

Sentiment Analysis of Public Opinion on Deforestation in Papua on YouTube Platform Using Long Short-Term Memory (LSTM) Method

Dolly Indra^{1✉}, Ramdaniah², Nada Kayatri Ode³

¹ Departement of information Systems, Computer Science, Universitas Muslim Indonesia, Indonesia

^{2,3} Department of Informatic Engineering, Computer Science, Universitas Muslim Indonesia, Indonesia

✉ **Corresponding Author** : Nada Kayatri Ode (e-mail: dolly.indra@umi.ac.id)

Article Information	ABSTRACT
Article History Received : August 15, 2025 Revised : September 02, 2025 Published : October 04, 2025	Deforestation in Papua has emerged as a significant environmental concern, attracting considerable attention due to its effects on biodiversity and the livelihoods of indigenous communities. This study seeks to examine public sentiment toward the issue by analyzing comments posted on the YouTube platform, employing the Long Short-Term Memory (LSTM) method. A dataset of 3,000 comments was gathered and processed through several stages, including text cleaning, tokenization, normalization, and Term Frequency–Inverse Document Frequency (TF-IDF) weighting. Subsequently, an LSTM model was developed and assessed using accuracy, precision, recall, and F1-score as evaluation metrics. The results reveal that the LSTM model achieved an accuracy of 88.43%, a precision of 90.01%, a recall of 97.49%, and an F1-score of 93.60%. Nevertheless, signs of overfitting were observed, indicated by lower validation performance compared to training results. These findings demonstrate that the LSTM approach is effective for identifying public opinion regarding deforestation and can serve as a valuable reference in decision-making and the formulation of environmental policies.
Keywords: <i>Analisis Sentimen; LSTM; Pembabatan Hutan Papua; TF-IDF; YouTube.</i>	

INTRODUCTION

Deforestation in Papua is one of the most important and complex environmental issues. The Papua region is known for its high biodiversity and its strategic role in maintaining the balance of global ecosystems (Kafiar et al., 2025). However, the exploitation of natural resources, especially deforestation, has caused significant negative impacts, including a decrease in environmental quality, loss of natural habitats, and social conflicts in local communities (Rinjani et al. 2024).

Along with the development of information and communication technology, public opinion on environmental issues is increasingly expressed through social media platforms, including YouTube (Memperoleh et al. 2025). Opinions spread on social media can provide important insights into public perceptions and views on deforestation in Papua (Ramadhanas, n.d.). Therefore, monitoring and analyzing public sentiment on social media is very relevant to understand social dynamics and support more appropriate and responsive policy making on this environmental issue (Muhammad et al. 2024).

In today's digital era, public opinion on environmental issues is often reflected through various social media platforms, including YouTube. Public opinions expressed on these platforms

can provide valuable insights into perceptions, concerns, and support for deforestation policies (Arifin et al., 2023.). Therefore, analyzing public sentiment on this issue is important to understand social dynamics as well as to help improve policies more effectively (Dwi Khotimah et al., n.d.).

The Long Short-Term Memory (LSTM) method has been widely used in sentiment analysis due to its ability to capture temporal context in text data (Khadapi et al., n.d.). LSTM is effective in detecting changes in sentiment nuances over time, which is particularly relevant given that public opinion can fluctuate dynamically (Made et al., n.d.). By applying this method to data obtained from the YouTube platform, it is hoped that a more accurate picture of public sentiment towards deforestation in Papua can be obtained.

This research aims to identify and map public sentiment related to forest clearing in Papua through data analysis from YouTube using the LSTM method. The results of the research are expected to contribute to natural resource management decision-making and encourage community involvement in environmental conservation dialog. In addition, this analysis can also serve as a basis for the formulation of public policies that are more responsive to the aspirations and needs of the community.

RESEARCH METHODS

Literature Study

A sentiment analysis was conducted to understand public perceptions of deforestation in Papua. This analysis included the identification of key themes that emerged, sentiment trends over time, and a comparison between positive and negative sentiments. Through this approach, the research provided in-depth insights into the public's concerns, support and opposition to the issue of deforestation, which can serve as a basis for policy-making and environmental conservation efforts (Junaidi et al., n.d.)

