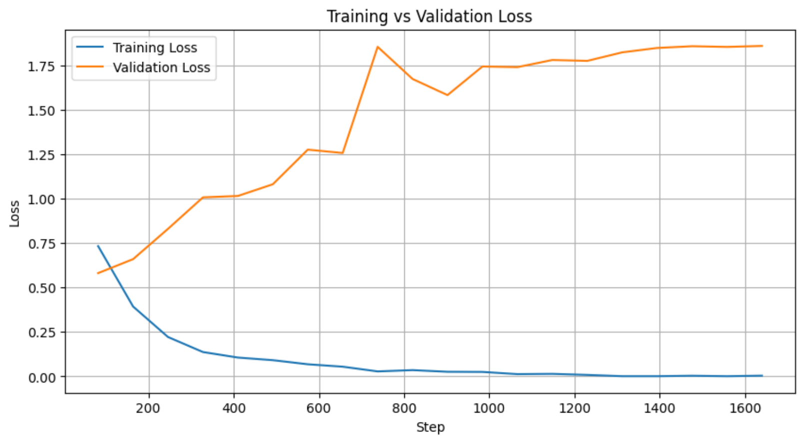

Figure 7. Training and validation loss Results

Based on Figure 7, the Training vs. Validation Loss curve shows a sharp contrast in learning characteristics between the training dataset and the validation dataset. The training loss value decreases drastically and consistently to near zero, proving that the IndoBERT model is very effective in absorbing linguistic features from the training data. However, on the other hand, the validation loss graph shows a steady upward trend as the number of iterations (steps) increases, from a range of 0.6 to above 1.75. Although this upward trend reflects significant linguistic challenges in social media text, the IndoBERT model still proves its superiority by maintaining an accuracy of 80.11%. This performance far exceeds other baseline methods, where the SVM model only achieves an accuracy of 70.18% and Naive Bayes 65.35%, as shown in Table 6.

Table 6 comparison table

Model	accuracy	Precision	Recall	F1-Score
IndoBERT	80,11 %	80,12%	80,11%	80%
SVM	70,18%	70%	70%	70%
Naïve Bayes	65,35%	68%	65%	62%

This significant performance gap confirms that despite the risk of overfitting, the transformer architecture in IndoBERT has much stronger semantic robustness in classifying complex sentiments compared to traditional machine learning methods. Furthermore, the high F1-score of 0.8301 for negative sentiment indicates that public criticism of programs tends to use explicit diction, which has practical implications for the government to utilize this system as a real-time policy evaluation tool to monitor digital aspirations and improve communication transparency.

4. CONCLUSION

This study concludes that the implementation of the IndoBERT model has proven effective in analyzing public sentiment related to the "Free Nutritious Meals" program on the X platform, with an accuracy rate of 80.11% and a weighted average F1-Score of 0.8000. This model shows very sharp performance in detecting negative sentiment with an F1-Score value of 0.8301, where overall public discussion is dominated by neutral and positive sentiment, reflecting high public expectations and support for this policy. Despite providing satisfactory results, this study has major limitations in handling linguistic ambiguity in neutral sentiment and indications of overfitting risk, as reflected in the upward trend of validation loss in the training curve. As an improvement for future research, it is recommended to expand the scope of slang and sarcasm normalization at the pre-processing stage, as well as to apply stricter model regulation techniques to improve the system's generalization ability in monitoring public opinion dynamics more accurately and comprehensively.

REFERENCES

[1] Suardi and A. S. Purmadani, “Jurnal Bisnis Net Volume : 8 No . 1 Juni , 2025 | ISSN : 2621 -3982 Jurnal Bisnis Net Volume : 8 No . 1 Juni , 2025 | ISSN : 2621 -3982,” *Bisnis Net J. Ekon. dan Bisnis*, vol. 8, no. 1, pp. 323–327, 2025, doi: https://doi.org/10.46576/bn.v8i1.6450.

[2] M. Basit and H. Ramadani, “Analisis Implementasi Program Makan Bergizi Gratis Terhadap Perkembangan Ekonomi,” *J. Econ. Dev. Res.*, vol. 1, no. 2, pp. 49–54, 2025, doi: https://doi.org/10.71094/joeder.v1i2.105.

[3] S. Altolyto, Y. Umaidah, and R. I. Adam, “Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis Pada Media,” *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: https://doi.org/10.23960/jitet.v12i3.4902.

[4] A. A. Pramita and F. Nugraha, “Aplikasi Lazada Menggunakan Metode Naive Bayes,” *J. Digit*, vol. 14, no. 1, pp. 23–30, 2024, doi: https://doi.org/10.51920/jd.v14i1.362.

[5] W. Romadhona and R. A. Isnain, “Analisis Sentimen Pengguna Media Sosial Terhadap Kebijakan Kenaikan Pajak Hiburan Menggunakan Metode SVM (Support Vector Machine),” *J. Ilm. Penelit. dan Pembelajaran Inform.*, vol. 9, no. 4, pp. 2185–2195, 2024, doi: https://doi.org/10.29100/jipi.v9i4.5603.

