JSCE (Journal of System and Computer Engineering)
Vol.6, No.2, April 2025, pp. 197-203
ISSN (online): 2723-1240
DOI : https://doi.org/10.61628/jsce.v6i2.1867

Sentiment Analysis of the 2024 Indonesian Presidential Election Using Transfomer (Distilbert-Base-Uncased)

Andi Aljabar [1,a]; Binti Mamluatul Karomah [2,b]; Nahla Tarisafitri [3,c]; Jeffry [4,d,*]
[1,2,3] Computer Engineering Department, Faculity of Engineering and Computer Science, University of Nahdlatul Ulama Indonesia Jakarta 10320, Indonesia
[4] Department of Production and Industrial Technology, Computer Science Study Program, Bacharuddin Jusuf Habibie Institute of Technology, Parepare 91125, Indonesia
[a] andialjabar@unusia.ac.id; [b] mamluatul93@unusia.ac.id; [c] nahlatafi@gmail.com; [d] jeffry@ith.ac.id
** Corresponding author*

Abstract

Utilizing a transformer-based natural language processing model called DistilBERT-base-uncased, this study investigates the use of sentiment analysis in relation to Indonesia's 2024 presidential election. Particularly during political events, sentiment analysis is a potent tool for gaining insight into public opinion. The program divides public posts' sentiment into positive and negative categories by examining social media data (twitter). In order to assure consistency and correctness, the dataset used in the research has been carefully selected. DistilBERT is then used to train the model. The result shows that, out of 19920 row of data, only 4.47% of Indonesian citizens posted positive comments.

Keywords— *Sentiment Analysis, DistilBERT, Transformer Model, Indonesian Presidential Election*

1. Introduction

Understanding public opinion could be facilitated by sentiment analysis, which is particularly useful in situations such as Indonesia's 2024 presidential election (Resti et al., 2024; Damayanti & Lhaksmana, 2024; Sulistianingsih & Switrayana, 2024; Baharuddin et al., 2022). Sentiment analysis enables researchers and political strategists to assess the general mood and attitudes toward candidates, parties, and important issues, as millions of citizens share their opinions on a variety of platforms (Damayanti & Lhaksmana, 2024; Sulistianingsih & Switrayana, 2024; Akpatsa et al., 2022). Sentiment analysis helps determine if public opinion is good or negative regarding political developments by examining large volumes of textual data from social media (Geni et al., 2023).

Sentiment analysis requires the use of sophisticated natural language processing (NLP) models, such as those proposed by Akpatsa et al. (2022), Berfu et al. (2020), and Igali et al. (2024). A more condensed variation of the well-known BERT model, DistilBERT provides enhanced language contextual awareness at a faster rate. Because the Indonesian language is typically written informally online and frequently uses both lowercase and uppercase letters, the uncased version is very helpful for this type of analysis (Nair et al., 2024). Accurate and effective public sentiment classification is made feasible by utilizing this model. Additional modeling techniques, such as convolutional neural networks (CNNs), have also demonstrated effectiveness in classification tasks, including in domains like food categorization and pharmaceutical safety, further supporting their potential in NLP-based sentiment analysis (Abasa et al., 2025; Ahyana et al., 2025).

Sentiment research can shed light on how various voter groups feel about candidates, policies, or political developments in the context of the 2024 presidential election. Analysts can identify patterns (Nair et al., 2024) and trends in public opinion by gathering these opinions from many web sources (Firdaus et al., 2024), which can assist in guiding campaign tactics. For example, candidates might have to speak out more about a certain subject if unfavorable sentiment is observed surrounding it. Positive comments in a similar vein might reveal which campaign themes are striking a chord with voters (Fattah & Ratnasari, 2023; Wahyudi et al., 2024; Palani et al., 2021).

2. Method

Figure 1 displays the research methodologies. Data collection is the first step. Utilizing the scraping technique from social media, data was gathered from Twitter (Firdaus et al., 2024). Data cleaning and preparation come next after data collection. Anomalous data includes things like duplicates and mismatched entries, among other issues (Palani et al., 2021).

Tokenization comes first when employing the BERT approach, followed by data preprocessing (Digitus et al., 2023). Next, training and validation are conducted using the uncased version of the DistilBERT model. Confusion matrix analysis is the following phase, after which the model is evaluated (Digitus et al., 2023; Iparraguirre-Villanueva et al., 2023).

To find the optimal accuracy and loss, this process is repeated while varying a few variables, such as batch size, number of epochs, and so on. Finally, based on the data received from Twitter, the model predicts the results and provides output regarding public opinion about the Indonesian presidential election of 2024 on Twitter.

