Smart Techno (Smart Technology, Informatic, and Technopreneurship)
Volume 08, Number 01, 2026 pp. 142-150
E-ISSN : 2541-0679
Open Access: https://lppm.primakara.ac.id/jurnal/index.php/smart-techno/

Sentiment Analysis of YouTube Comments for the Jumbo Movie Trailer Using IndoBERT

Fardan Zamakhsyari[1*], Rizka Suhana[2], Irfan Ramadhani[3], Dwi Priyo Santoso[4]
[1,2,3,4] Department of Informatics Engineering, University of Cahaya Surya, Kediri, Indonesia

ARTICLE INFO
Article history:
Initial submission
14-01-2026
Received in revised form
26-01-2026
Accepted 17-01-2026
Available online 27-02-2026

Keywords:
Sentiment Analysis, Crawl Data, IndoBERT, Confusion Matrix

DOI:
https://doi.org/10.59356/smart-techno.v8i1.198

ABSTRACT
The film industry in Indonesia has experienced significant growth, ranging from cinematography to animation. Along with this growth, public opinion has also varied, ranging from assessments of storylines to production processes. To analyze public sentiment on social media, a system that supports this process is needed. This study aims to analyze public sentiment towards the trailer for the animated film "Jumbo" released on YouTube. Using an NLP approach, two customized IndoBERT models were compared: 'Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis' and 'rikidharmawan/finetuning-sentiment-model-indobertweet-v2' and were fine-tuned again using the data obtained. The processed data was obtained from 1,468 YouTube comments through a crawling process using the YouTube API. The data was then analyzed using both models to classify comments into positive, neutral, and negative sentiments. The evaluation was carried out using a confusion matrix with accuracy, precision, recall, and F1-score metrics. The evaluation results showed that 'Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis' was superior, with an accuracy of 73.7% and a higher average F1 score compared to 'rikidharmawan/finetuning-sentiment-model-indobertweet-v2', which had an accuracy of 72.0%. This research contributes to the selection of sentiment analysis models for Indonesian-language data, particularly in the fields of social media and the film industry.

1. INTRODUCTION

In recent years, animated films have continued to evolve, with many new animations in the digital era, such as live-action films The Lord of the Rings, Transformers, Narnia, and The Hobbit, as well as 3D animations such as Madagascar, Finding Nemo, Cars, Toy Story 3, Monster Inc, Shrek, Ice Age, Brave, and others, being released almost continuously throughout the year. These animations are far more complex than those of previous eras (Fadly et al., n.d.). Animation has undergone significant development throughout its history, beginning in the early 20th century, when silent films gained popularity. The creation of animated content encompasses various aspects, including the use of sound and visuals, as well as the application of modern technology (Audi et al., 2024). Even in Indonesia, an animated film with extraordinary animation, Jumbo, has been successfully produced.

The release of this film was accompanied by various public sentiments. In this study, we used YouTube, one of the main social media platforms where many comments and public opinions are expressed. The comments section on YouTube is a place for users to respond to uploaded content. Sentiment analysis of YouTube comments can be an alternative way to understand public opinion and sentiment towards a piece of content (Oswari et al., 2024). Movie trailers, as an initial representation of a cinematic work, are an interesting focal point for analysing how audiences respond before the film is fully released. Understanding the sentiment in trailer comments can provide valuable insights for filmmakers, distributors, and marketers in gauging public expectations and the potential reception of a film (Hutto & Gilbert, 2014).

Sentiment analysis, a field within natural language processing (NLP), aims to identify and extract subjective opinions from text. Traditional methods in sentiment analysis often rely on lexicon-based approaches or conventional machine learning. However, the emergence of transformer models, such as Bidirectional Encoder Representations from Transformers (BERT), has revolutionised the field of NLP, including sentiment analysis (Zamakhsyari & Fatwanto, 2025). These models, which are pre-trained on a large scale using massive text corpora, can capture context and linguistic nuances more effectively, resulting in superior performance in various text classification tasks (Wu & Dredze, 2020).

