Smart Techno (Smart Technology, Informatic, and Technopreneurship)
Volume 8, Number 2, 2026 pp. 303-313
E-ISSN : 2541-0679
Open Access: https://lppm.primakara.ac.id/jurnal/index.php/smart-techno/

Comparative Analysis of Naïve Bayes and SMOTE-Based Long Short-Term Memory (LSTM) for Electric Vehicle Sentiment Analysis on YouTube

Rayhan Gimnastiar[1]*, Fania Indah Lestari[2], Nadhif Daniswara Prasetyo[3], Eko Prasetyo[4], Fardanto Setyatama[5]
[1,2,3,4,5] Informatics Engineering, Universitas Bhayangkara Surabaya, Surabaya, Indonesia

ARTICLE INFO
Article history:
Initial submission
15-06-2026
Received in revised form
08-08-2026
Accepted 21-08-2026
Available online 30-09-2026

Keywords:
Sentiment Analysis,
Electric Vehicles,
YouTube, Naïve Bayes,
LSTM, SMOTE

DOI:
https://doi.org/10.59356/smart-techno.v8i2.230

ABSTRACT

The transition to electric vehicles in Indonesia has generated diverse public opinions on social media. Most previous sentiment analysis studies have tended to employ a single classification method without in-depth comparison and have overlooked the issue of extreme data imbalance, which can introduce bias into classification models. Furthermore, a research gap remains regarding the effectiveness of complex deep learning models compared with simpler statistical models when applied to medium-dimensional Indonesian-language opinion texts. This study aims to address this gap by conducting a comparative evaluation of model performance. The dataset consists of 1,512 unstructured public opinion comments collected from YouTube, distributed across three sentiment classes: 868 negative, 503 positive, and 141 neutral comments. To address the class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. The study then compares the statistical Naïve Bayes method using TF-IDF weighting with the Deep Learning Long Short-Term Memory (LSTM) method using Word Embedding. Evaluation using a Confusion Matrix demonstrates that Naïve Bayes outperformed LSTM and exhibited greater stability in sentiment classification for this dataset, achieving an accuracy of 63.04%, compared with 57.43% for the LSTM model.

1. INTRODUCTION

The transition to electric vehicles in Indonesia as an environmentally friendly transportation solution has triggered polarized public opinions, with public enthusiasm often accompanied by concerns regarding selling prices, infrastructure limitations, subsidy policies, and battery sustainability (Huzna et al., 2024). This debate is extensively reflected in the form of unstructured textual data on digital platforms such as YouTube (Sri Widagdo Adika Nuresa Qodri et al., 2023). Therefore, a Text Mining approach through Sentiment Analysis is essential for computationally extracting and categorizing public perceptions (Suryani et al., 2023). Objective sentiment mapping is crucial as a basis for strategic evaluation by the government and automotive industry in formulating effective policies and marketing strategies to accelerate electric vehicle adoption in Indonesia (Bayu & Wurijanto, 2024).

Research on electric vehicle sentiment analysis has been extensively conducted using various Machine Learning algorithms. Most previous studies have relied on traditional statistical classification models, such as the Naïve Bayes Classifier (NBC) and Support Vector Machine (SVM), which have demonstrated speed and efficiency in handling medium-sized datasets (Bayu & Wurijanto, 2024). Meanwhile, advances in artificial intelligence have introduced Deep Learning approaches, such as Long Short-Term Memory (LSTM), which are considered more capable of capturing sequential information and semantic context in longer sentences (Sri Widagdo Adika Nuresa Qodri et al., 2023). This situation raises a research gap concerning whether complex Deep Learning models are genuinely more accurate than simpler

* *Corresponding Author: Eko Prasetyo: eko@ubhara.ac.id*

statistical models when applied to Indonesian public opinion texts characterized by short, informal, and medium-dimensional textual data.

To address this research gap, this study proposes a dual contribution. First, it directly compares the performance of Naïve Bayes and LSTM on YouTube comments containing relatively complex and potentially sarcastic expressions. Second, it specifically applies the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance (Persada Pulungan et al., 2024), a critical methodological issue that has often been overlooked in previous research. The application of SMOTE is intended to ensure that the accuracy comparison between the two models is more objective and less susceptible to class-distribution bias.

Based on the identified background and research gap, the research questions focus on the distribution of public sentiment (positive, negative, and neutral) regarding electric vehicles in Indonesia on the YouTube platform and the comparative classification performance of the statistical Naïve Bayes method and the Deep Learning Long Short-Term Memory (LSTM) approach after applying SMOTE to address class imbalance (Rezki et al., 2024). The study also seeks to identify textual characteristics that may influence the superior performance of either algorithm on this dataset.

To address these research questions, this study proposes a comparative sentiment analysis approach. The primary objective is to implement and objectively compare the performance of the two classification models. Raw textual data are systematically processed through a series of preprocessing steps. Subsequently, Term Frequency-Inverse Document Frequency (TF-IDF) feature extraction is combined with Naïve Bayes, while a Word Embedding and sequence-based representation is fed into the LSTM model. The classification performance of both approaches is then evaluated using a Confusion Matrix to identify the model that is most effective and appropriate for the characteristics of Indonesian public opinion texts concerning electric vehicles.