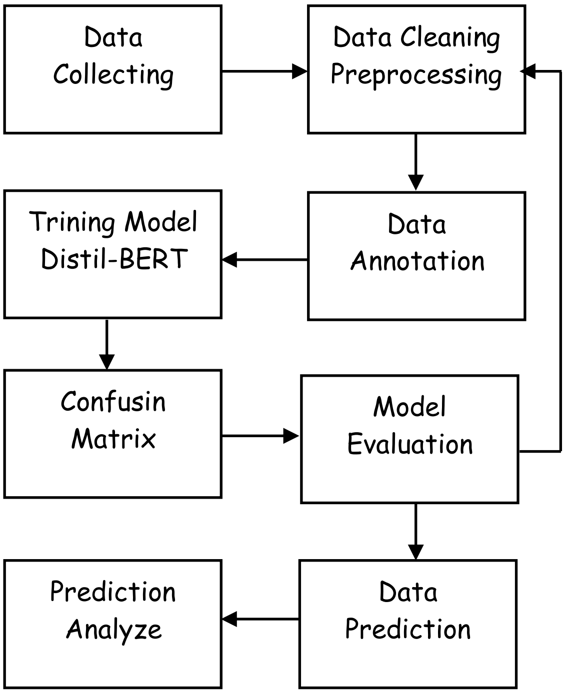

Figure 1 Research Method

3. Results And Discussion

The data was collected in every special moment, like debates performa, then the data rows are 19920 rows, based on the data gathering process. Data trining and validation accounted for 25% of the data. Of the 25% of data, 80% were utilized for trining, while the remaining 20% were used for validation. Model Distil-BERT-Uncaused predicted 75% of the remaining data.

DOI : https://doi.org/10.61628/jsce.v6i2.1867

In Figure 2, model architecture is displayed. The distilbert, pre_clissifier (dense), another classifier (dense), and dropout are the four laters. For the best model outcome in this research, only two epochs were used. It is indicated that the complete trinable parameters have 255.42 MB of data capacity (66.955.779).

Layer (type)	Output Shape	Param #
distilbert (TFDistilBertMainLayer)	multiple	66362880
pre_classifier (Dense)	multiple	590592
classifier (Dense)	multiple	2307
dropout_39 (Dropout)	multiple	0

Total params: 66955779 (255.42 MB)
Trainable params: 66955779 (255.42 MB)
Non-trainable params: 0 (0.00 Byte)

Figure 2 Model Architecture

Accuracy curves explained (figure 3) that the model's performance is showing positive trends in the plot, especially with its rising training accuracy. Training accuracy is shown by the red line, which rises gradually from 90% to over 94%, demonstrating how well the model is selecting and modifying the patterns in the training data. This implies that as training progresses, the model's ability to identify sentiment or categories is improving, indicating positive internal learning dynamics.

Even though the validation accuracy (blue line) increases more slowly across the training phase, staying slightly below 93%, it does so steadily. This stability indicates that there are No. significant fluctuations or losses in accuracy when the model is applied to unobserved data, which is a good sign of the model's robustness. The model is not badly overfitting and retains a great ability to generalize, as evidenced by the fact that the validation accuracy tart sat a high point and stays near to the training accuracy.

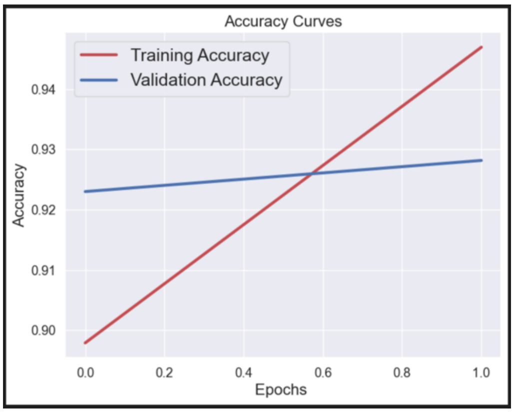

Figure 3 Accuracy Curves

Positive performance indicators are displayed on the graph (loss curves figure 4). The model appears to be learning efficiently when the training loss (red line) gradually drops. The model is still showing strong generalization on untested data, even though the validation loss (blue line) is somewhat increasing. Though it still performs well overall, the model is becoming better during training, and its output can still be further optimized.