This study focuses on analysing the sentiment of YouTube comments on the trailer for the movie Jumbo. With its premise and potential visual appeal, the movie Jumbo is likely to elicit a variety of reactions and opinions from viewers. Understanding the sentiment in the comments on this movie trailer can provide an initial picture of public expectations and sentiment toward the film. Furthermore, this study aims to compare the efficiency and effectiveness of two different IndoBERT model variants after undergoing a fine-tuning process for sentiment classification tasks. The two IndoBERT models to be compared are “Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis” and “rikidharmawan/finetuning-sentiment-model-indobertweet-v2”. This comparison will be assessed based on the resulting sentiment classification performance, which will be evaluated using a confusion matrix. The confusion matrix provides a visualisation of the classification model's performance, including metrics such as accuracy, precision, recall, and F1-score, which are important for understanding the strengths and weaknesses of each model (Zamakhsyari et al., 2025).

By comparing two fine-tuned IndoBERT models, this study aims to provide insights into which model is more effective and efficient in analysing the sentiment of Indonesian-language comments in the context of movie trailers. The results of this study have the potential to contribute to the understanding of transformer model applications in Indonesian-language sentiment analysis, particularly in the domain of film, as well as provide practical guidance in selecting the appropriate model for similar tasks.

2. METHOD

This study was conducted to analyse public sentiment in YouTube comments on the Jumbo trailer. The models used were “Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis” and “rikidharmawan/finetuning-sentiment-model-indobertweet-v2.” In the research process, there will be several stages to go through to obtain maximum results. The research methods and flow are shown in Figure 1.

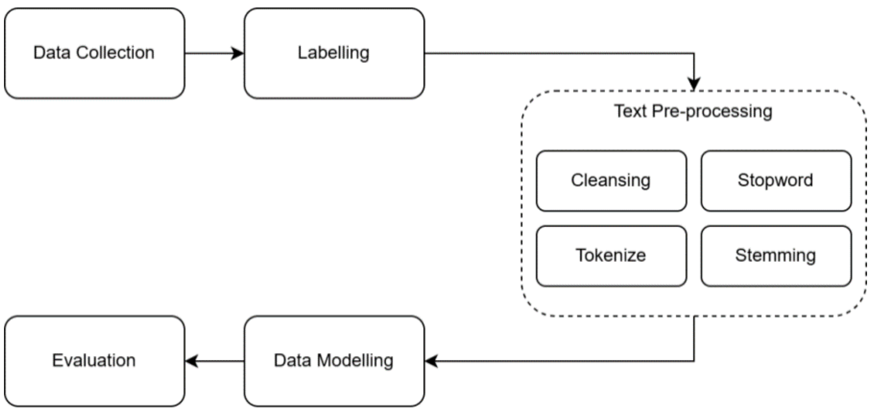

Figure 1. Research Process

2.1. Data Collection

The initial stage of this research is collecting comment data from the Jumbo movie trailer on YouTube. These comments will be the main source of textual data for sentiment analysis. The data collection method will involve web scraping using appropriate libraries or tools to extract comments from the YouTube page of the Jumbo movie trailer. The collected data will include comment text and, if possible, additional information such as the time the comment was published and the number of likes.

At this stage, the process is carried out using YouTube's Data API to access comments on specific videos. The tools used for data crawling are Visual Studio Code with the Python language (Tohidi et al., 2024). The results obtained from data crawling will contain comments, the number of likes, and the date the comments were made. These comments are collected based on when the video was first uploaded. Next, the files are saved in CSV format to facilitate further data processing (Kristiana et al., 2023).

2.2. Data Labelling and Normalization

After the comment data has been collected, the next step is to label the sentiment of each comment. The purpose of this labelling is to assign an appropriate sentiment category to each comment, such as positive, negative, or neutral. The labelling process is done manually by annotators who have an understanding of the Indonesian language context and sentiment nuances (Merdiansah & Ali Ridha, 2024). In the process, we also perform data normalization. Data normalization is a term used to convey ideas by converting text formats for predetermined purposes. In the field of Natural Language Processing, many text normalization processes depend on their use, namely: i) the type of text entered; ii) the desired output format; iii) the purpose of text normalization; iv) the methods used in the text normalization process (Nur & Prasetija, 2022).