[6] A. Syahrir, Harlinda, and F. Umar, “Analisis Sentimen Masyarakat Terhadap Kebijakan Pemerintah Vaksinasi Booster 2 Menggunakan Metode Naïve Bayes Classifier,” *Zo. J. Sist. Inf.*, vol. 4, no. 1, pp. 347–359, 2023, doi: https://doi.org/10.33096/busiti.v4i4.1835.

[7] R. D. Himawan and E. Eliyani, “Perbandingan Akurasi Analisis Sentimen Tweet terhadap Pemerintah Provinsi DKI Jakarta di Masa Pandemi,” *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 1, p. 58, 2021, doi: 10.26418/jp.v7i1.41728.

[8] A. B. Sasmita, B. Rahayudi, and L. Muflikhah, “Analisis Sentimen Komentar pada Media Sosial Twitter tentang PPKM Covid-19 di Indonesia dengan Metode Naïve Bayes,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 3, pp. 1208–1214, 2022, [Online]. Available: http://j-ptiik.ub.ac.id

[9] B. P. Aji, C. Sri, and K. Aditya, "Klasifikasi Sentimen Ulasan Produk pada Platform E-Commerce di Indonesia dengan Menggunakan Model Pre-Trained IndoBERT," *Build. Informatics, Technol. Sci.*, vol. 6, no. 4, 2025, doi: 10.47065/bits.v6i4.6968.

[10] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," in *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*, 2020, pp. 757–770. doi: 10.18653/v1/2020.coling-main.66.

[11] A. Alfarel, Hasmawati, Bunyamin, J. Teknologi, and A. Alfarel, "Sentiment Analysis of Telkom University using the Long Short- Term Memory and Word2Vec Feature Expansion," *J. Teknol. Inf. dan Pendidik.*, vol. 17, no. 2, pp. 468–481, 2024, doi: https://doi.org/10.24036/jtip.v17i1.

[12] S. R. Putri *et al.*, "Analisis Sentimen Publik terhadap Nadiem Makarim sebagai Mendikbudrisktek menggunakan Support Vector Machine (SVM) Public Sentiment Analysis of Nadiem Makarim as Minister of using Support Vector Machine (SVM)," *Sist. J. Sist. Inf.*, vol. 14, no. 2, pp. 826–834, 2025, doi: https://doi.org/10.32520/stmsi.v14i2.5067.

[13] A. S. A. Rafsanjani *et al.*, "Analisis Sentimen Ulasan Pengguna Aplikasi KitaLulus pada Google Play Store dengan menggunakan algoritma Support Vector Machine (SVM) Sentiment Analysis of User Reviews of the KitaLulus Application on Google Play Store using the Support Vector Machine (," *Sist. J. Sist. Inf.*, vol. 14, no. 5, pp. 2519–2530, 2025, doi: https://doi.org/10.32520/stmsi.v14i5.5519.

[14] D. A. Dzulhijjah *et al.*, "Perbandingan Metode Random Forest dan KNN pada Analisis Sentimen Twitter," *Smart Comp Jurnalnya Orang Pint. Komput.*, vol. 12, no. 3, pp. 767–772, 2023, doi: https://doi.org/10.30591/smartcomp.v12i3.5106.

[15] J. Sanjaya, B. Priyatna, S. S. Hilabi, T. Tukino, B. Priyatna, and S. S. Hilabi, "Analisis Sentimen Terhadap Opini Proyek Kereta Cepat Menggunakan Metode Naïve Bayes Classifier," *J. Fasilkom*, vol. 14, no. 1, pp. 263–270, 2024, doi: https://doi.org/10.37859/jf.v14i1.6598.

[16] W. Ramadlan *et al.*, "Analisis Sentimen Ulasan Pengguna Aplikasi Spotify Di Google Play Store Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 2, pp. 3600–3607, 2025, doi: https://doi.org/10.36040/jati.v9i2.13543.

[17] B. Wilie *et al.*, "IndoNLU : Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," in *AACL-IJCNLP 2020 - the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, Association for Computational Linguistics, 2020, pp. 843–857. doi: 10.18653/v1/2020.aacl-main.

[18] V. Anggraeni, Y. Sibaroni, and S. S. Prasetyowati, "Sentiment Analysis of 2024 Election Reviews in Twitter using BERT with Emoticon Feature Extraction," *J. Teknol. Inf. dan Pendidik.*, vol. 18, no. 2, pp. 955–967, 2025, doi: https://doi.org/10.24036/jtip.v18i2.993.

[19] U. Fathon, M. Arifin, and A. Setiawan, "Analisis Sentimen Ulasan Pengguna Aplikasi Crunchyroll , iQIYI , Wibuku di Google Play Store Menggunakan Metode Random Forest," *JUSIBI (JURNAL Sist. Inf. DAN E-BISNIS*, vol. 7, no. 2, pp. 63–72, 2025, doi: https://doi.org/10.54650/jusibi.v7i2.603.

[20] R. Apriliyanti, D. Kurniadi, D. Novaliendry, S. Rahmadika, and M. Farhan, "Real-Time Color Classification of Objects with an Improved MobileNetV2 CNN Model," *J. Teknol. Inf. dan Pendidik.*, vol. 18, no. 2, pp. 900–914, 2025, doi: https://doi.org/10.24036/jtip.v18i2.969.