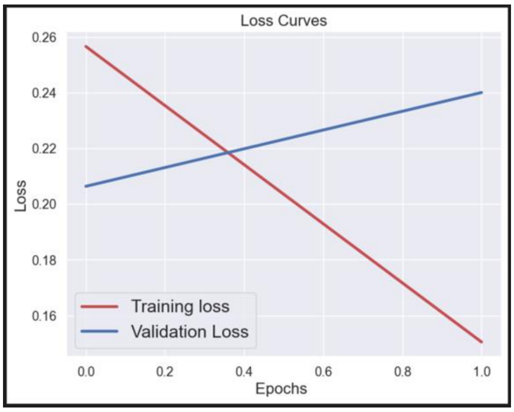

Figure 4 Loss Curves

An understanding of the model's classification performance can be obtained from the confusion matrix. In 157 cases, or "true positives," the algorithm accurately predicted a positive attitude; in 68 cases, or "true negatives," it properly labeled a negative sentiment (TN). False positives (FP) refer to the six instances in which the model predicted a positive sentiment when the actual sentiment was negative. Furthermore, in nine cases—referred to as false negatives (FN)—the model predicted a negative attitude when the true label was positive, failing to predict a positive sentiment.

Overall, the model performs well, demonstrating its effectiveness in detecting positive cases with 157 true positives and only 9 false negatives, especially in predicting positive emotion. The model appears to be dependable in preventing false positive predictions, as evidenced by the low number of false positives (6). The model's ability to identify negative sentiment is further demonstrated by the ratio of true negatives (68) to comparatively few false positives. Shown in the figure 5

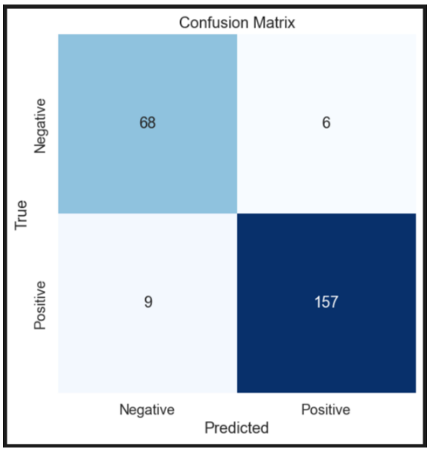

Figure 5 Confusion Matrix

Both classes demonstrate good performance from the model; Class 0 achieves a precision of 0.88, while Class 1 achieves a greater precision of 0.96. This suggests that Class 1 predicts positive events with a marginally higher degree of accuracy. Class 1 has a recall of 0.95 compared to 0.92 for Class 0 in terms of recall, indicating that Class 1 is more adept at identifying true positives. The F1 Scores of 0.95 for all courses are the same, indicating that each class performed well in terms of both recall and precision.

The F1 Score, which measures overall accuracy, is 0.94, suggesting a good degree of overall model performance. The average precision, recall, and F1 Score are 0.92, 0.93, and 0.93, respectively, according to the macro average measures, which treat every class equally regardless of how frequently it occurs. These numbers show that, in the absence of class imbalance, the model operates well in all classes. The precision, recall, and F1 Score of the weighted average metrics—which take into consideration the quantity of examples in each class—are all somewhat higher at 0.94. This implies that the model performs robustly throughout the full dataset and is especially successful when considering the support (number of occurrences) of each class. It is shown in the figure 6.

	precision	recall	f1-score	support
0	0.88	0.92	0.90	74
1	0.96	0.95	0.95	166
accuracy			0.94	240
macro avg	0.92	0.93	0.93	240
weighted avg	0.94	0.94	0.94	240

Figure 6 Classification Reprot

Figure 7 explains that most comments made by users on social media, particularly on Twitter, are classified as unfavorable. As per the results, only 4.47% of them left good comments.

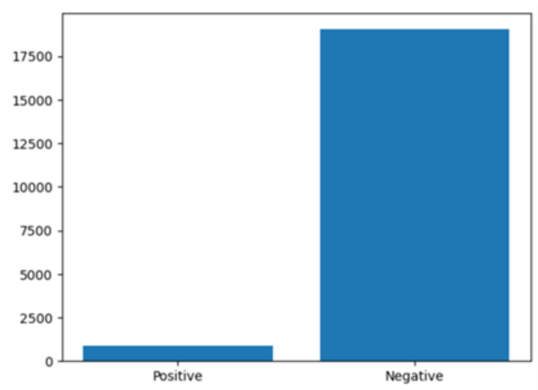

Figure 7 Prediction Analyze

Using a word cloud, figure 8—which supports figure 7—shows that most of the terms are in criticism. Words like "No More, Wakanda, Asal Bukan" and others are commonly used to give orders.