2.3. Text Pre-Processing and Handling Imbalanced Data

The collected data will be reprocessed in the pre-processing stage to maximise the results in data processing. In this study, several pre-processing stages will be carried out (Vidya Chandradev et al., 2023), the first of which is cleaning, by removing irrelevant characters such as URLs, hashtags, mentions, special symbols, and numbers that have no meaning. Next is stopword removal, at this stage, common words that often appear and are considered to have no significant meaning, such as “is”, ‘and’, “which”, and so on, will be removed. After that, the Tokenise stage is carried out, which breaks the text into tokens. Every word or punctuation mark used in the comments will be considered a token. Finally, stemming is carried out to change words into their basic form (root words). After the data is processed at this stage, it will then proceed to the data modelling process to assess sentiment based on the model used (Qiu et al., 2020). The collected data is imbalanced. This can cause problems in machine learning modelling because the model tends to better predict classes that appear more frequently than classes that appear less frequently, so the data needs to be handled. In this study, to handle imbalanced data, the class weight method is used by giving higher weights to samples from minority classes (Sains et al., 2023).

2.4. Data Modelling and Fine-tuning

At this stage, the data that has been collected and pre-processed will be processed using the two IndoBERT models mentioned earlier to perform sentiment analysis. This study will use two IndoBERT models, namely the IndoBERT Model from Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis and the IndoBERT Model from rikidharmawan/finetuning-sentiment-model-indobertweet-v2. Both models are fine-tuned versions of IndoBERT and are specifically designed for sentiment analysis tasks on Indonesian language datasets.

In this study, both BERT models will be fine-tuned for sentiment analysis tasks on the “Jumbo” movie trailer dataset. The training data consists of a large number of reviews of the movie trailer “Jumbo” that have been labelled as positive, negative, or neutral sentiment. After the BERT model undergoes the fine-tuning process, the next step is to test it on the dataset. To evaluate the model's performance, various evaluation metrics will be calculated, such as accuracy, precision, recall, and F1-score (Sjoraida et al., 2024).

2.5. Data Evaluation

After the data modelling process is complete, to calculate the performance results of the two IndoBERT models, an evaluation process will be carried out using a confusion matrix. Calculations using a confusion matrix will provide information about the number of correct and incorrect predictions for each sentiment class (positive, negative, neutral) (Putri, 2020). Based on the confusion matrix calculation, several evaluation metrics will be calculated to measure the effectiveness and efficiency of both models, including:

Table 1. Confusion Matrix Table

Student Name	**Actual Values**	
	Positive	**Negative**
Positive	TP	FN
Negative	FN	TN

Table 1 explains the confusion matrix assessment, where TP (True Positive) is the number of correct predictions in the positive class, and FP (False Positive) is the number of incorrect predictions in the positive class. Next is FN (False Negative), which is the number of incorrect predictions in the negative class, and finally TN (True Negative), which is the number of correct predictions in the negative class (Savana et al., 2025).

$$Precision = \frac{TP}{TP+FP} \qquad (1)$$

The above formula is used to calculate the Precision (p) value, which is the number of correct predictions in the positive class divided by the number of actual positives in the study.

$$Recall = \frac{TP}{TP+FN} \qquad (2)$$

Meanwhile, the Recall value can be obtained by assessing the number of correct predictions in the positive class divided by the number classified as positive.

$$Precision = \frac{TP+TN}{TP+TN+FP+FN} \qquad (3)$$

Finally, accuracy is obtained by assessing the ratio of correct predictions (positive and negative) to the total data. Thus, accuracy can determine the correct predictions from the available data (Maula et al., 2023).