Data Collection

The data collection process in this research was carried out by utilizing various platforms and software, including YouTube, Tweepy, Snsrape, as well as JSON and CSV data storage formats. The data was collected through the technique of scraping tweets from the Twitter platform using the Twitter API with the selection of keywords relevant to the research topic. Furthermore, to prepare the data before analysis, word weighting is performed using the Term Frequency Inverse Document Frequency (TF-IDF) method to improve the representation of text features in the sentiment classification process (Putra, 2023).

Sentiment Analysis

Sentiment analysis is a research domain concerned with detecting and interpreting opinions, feelings, assessments, attitudes, and emotions conveyed by individuals in written form (Leonardi et al., n.d.). The scope of sentiment analysis includes various aspects, including emotion analysis, Opinion extraction, sentiment mining, subjectivity analysis, affect analysis, and review mining are techniques designed to identify meaning and emotional nuances within textual data, thereby aiding decision-making processes and enhancing the understanding of public perception (Safitri et al., 2024).

Text Mining

Text mining is the technique of identifying valuable patterns from textual data for the purposes of research and analysis. This field combines elements from information retrieval, data mining, machine learning, statistics, and computational linguistics (Badriyah, 2024) Through the integration of these approaches, text mining is able to process and interpret text data on a large scale to produce insights that are useful in various applications (Kurniasari, 2018).

TF-IDF Weighting

TF-IDF (Term Frequency - Inverse Document Frequency) is a method in Natural Language Processing (NLP) used to measure the importance of a word in a document relative to

a collection of other documents (corpus). This method is often used in information retrieval, text processing, and document classification to determine how important a word is in a text (Septiani et al.2022).

The TF-IDF method consists of two main components: Term Frequency (TF) and Inverse Document Frequency (IDF). Term Frequency (TF) measures how frequently a specific term appears in a document, where higher occurrence leads to a greater TF value (Badriyah, 2024). Nevertheless, commonly used words such as “and,” “is,” or “it” may obtain high TF scores while offering minimal semantic value. To mitigate this, Inverse Document Frequency (IDF) is utilized to reduce the weight of terms that appear across numerous documents and increase the weight of less frequent terms. IDF is computed by taking the logarithm of the ratio between the total number of documents and the number of documents containing the term (Septiani et al., 2024). The TF-IDF value is then obtained, with the TF component calculated systematically as described below formula:

$$TF = \frac{\text{Number of occurrences of a word in a document}}{\text{Total number of words in the document}} \quad (1)$$

The Inverse Document Frequency (IDF) is calculated using the following formula:

$$IDF = \log \left(\frac{\text{Total number of documents}}{\text{Number of documents containing the word}} + 1 \right) \quad (2)$$

Then, the TF-IDF score is obtained by multiplying the values of TF and IDF:

$$TF - IDF = TF \times IDF \quad (3)$$

Explanation:

TF (Term Frequency): a measure of how frequently a word appears in a document.

IDF (Inverse Document Frequency): a measure indicating how rare or unique a word is across all documents.

With this approach, words that appear frequently in one document but infrequently in another will get a higher weight, indicating that they have more important information value. Conversely, words that appear in almost all documents will have a lower weight

Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is composed of memory cells that contain an essential element known as the cell state, which is regulated by three types of gates: the forget gate, the input gate, and the output gate (Larasati, 2020). The forget gate (f_t) determines which parts of the previous cell state are retained or removed, whereas the input gate (i_t) specifies which values in the cell state are modified based on the new input. The output gate (o_t) subsequently manages whether the cell state influences other neurons (Alawiyah, 2021). This architecture enables LSTM to effectively capture long-term dependencies in sequential data while overcoming the vanishing gradient issue. The mathematical formulations for these gates are presented in Equations (1) to (6). The three equations for this gate can be seen in equations (4) to (9)

$$f_t = \sigma(W_{vf}V_t + W_{hf}h_{t-1} + b_f) \quad (4)$$

$$i_t = \sigma(W_{vi}V_t + W_{hi}h_{t-1} + b_i) \quad (5)$$

$$C_t = \tanh(W_{vc}V_t + W_{hc}h_{t-1} + b_c) \quad (6)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot b_o \quad (7)$$