Figure 8 Words Cloud

4. Conclusions

In conclusion, according to the model shows the average of accuracy is 94% and 18% of loss average. The result of prediction of sentiment analysis of presential election 2024 in Indonesia explain that only 4.47% of data is categorize as positive comments on the twitter platform.

References

Abasa, S., Aziz, F., Ishak, P., & Jeffry. (2025). Detection of persistent vs non-persistent medications in pharmacy using artificial intelligence: Development of intelligent algorithms for pharmaceutical product safety. Journal of System and Computer Engineering (JSCE), 6(1), 137–143.

Ahyana, N., Jeffry, J., Fadliana, N., Rimalia, W., & Iskandar, I. (2025). CNN modeling for classification of Bugis traditional cakes. Journal of System and Computer Engineering, 6(1), 137–143. https://doi.org/10.61628/jsce.v6i1.1685.

Akpatsa, S. K., et al. (2022). Online news sentiment classification using DistilBERT. Journal of Quantum Computing, 4(1), 1–11. https://doi.org/10.32604/jqc.2022.026658

Baharuddin, T., Qodir, Z., Jubba, H., & Nurmandi, A. (2022). Prediction of Indonesian presidential candidates in 2024 using sentiment analysis and text search on Twitter. International Journal of Communication and Society, 4(2), 204–213. https://doi.org/10.31763/ijcs.v4i2.512

Berfu, B., Ali, H., Arzucan, A., & Arzucanozgür, A. A. (2020). Analyzing ELMo and DistilBERT on socio-political news classification. GitHub. https://github.com/huggingface/

Damayanti, L., & Lhaksmana, K. M. (2024). Sentiment analysis of the 2024 Indonesia presidential election on Twitter. Jurnal dan Penelitian Teknik Informatika, 8(2). https://doi.org/10.33395/v8i2.13379

Digitus, K., Diantoro, A., Soderi, A., Rohman, A., & Sitorus, A. T. (2023). Sentiment analysis of public opinion on the 2024 presidential election in Indonesia using Twitter data with the K-NN method. Unpublished manuscript.

Fattah, F. S., & Ratnasari, C. I. (2023). Sentiment analysis of Indonesian presidential candidate. Fardhan Saifullah Fattah. [Self-published].

Firdaus, A. A., Yudhana, A., & Riadi, I. (2024). Prediction of presidential election results using sentiment analysis with pre and post candidate registration data. Unpublished manuscript.

Geni, L., Yulianti, E., & Sensuse, D. I. (2023). Sentiment analysis of tweets before the 2024 elections in Indonesia using IndoBERT language models. Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI), 9(3), 746–757. https://doi.org/10.26555/jiteki.v9i3.26490

Igali, A., Abdrakhman, A., Torekhan, Y., & Shamoi, P. (2024, August). Tracking emotional dynamics in chat conversations: A hybrid approach using DistilBERT and emoji sentiment analysis. arXiv. http://arxiv.org/abs/2408.01838

Iparraguirre-Villanueva, O., et al. (2023, February). The public health contribution of sentiment analysis of monkeypox tweets to detect polarities using the CNN-LSTM model. Vaccines (Basel), 11(2). https://doi.org/10.3390/vaccines11020312

Nair, A. R., Singh, R. P., Gupta, D., & Kumar, P. (2024). Evaluating the impact of text data augmentation on text classification tasks using DistilBERT. In Procedia Computer Science (pp. 102–111). Elsevier B.V. https://doi.org/10.1016/j.procs.2024.04.013

Palani, S., Rajagopal, P., & Pancholi, S. (2021, June). T-BERT -- Model for sentiment analysis of micro-blogs integrating topic model and BERT. arXiv. http://arxiv.org/abs/2106.01097

Resti, J., Cahyanti, R., Maftuhah, D. N., Santoso, A. B., & Budi, I. (2024). Number 1 of 2026 candidates of the 2024 Indonesian presidential election. Teknologi Informasi (RESTI), 10(4), 516–524. https://doi.org/10.29207/resti.v8i4.5839

Sulistianingsih, N., & Switrayana, I. N. (2024). Enhancing sentiment analysis for the 2024 Indonesia election using SMOTE-Tomek Links and binary logistic regression. International Journal of Education and Management Engineering, 14(3), 22–32. https://doi.org/10.5815/ijeme.2024.03.03

Wahyudi, A., Santoso, G. B., & Sholihah, B. (2024, February). Analysis of the sentiment of Indonesian presidential candidates for 2024 on the YouTube social media platform using the support vector machine method. Intelmatics, 4(1), 22–30. https://doi.org/10.25105/itm.v4i1.17636