3. RESULT AND DISCUSSION

3.1. Data Evaluation

The dataset used in this study was collected from YouTube videos using YouTube crawling techniques. This method is a process of retrieving data from the YouTube platform aimed at collecting information from YouTube videos (Mutiara Puspita). Using an Application Programming Interface (API) facilitates the process of scraping data from YouTube comments. The data collected consists of comments from the trailer video for the animated film Jumbo. This video is titled “Trailer Film Jumbo” and was released by Visinema Pictures. A total of 1,453 comments from the video were successfully retrieved during this research process. The collected data is stored in a CSV file and will proceed to the labelling and text pre-processing stages.

3.2. Data Labelling and Normalization

Data labelling on the collected comments was done manually, by assessing the meaning based on words and sentences. 1,453 comment data will be labelled with 3 variables, namely positive, neutral, and negative. This data will later be stored in CSV format to be continued in the text pre-processing stage. A comparison of the data amounts for each variable is presented in Table 2.

Table 2. Data Labelling Results Table

Total Sentiment Labelling		
Positive	**Negative**	**Netral**
767	46	637

The sentiment analysis model is highly sensitive to label inconsistencies, so data normalization is performed during the process to reduce ambiguity and noise in the training data. Data normalization is done by removing missing or ambiguous sentiment label data to maintain data quality.

3.3. Text Pre-Processing and Handling Imbalanced Data

The collected data will be processed in several stages beforehand, so that later, when the model is applied, maximum results can be achieved. In the text pre-processing section, several stages of adjustment will be carried out, starting from cleaning, stopword removal, tokenisation, and stemming. The processed data results are shown in Table 3.

Table 3. Text Pre-processing Data Results

Stages	Result
Initial Data	animasinya jadi ingat film animasi up karakternya masih kebaratan kalo saya liat ya tp semoga sukses dengan pencapaiannya maju terus film Indonesia.
Cleaning Data	animasinya jadi ingat film animasi up karakternya masih kebaratan kalo saya liat ya tp semoga sukses dengan pencapaiannya maju terus film Indonesia.
Stopword Removal	animasinya ingat film animasi up karakternya kebaratan kalo liat tp semoga sukses pencapaiannya maju film Indonesia.
Tokenize	['animasinya', 'ingat', 'film', 'animasi', 'up', 'karakternya', 'kebaratan', 'kalo', 'liat', 'tp', 'semoga', 'sukses', 'pencapaiannya', 'maju', 'film', 'indonesia']
Stemming	['animasi', 'ingat', 'film', 'animasi', 'up', 'karakter', 'barat', 'kalo', 'lihat', 'tp', 'moga', 'sukses', 'capai', 'maju', 'film', 'indonesia']

Table 3 provides an overview of the flow in text pre-processing. By following the steps above, each comment can be assessed on a word-by-word and sentence-by-sentence basis, thereby facilitating data modelling and evaluation.

After the data goes through the pre-processing stage, the next step is to handle the data imbalance. Class weights are calculated using a balanced weighting scheme, which assigns a higher loss penalty to underrepresented classes. These weights are incorporated into the loss function during model refinement, ensuring that minority classes, such as neutral, are well learned by the model. This approach allows the model to minimize bias toward dominant classes and improves overall classification fairness.

3.4. Fine-Tuning Model and Data Evaluation

Data that has undergone labelling and text pre-processing will be processed and evaluated using sentiment analysis with the IndoBERT model. Two IndoBERT models will be used in this study, namely the IndoBERT Model from Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis and the IndoBERT Model from rikidharmawan/finetuning-sentiment-model-indobertweet-v2. These two models will be compared based on their performance in providing sentiment analysis. The model calling process can be seen in Figure 2.

```
model_a = pipeline("sentiment-analysis", model="Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis")
model_b = pipeline("sentiment-analysis", model="rikidharmawan/finetuning-sentiment-model-indobertweet-v2")
```

Figure 2. Load Model IndoBERT

Both models were fine-tuned using the same experimental configuration to ensure a fair comparison. The models were trained using a learning rate of 2e-5, batch size of 16, and trained for 5 epochs. Tokenization was performed using the corresponding IndoBERT tokenizer with a maximum sequence length of 128 tokens. Stratified train-test splitting (80:20) was applied to preserve the sentiment distribution across training and testing sets.