$$o_t = \sigma(W_{vo}V_t + W_{ho}h_{t-1} + b_o) \quad (8)$$

$$h_t \cdot o_{t-1} \cdot \tanh(C_t) \quad (9)$$

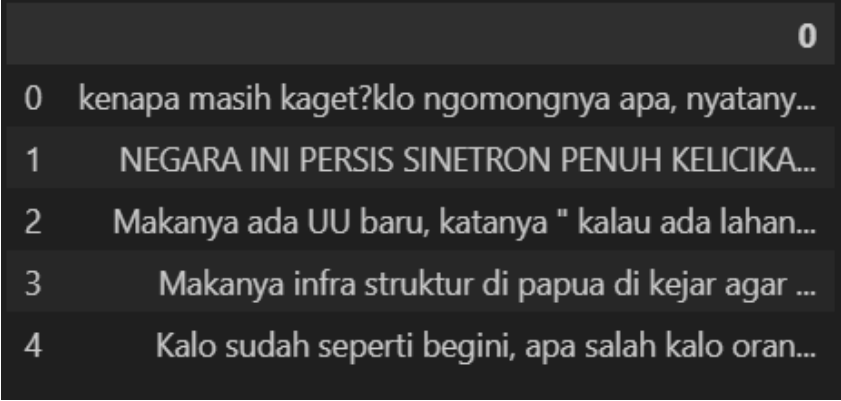
Data Visualization

After preprocessing is complete, data visualization is performed to understand the sentiment distribution, the most frequent words, and the relationship between words in the dataset. Visualizations can be bar graphs, word clouds, or other diagrams that help in exploratory analysis (Agusia et al., n.d.).

RESULTS AND DISCUSSION

Data Initialization

The research began with a data crawling process on the YouTube platform, which served not only as a technical step but also as a strategic one. By carefully selecting appropriate keywords, the researchers successfully gathered thousands of user comments—digital reflections of public concerns, hopes, and even frustration regarding the issue of deforestation. The dataset obtained from this crawling process is presented in Figure 1.



0	kenapa masih kaget?klo ngomongnya apa, nyatany...
1	NEGARA INI PERSIS SINETRON PENUH KELICIKA...
2	Makanya ada UU baru, katanya " kalau ada lahan...
3	Makanya infra struktur di papua di kejar agar ...
4	Kalo sudah seperti begini, apa salah kalo oran...

Figure 1. Dataset Crawling Results

Labelling

The collected data was categorized into two sentiment classes—positive and negative—using the RoBERTa model, which is known for its high accuracy in text classification. The labeling process began by reading data from *cleaned_dataset.csv* and limiting it to 5,000 rows to improve processing efficiency. The model was executed in evaluation mode to generate predictions without additional training. Preprocessed comments were tokenized using RoBERTa's built-in tokenizer, then analyzed to produce logits, which serve as raw scores for each sentiment category (Purwati, n.d.). The highest score was identified using the *argmax* function and assigned as either a positive or negative label. These labels were stored in a new column named *predicted_label_roberta* within the dataset, which was then saved as a separate file. Subsequently, the dataset was cleaned of any empty values, and the predicted labels were converted into numeric form to facilitate further analysis, such as computing the frequency of each sentiment category. The results of this labeling process are illustrated in Figure 2.

```

predicted_label_roberta
1      3767
0       597
Name: count, dtype: int64

```

Figure 2. Labeling Process Result

Preprocessing

Comments obtained from social media have a variety of structures and language styles, so a pre-processing stage is needed to simplify and normalize the text form. This process includes cleaning non-alphabetic characters, changing letters into small forms (case folding), word tokenization, stemming using Sastrawi, and normalizing non-standard words. The following preprocessing stages play a very important role.

a) Case Folding

In this stage, the text cleaning process is carried out, there are several stages, one of which is the Case Folding stage which involves the process of converting the dataset into lowercase letters. The results of case folding data cleaning performed on the dataset can be shown in Figure 3.

```

Data Mentah: Alam indonesia 🇮🇩..sedih saya melihat alam indonesia tercinta 🇮🇩..manusia laknat merusak hutan 🇮🇩..pantesan pemanasan global semakin meningkat 🇮🇩
Cleaning: Alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Case Folding: alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat

```

Figure 3. Case Folding Result

b) Tokenization

After the data cleaning stage, the text is segmented into tokens or smaller textual units. This tokenization process is carried out using the Natural Language Toolkit (NLTK), where the text is split into individual words. The outcome of this tokenization process is presented in Figure 4.

```

Data Mentah: Alam indonesia 🇮🇩..sedih saya melihat alam indonesia tercinta 🇮🇩..manusia laknat merusak hutan 🇮🇩..pantesan pemanasan global semakin meningkat 🇮🇩
Cleaning: Alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Case Folding: alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Normalisasi: alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Tokenisasi: ['alam', 'indonesiasedih', 'alam', 'indonesia', 'tercintamanusia', 'laknat', 'merusak', 'hutanpantesan', 'pemanasan', 'global', 'meningkat']
stemming: ['alam', 'indonesiasedih', 'alam', 'indonesia', 'tercintamanusia', 'laknat', 'rusak', 'hutanpantesan', 'panas', 'global', 'tingkat']

'alam indonesiasedih alam indonesia tercintamanusia laknat rusak hutanpantesan panas global tingkat'

```

Figure 4. Tokenization Result

c) Stemming

During the stemming stage, words containing affixes are transformed into their root or base form. This process utilizes a specialized library designed for the Indonesian language. The stemming results are displayed in Figure 5.

```

Data Mentah: Alam indonesia 🇮🇩..sedih saya melihat alam indonesia tercinta 🇮🇩..manusia laknat merusak hutan 🇮🇩..pantesan pemanasan global semakin meningkat 🇮🇩
Cleaning: Alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Case Folding: alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Normalisasi: alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Tokenisasi: ['alam', 'indonesiasedih', 'alam', 'indonesia', 'tercintamanusia', 'laknat', 'merusak', 'hutanpantesan', 'pemanasan', 'global', 'meningkat']
stemming: ['alam', 'indonesiasedih', 'alam', 'indonesia', 'tercintamanusia', 'laknat', 'rusak', 'hutanpantesan', 'panas', 'global', 'tingkat']

'alam indonesiasedih alam indonesia tercintamanusia laknat rusak hutanpantesan panas global tingkat'

```

Figure 5. Stemming Results

d) Normalization

The results of the sentence that has been modified will be forwarded to the normalization stage of the word that is not standard will be converted into a standard form. The results of normalization can be seen in Figure 6.

```
Data Mentah: Alam indonesia..sedih saya melihat alam indonesia tercinta..manusia laknat merusak hutan..pantesan pemanasan global semakin meningkat
Cleaning: Alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Case Folding: alam indonesiasedih saya melihat alam indonesia tercintamanusia laknat merusak hutanpantesan pemanasan global semakin meningkat
Normalisasi: alam indonesiasedih, 'alam', 'indonesia', 'tercintamanusia', 'laknat', 'merusak', 'hutanpantesan', 'pemanasan', 'global', 'meningkat'
Tokenisasi: ['alam', 'indonesiasedih', 'alam', 'indonesia', 'tercintamanusia', 'laknat', 'merusak', 'hutanpantesan', 'pemanasan', 'global', 'meningkat']
stemming: ['alam', 'indonesiasedih', 'alam', 'indonesia', 'tercintamanusia', 'laknat', 'rusak', 'hutanpantesan', 'panas', 'global', 'tingkat']

'alam indonesiasedih alam indonesia tercintamanusia laknat rusak hutanpantesan panas global tingkat'
```

Figure 6. Normalization Result

TF-IDF

After the preprocessing stage is carried out, the TF-IDF stage is carried out to weight each word based on the frequency of appearance in the dataset. Words that rarely appear but have a strong meaning will get a higher weight, while common words will have a low weight, the following TD-IDF results can be shown in Table 1.