After the data is retrieved and processed, the results of both models will be evaluated using the confusion matrix method to assess precision, recall, and F1-Score based on the labels provided. The results of this comparison will provide an overview of each model's performance in processing sentiment analysis using the Jumbo movie trailer comment dataset.

Table 4. Confusion Matrix Test Results Table

Model	Evaluasi	Precision	Recall	F1 Score
Model A	Accuracy	-	-	**0.7379**
	Macro Avg	**0.5449**	**0.5463**	**0.5449**
	Weighted Avg	**0.7404**	**0.7379**	**0.7382**
Model B	Accuracy	-	-	0.7207
	Macro Avg	0.5697	0.6147	0.5833
	Weighted Avg	0.7357	0.7207	0.7251

Table 4 explains the evaluation results conducted on both models. Based on the evaluation results conducted on both models, model A (“Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis”) performed better than model B (“rikidharmawan/finetuning-sentiment-model-indobertweet-v2”). From the accuracy test results, model A obtained a score of 73.7%, while model B obtained a score of 72.0%. These results were obtained by evaluating split data with a ratio of 80:20, using a confusion matrix testing model.

3.5. Discussion

The evaluation results show the performance of both models, where model A has a better performance than model B, with an accuracy value of 73.7%. Looking at the sentiment analysis results, there are several significant differences in the test results, especially in the classifications produced.

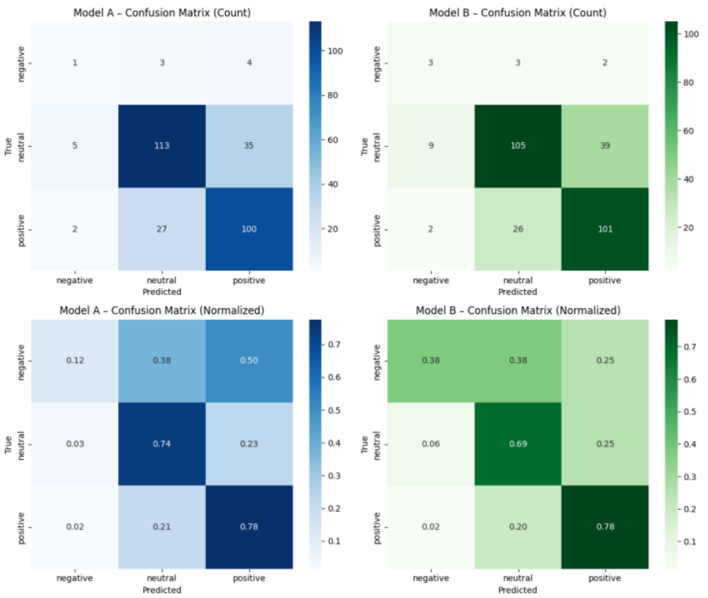

Figure 3. Confusion Matrix Model A and Model B

Figure 3 presents the confusion matrix of Model A and Model B, displayed in both absolute count and normalized form. The normalized confusion matrix illustrates the proportion of correctly and incorrectly classified instances for each sentiment class, allowing a fair comparison between models despite class imbalance.Model A shows good performance in classifying positive sentiment, with an accuracy rate of 78%. This is followed by an accuracy rate of 74% for correctly categorized neutral sentiment. Meanwhile, negative sentiment only has an accuracy rate of 12% in its predictions, and 50% of these are miscategorized as positive, indicating poor performance in the negative class. Thus, it can be said that Model A often misinterprets negative phrases as positive, possibly due to subtle emotional signals or a lack of negative training data.

Slightly different from Model A, Model B has an accuracy of 38% on correctly classified negative samples. Model B outperforms Model A in handling the negative sentiment class. The neutral class achieves an accuracy of 69%, while the positive class maintains a high accuracy of 78%. Model B shows more balanced categorization across all sentiment categories, particularly by reducing severe classification errors on negative samples, although neutral performance is slightly lower than Model A.

Model B has more balanced performance across all sentiment groups, based on a comparison between the two models. Model B significantly improves the classification of negative emotions, although both models are comparable in recognizing positive sentiment. This suggests that improving the model's sensitivity to minority or more complex sentiment expressions requires fine-tuning. The impact of class imbalance in the dataset is further shown by the confusion matrix. Because the neutral class predominates in the sample, its prediction accuracy is higher than that of the negative class, which has lower recall. This problem is lessened by using class weighting during fine-tuning, especially in Model B, which improves the identification of sentiment classes that are underrepresented.

The confusion matrix shows that Model A performs poorly on the negative sentiment class despite achieving a slightly higher overall accuracy. Model B, on the other hand, performs more evenly across all emotion categories, especially by greatly enhancing the identification of negative sentiment. This implies that assessing sentiment classification ability on unbalanced datasets requires more than just overall accuracy. For real-world applications where minority emotion recognition is crucial, Model B is therefore thought to be more dependable.

4. CONCLUSION

This study successfully analyzed the sentiment of YouTube user comments on the trailer for the animated film “Jumbo” using a transformer-based approach, specifically IndoBERT. Two fine-tuned IndoBERT models were compared, namely IndoBERT (Model A) Aardiiiiy/indobertweet-base-Indonesian-sentiment-analysis and IndoBERT (Model B) rikidharmawan/finetuning-sentiment-model-indobertweet-v2. The sentiment analysis process consisted of several stages, including data collection, manual sentiment labelling, text preprocessing, data imbalance handling, model refinement, and evaluation using confusion matrix analysis.

The data set was gathered from 1,453 YouTube comments on the Jumbo movie trailer that was posted to the Visinema Pictures YouTube channel using the YouTube API. According to the evaluation results, Model A's total accuracy (73.7%) was somewhat greater than Model B's (72.0%). Confusion matrix research, however, demonstrated that Model B performed more evenly across all sentiment classes, particularly when it came to detecting minority sentiment classes. This result emphasizes that assessing sentiment classification effectiveness on unbalanced datasets requires more than just overall accuracy.

This study is limited by the scope of the dataset, as the data was collected from a single YouTube video. For further research, several improvements can be considered: (1) expanding data collection to include multiple trailers or channels to increase data diversity and (2) exploring alternative deep learning architectures such as LSTM, BiLSTM, or hybrid transformer models to improve sentiment classification performance.

5. REFERENCES

Audi, A., Saputra, I. M. A. A., & Pramayasa, I. M. H. M. (2024). PERKEMBANGAN GAYA ANIMASI DUNIA DAN INDONESIA. Anima Rupa, 1(2), 72–78. https://doi.org/10.59997/animarupa.v1i2.3589

Fadly, M., Kom, S. I., & Kom, M. I. (n.d.). Analisis Perkembangan Animasi Indonesia Dalam Konten Animasi Dunia. In profilmjournals.com. http://profilmjournals.com/wp-content/uploads/2022/11/Analisis-Perkembangan-Animasi-Indonesia-Dalam-Kontels-Animasi-Dunia-Fadly.pdf

Hutto, C. J., & Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for. Eighth International AAAI Conference on Weblogs and Social Media, 216–225. https://ojs.aaai.org/index.php/ICWSM/article/view/14550

Kristiana, U., Rizqi Lahitani, A., Budi Setiawan, C., & Alfi Sa'diya, N. (2023). Analisis Forensik Digital Pada Komentar Youtube Live Menggunakan Sentiment Analysis. Teknomatika: Jurnal Informatika Dan Komputer, 15(1), 23–27. https://doi.org/10.30989/teknomatika.v15i1.1115

Maula, M. A., Wibowo, A. T., & Izzuddin, M. A. (2023). Eye Aspect Ratio based on Histogram Oriented Gradient and Linear Support Vector Machine to Microsleep Detection. Jurnal Matics, 15.