Table 1. TF-IDF Result

Word	TF-IDF
Broken	0.85
Forest	0.80
Polluted	0.75
River	0.70
Sea	0.70
Flood	0.70
Greedy	0.65
Oligarchy	0.60
Palm	0.60
Mine	0.60
Food	0.55
Estate	0.55
Hard	0.50
Life	0.50
People	0.50
Mulyono	0.45
Law	0.45
Prison	0.45
Wise	0.40

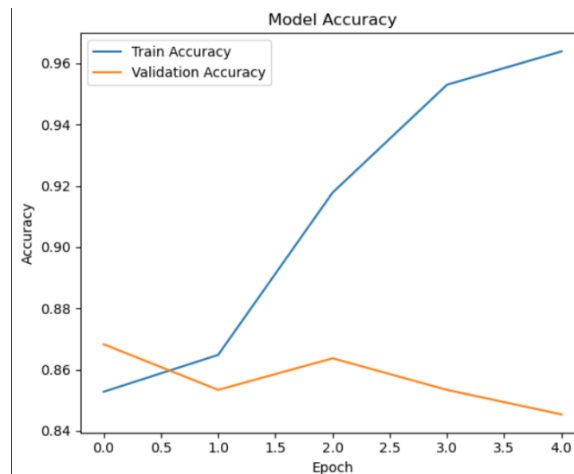
LSTM Evaluation

After the data is collected, the comments go through a pre-processing stage which includes text cleaning, stopword removal, language normalization, and sentiment labeling (positive or negative). Next, the data is divided into two parts, 80% for model training and 20% for testing. The model used is LSTM. The accuracy results obtained by the LSTM model for sentiment analysis with a split data division of 80:20 can be seen in Table 2.

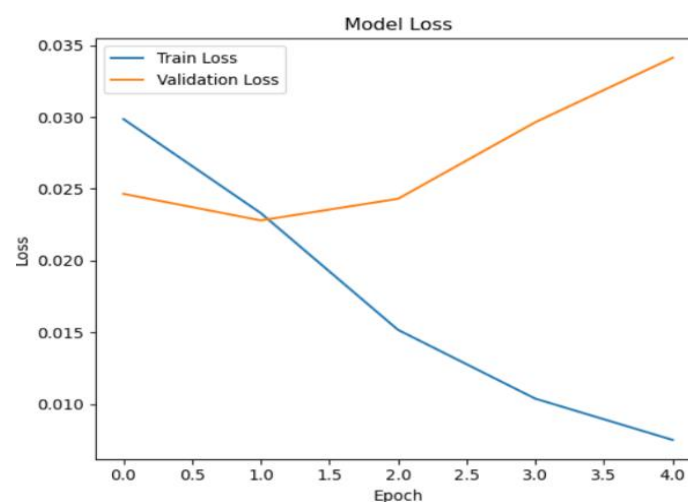
Table 2. LSTM Evaluation Results

Evaluation Metric	Value %
Accuracy	88.43
Precision	90.1
Recall	97.49
F1-Score	93.60

These results show that the LSTM model built is able to perform classification with a fairly good level of accuracy. The model performance graph also shows an increase in accuracy on training data as epochs increase, as can be seen in Figure 7.

**Figure 7.** LSTM Accuracy Model Results

The accuracy graph shows a steady rise in training accuracy as the number of epochs increases, starting at roughly 85% and surpassing 96% by the 4th epoch. This indicates that the model is progressively improving its ability to recognize patterns in the training data. In contrast, the validation accuracy displays a different trend. It began at around 86.7%, dropped in the 1st epoch, experienced a slight increase in the 2nd epoch, and then declined again in subsequent epochs, eventually falling below 85% by the 4th epoch. The corresponding Model Loss graph is presented in Figure 8.

**Figure 8.** LSTM Model Loss Results

The loss graph highlights the contrasting trends between training and validation data. For the training set, the loss value decreases sharply from about 0.03 to below 0.01 by the end of training, indicating that the model has learned the training data very effectively. In contrast, the validation loss shows the opposite pattern, rising from approximately 0.025 to over 0.034 in the final epoch. This trend strongly suggests that the model is experiencing overfitting—performing exceptionally well on training data but struggling to generalize to unseen data. The Confusion Matrix results for the LSTM model are presented in Figure 9.

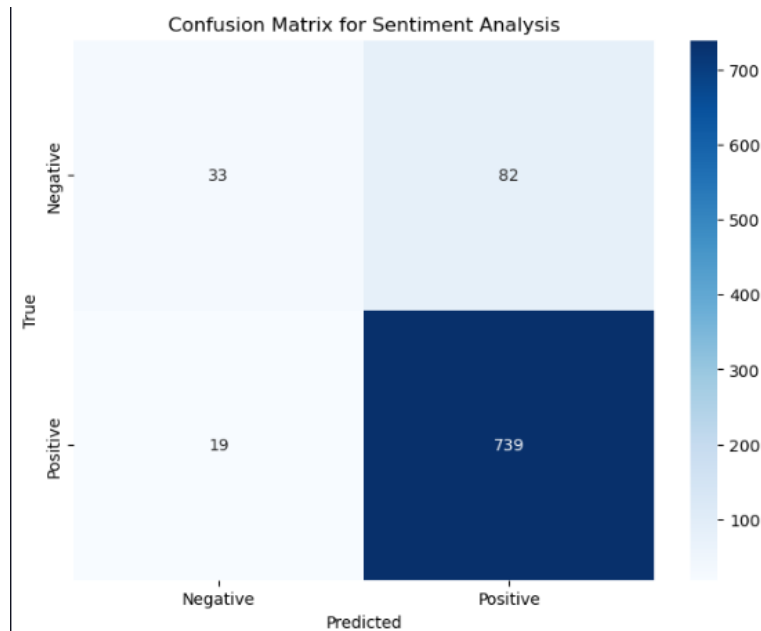


Figure 9. Confusion Matrix Results

Data Visualization

The LSTM model in this study demonstrated strong performance in detecting positive comments, as reflected in its high accuracy for correctly classifying such inputs. Out of all positive comments, 739 were accurately identified, with only 19 misclassified, yielding a recall rate of approximately 97.5%. This indicates that the model is highly responsive to linguistic patterns associated with positive sentiment. However, its capability to detect negative comments remains limited. The majority of negative comments were incorrectly labeled as positive, with only 33 correctly classified. This suggests a bias toward positive classification, likely due to the significantly larger proportion of positive data compared to negative. To enhance overall model performance, measures such as balancing the dataset, fine-tuning parameters, and increasing the amount of negative data are recommended. These improvements are expected to help the model identify both sentiment polarities more evenly and accurately. The visualization of frequently occurring words during the training process is shown in Figure 10.

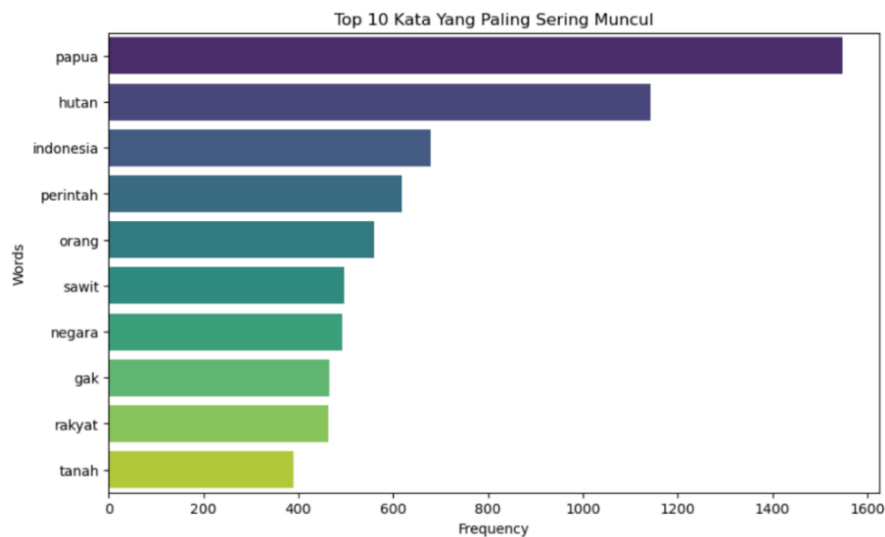


Figure 10. Visualization results of frequently occurring words

CONCLUSION

This research successfully analyzed public sentiment on the issue of deforestation in Papua using Natural Language Processing and Deep Learning approaches. The LSTM model demonstrated strong performance with an accuracy of 88.43%, a precision of 90.01%, a recall of 97.49%, and an F1-score of 93.60%. These results indicate that the LSTM method is effective in classifying positive and negative sentiments from YouTube comments. However, signs of overfitting were observed, as shown by the gap between training and validation performance.