Merdiansah, R., & Ali Ridha, A. (2024). Sentiment Analysis of Indonesian X Users Regarding Electric Vehicles Using IndoBERT. Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI, 7(1), 221–228.

Nur, R., & Prasetija, S. (2022). Analisis Pengaruh Normalisasi Teks pada Klasifikasi Sentimen Ulasan Produk Kecantikan. E-Proceeding of Engineering, 9(3).

Oswari, T., Murniyati, M., Yusnitasari, T., Nurasiah, N., & Wijay, S. (2024). Sentiment Analysis of Indonesian Youtube Reviews About Lesbian, Guy, Bisexual and Transgender (LGBT) using IndoBERT Fine Tuning. Lontar Komputer : Jurnal Ilmiah Teknologi Informasi, 15(1), 26. https://doi.org/10.24843//lkjiti.2024.v15.i01.p03

Putri, C. A. (2020). Analisis Sentimen Review Film Berbahasa Inggris Dengan Pendekatan Bidirectional Encoder Representations from Transformers. JATISI (Jurnal Teknik Informatika Dan Sistem Informasi), 6(2), 181–193. https://doi.org/10.35957/jatisi.v6i2.206

Qiu, X. P., Sun, T. X., Xu, Y. G., Shao, Y. F., Dai, N., & Huang, X. J. (2020). Pre-trained models for natural language processing: A survey. Science China Technological Sciences, 63(10), 1872–1897. https://doi.org/10.1007/s11431-020-1647-3

Sains, J., Akbar, H., Selatan, T., & Sanjaya, W. K. (2023). Kajian Performa Metode Class Weight Random Forest pada Klasifikasi Imbalance Data Kelas Curah Hujan. Jurnal Sains, Nalar, Dan Aplikasi Teknologi Informasi, 3(1). https://doi.org/10.20885/snati.v3i1.30

Savana, B. R. K., Arifianto, D., & Muharom, L. A. (2025). SENTIMENT ANALYSIS OF COMMENTS ON INDONESIAN POLITICAL SPEECH VIDEOS ON YOUTUBE USING. Smart Techno (Smart Technology, Informatic, and Technopreneurship), 07(02), 34–44. https://doi.org/https://doi.org/10.59356/smart-techno.v7i2.159

Sjoraida, D. F., Guna, B. W. K., & Yudhakusuma, D. (2024). Analisis Sentimen Film Dirty Vote Menggunakan BERT (Bidirectional Encoder Representations from Transformers). Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi), 8(2), 393–404. https://doi.org/10.35870/jtik.v8i2.1580

Tohidi, E., Perdana Herdiansyah, R., Wahyudin, E., & Kaslani, K. (2024). Analisa Komentar Video Youtube di Channel TV One News Tentang Calon Presiden Prabowo Subianto Menggunakan Support Vector Machine. JATI (Jurnal Mahasiswa Teknik Informatika), 8(1), 660–667. https://doi.org/10.36040/jati.v8i1.8560

Vidya Chandradev, I Made Agus Dwi Suarjaya, & I Putu Agung Bayupati. (2023). Analisis Sentimen Review Hotel Menggunakan Metode Deep Learning BERT. Jurnal Buana Informatika, 14(02), 107–116. https://doi.org/10.24002/jbi.v14i02.7244

Wu, S., & Dredze, M. (2020). Are all languages created equal in multilingual BERT? Proceedings of the Annual Meeting of the Association for Computational Linguistics, July, 120–130. https://doi.org/10.18653/v1/2020.repl4nlp-1.16

Zamakhsyari, F., & Fatwanto, A. (2025). Language model optimization for mental health question answering application. International Journal of Electrical and Computer Engineering (IJECE), 15(5), 4829–4836. https://doi.org/10.11591/ijece.v15i5.pp4829-4836

Zamakhsyari, F., Makayasa, B. A., & Hamami, R. A. (2025). Comparison of KNN and Random Forest Algorithms on E- Commerce Service Chatbot. Jurnal Informatika Sunan Kalijaga, 10(1), 100–109. https://doi.org/10.14421/jiska.2025.10.1.100-109