For future research, several improvements are recommended. Expanding the dataset and incorporating data from multiple social media platforms (e.g., Twitter, Instagram, TikTok) would provide a broader representation of public opinion. Addressing class imbalance through oversampling, undersampling, or SMOTE is necessary to improve the detection of negative sentiments. In addition, employing more advanced models such as BERT, RoBERTa, or Transformer-based architectures could further enhance contextual understanding. Aspect-based sentiment analysis is also suggested to capture public concerns related to specific dimensions such as social, economic, and environmental impacts. To mitigate overfitting, future studies could apply regularization techniques, dropout, or cross-validation. Finally, integrating sentiment analysis with spatial visualization (GIS) and adopting an interdisciplinary approach may provide richer insights to support policymaking and community engagement in forest conservation.

ACKNOWLEDGEMENTS

The authors would like to express their deepest gratitude to Universitas Muslim Indonesia, Faculty of Computer Science, for providing academic support and research facilities that contributed to the completion of this study. Special appreciation is extended to colleagues within the Faculty of Computer Science for their valuable input and constructive discussions. The authors also acknowledge the assistance of students who helped in data collection and preprocessing. This research was conducted independently without external funding.

REFERENCES

- Ramadhanas, 2024,. (n.d.). Collective Action in the Digital World:# Alleyesonpapua and the Struggle of Awyu and Moi Indigenous Peoples for Indigenous Forests. Governance.Lkispol.or.Id. Retrieved August 10, 2025, from <https://www.governance.lkispol.or.id/index.php/description/article/view/320>

- Agusia, P., Uli, M., Manurung, A., Calista, V., & Mawardi, V. C. (n.d.). Utilization of Word Cloud on Sentiment Analysis in Exploring Public Perception. *Corisindo.Utb-Univ.Ac.Id*. Retrieved August 10, 2025, from <https://corisindo.utb-univ.ac.id/index.php/penelitian/article/view/11>
- Alawiyah, S. (2021). Modeling Using Recurrent Neural Network Long Short Term Memory (Rnnlstm) Approach on World Gold Prices. <http://repository.unimus.ac.id/id/eprint/4855>
- Arifin, M., ... M. S.-L. K., & 2023, undefined. (n.d.). . Improving the Digital Literacy Capability of the Community in Using Social Media to Realize a Conducive Digital Space in the Framework. *Lib.Lemhannas.Go.Id*. Retrieved August 10, 2025.
- Badriyah, B. (2024). Application of Backpropagation Neural Network and Smote on Sentiment analysis of MYXL user reviews on Google Play Store. <http://etheses.uin-malang.ac.id/72767/>
- Dwi Khotimah, M., Nadia Fatimah, A., Fajar Rizki, M., Puspitasari, W., & Soegiarto, A. (n.d.). Analysis of Public Sentiment towards Pertamina's Pertamina Fuel Modification Issue on Social Media. *J-Innovative.Org*. Retrieved August 10, 2025, from <http://j-innovative.org/index.php/Innovative/article/view/19219>
- Junaidi, A., Susanto, E., Komunikasi, R. L.-, & Media, M. & N. (n.d.). Communication of Diversity: Analysis of News Topics About Conflict and Harmony in Cyber Media. *Researchgate.Net*. Retrieved August 10, 2025.
- Kafiar, A., Teknik, S. S.-J. (Jurnal M., & 2024, undefined. (n.d.). Sentiment Analysis of Netizens on the Issue of Clearing Papua's Indigenous Forests through the Hashtag #Alleyesonpapua Using Algorithms. *Ejournal.Itn.Ac.Id*. Retrieved August 10, 2025, from <https://www.ejournal.itn.ac.id/index.php/jati/article/view/10672>
- Khadapi, M., Maruli Pakpahan, V., Kaputama, S., Veteran No, J., Binjai Kota, K., Binjai, K., & Utara, S. (n.d.). . LSTM and BERT Network-based Sentiment Analysis of Twitter Discussions about the 2024 Election. *Ioinformatic.Org*. Retrieved August 10, 2025, from <https://ioinformatic.org/index.php/JUKI/article/view/681>
- Kurniasari, S. (2018). *Implementation of SVM and Association Analysis For Sentiment Analysis of The Phoenix Hotel Yogyakarta pada Situs Tripadvisor*. <https://dspace.uui.ac.id/handle/123456789/6350>
- Larasati, K. (2020). *Bitcoin Price Prediction Based on Blockchain Information Using Long-Short Term Memory Method*. <https://dspace.uui.ac.id/handle/123456789/23614>
- Leonardi, V., Ibrahim, A., Kurnia, R., Algoritma, M. A.-J., & 2025. Analysis of Language Learning Application Reviews Using the VADER Method. *Journal.Itg.Ac.Id*. <https://doi.org/10.33364/algoritma/v.22-1.2285>
- Made, N., Sedana, K., Nyoman, I., Wijaya, S., Ketut, I., Artana, R., & Korespondensi, P. (n.d.). English Sentiment Analysis with Lstm Method Case Study of Bali Tourism Online News. *Jtiik.Ub.Ac.Id*. <https://doi.org/10.25126/jtiik.2024118792>
- Obtainings, G., Sarjana, G., Dalam, (s-1, Tarbiyah, I., & Tarbiyah, F. (2025). *Analysis of Cyberpragmatic Views on Youtube on Deddy Corbuzier's Social Media Account Podcast*. <http://etheses.iaincurup.ac.id/id/eprint/7908>

- Muhammad, R., Review, B. T.-U. L.,. (2024) The Influence of Social Media on Public Perceptions of the Justice System: A Sentiment Analysis on Twitter. Review Unes.Com, 7(1). <https://doi.org/10.31933/unesrev.v7i1>
- Pipin, S. J., Kurniawan, H., Jurnal, |, & Mikroskil, S. (2022). Sentiment Analysis of MBKM Policy Based on Public Opinion on Twitter Using LSTM. *Pdfs.Semanticscholar.OrgSJ Pipin, H KurniawanJournal SIFO Mikroskil, 2022-pdfs.Semanticscholar.Org, 23, 1–5.* <https://pdfs.semanticscholar.org/8166/8b5accf7b144802cb0a5dd4d11094552f423.pdf>
- Purwati, K. (n.d.). Sentiment Analysis of Covid-19 Vaccine News with Robustly Optimized Bert Pre-Training Approach (Roberta). Repository.Uinjkt.Ac.Id. Retrieved August 10, 2025, from <https://repository.uinjkt.ac.id/dspace/handle/123456789/73319>
- Putra, A. (2023). Sentiment Analysis of Public Responses to Youtube Videos about the 2023 Recession using Naïve Bayes Dan K-Nearest Neighbor (KNN) Method. <https://repository.uinjkt.ac.id/dspace/handle/123456789/72716>
- Rinjani, A., Shadewi, E., K. N.- J of E., (2024). Economic and Environmental Impacts of Forest Exploitation. *Journal.Academiacenter.Org.* <https://jurnal.academiacenter.org/index.php/IJEN/article/view/529>
- Safitri, L., Nurhasan, R., E-Bis, F. R.-J. (2024). Sentiment Analysis of Starbucks Brand Image on Twitter Social Media. Jurnal.Politeknik-Kebumen.Ac.Id, 8(1), 193–202. <https://doi.org/10.37339/e-bis.v8i1.1610>
- Septiani, D., Indonesia, I. I.-S. dan T. I., & 2022. (n.d.-a). Term frequency inverse document frequency (tf-idf) analysis in information retrieval on text documents. 103.8.12.212. Retrieved August 10, 2025, from <http://103.8.12.212:33180/unj/index.php/SINTESIA/article/view/39364>
- Septiani, D., Indonesia, I. I.-S. dan T. I., & 2022. Term frequency inverse document frequency (tf-idf) analysis in information retrieval on text documents. 103.8.12.212. Retrieved August 10, 2025, from <http://103.8.12.212:33180/unj/index.php/SINTESIA/article/view/39364>