2. LITERATURE REVIEW

Although successfully implementing a Deep Learning model, namely Long Short-Term Memory (LSTM), specifically designed to process sequential textual data, the study has several limitations. It employs only a single algorithm without comparative models and trains the data without addressing class imbalance. Consequently, its classification accuracy varies across different classes (Sri Widagdo Adika Nuresa Qodri et al., 2023). The strength of the comparative evaluation between Naïve Bayes and Support Vector Machine (SVM) using K-Fold Cross-Validation in this study is unfortunately not complemented by an appropriate strategy for addressing the severe class imbalance in the dataset, which consists of 71% Neutral, 20% Positive, and 9% Negative samples, prior to the training process (Suryani et al., 2023).

This study can be considered comprehensive as it directly compares three algorithms SVM, Random Forest, and Naïve Bayes. However, it presents a methodological gap by overlooking the application of data-balancing techniques despite the highly imbalanced nature of the dataset (Tarigan & Yusupa, 2024). The application of Naïve Bayes in this study proved to be highly effective, achieving an accuracy of up to 100%. Nevertheless, its approach remains limited because it relies solely on a single classification method without conducting comparative evaluations against other models (Huzna et al., 2024). Although the research is highly structured through the adoption of the CRISP-DM methodology, it has several limitations, including a very small dataset consisting of only 214 textual data points, the absence of a stemming stage in the preprocessing process, and the use of only a single classification model (Bayu & Wurijanto, 2024).

This study is innovative in that it performs Aspect-Based Sentiment Analysis (ABSA) by integrating Latent Dirichlet Allocation (LDA) topic modeling with Naïve Bayes. However, its resulting classification performance remains relatively low, with an accuracy of only 53%, and

the model has not been further optimized using boosting techniques (Fatah et al., 2024). The successful integration of SVM and Particle Swarm Optimization (PSO) to significantly improve classification accuracy represents a major strength of the study. However, the evaluation was still conducted using an imbalanced data distribution without employing any data-balancing algorithm (Santoso et al., 2022).

This study demonstrates considerable strength by comparing three conventional algorithms K-Nearest Neighbors (KNN), Naïve Bayes, and SVM and addressing the class imbalance issue through the application of the Synthetic Minority Over-sampling Technique (SMOTE). Nevertheless, a remaining limitation is the lack of comparison between the performance of these classical algorithms and more recent Deep Learning-based approaches (Polgan et al., 2024). The study is notable for employing a combination of Term Frequency Inverse Document Frequency (TF-IDF) and N-Gram for word weighting, along with Chi-Square feature selection, to compare SVM and Naïve Bayes in terms of classification accuracy and training efficiency. However, the evaluation remains limited to traditional statistical models and does not consider Deep Learning-based neural network architectures (Ernawati et al., 2023).

Sentiment Analysis and Electric Vehicles

The transition from fossil fuel-powered vehicles to electric vehicles represents an important technological innovation, offering significant environmental benefits while reducing noise pollution (Polgan et al., 2024). The continuous reliance on conventional energy sources has led to a substantial depletion of fossil fuel resources (Suryani et al., 2023). This technological transition policy has consequently sparked both supportive and opposing opinions, generating a large volume of public comments across various social media platforms, particularly YouTube (Santoso et al., 2022).

Public opinions and responses play a crucial role in determining the successful adoption of electric vehicles in Indonesia (Ernawati et al., 2023). Therefore, sentiment analysis can be defined as a computational technique for systematically identifying, extracting, and classifying users' opinions (Tarigan & Yusupa, 2024). The application of sentiment analysis is essential for providing stakeholders with deeper insights into public perceptions, thereby supporting efforts to strengthen public acceptance and adoption of emerging technologies (Huzna et al., 2024).

Text Preprocessing and TF-IDF Weighting

Opinion datasets obtained from social media are generally unstructured and characterized by free-form text (Fatah et al., 2024). Such raw data must undergo a series of text preprocessing steps to eliminate textual noise (Huzna et al., 2024). This initial cleaning process is crucial because opinion data collected from the internet often contain slang, informal expressions, and non-standard words (Tarigan & Yusupa, 2024). The series of text transformations includes normalization processes, such as removing numerical characters and punctuation marks and converting all text to lowercase. Once the text has been cleaned, the data must be transformed into numerical vectors using a word-weighting method (Bayu & Wurijanto, 2024). Term Frequency–Inverse Document Frequency (TF-IDF) is a weighting technique that measures how frequently a word occurs within a document while simultaneously determining the importance of that word across the entire dataset (Bayu & Wurijanto, 2024).

Naïve Bayes Classifier Algorithm

The Naïve Bayes Classifier is a text classification method based on Bayes' theorem, with the strong assumption that the features are conditionally independent of one another (Bayu & Wurijanto, 2024). This algorithm is considered highly effective and has been extensively evaluated by researchers for sentiment analysis of electric vehicle-related topics (Bayu & Wurijanto, 2024). The primary advantages of this probabilistic model are its ease of implementation and high computational efficiency, enabling rapid data processing (Ernawati et al., 2023).

In addition, the Naïve Bayes model is considered robust and has been shown to achieve superior precision compared with other standard classification algorithms (Suryani et al., 2023). Its computational efficiency makes it highly suitable for identifying variations in the polarity of public opinions on a large computational scale (Huzna et al., 2024).

Long Short-Term Memory (LSTM) Algorithm

The Long Short-Term Memory (LSTM) architecture is a specialized variant of an advanced artificial neural network known as a Recurrent Neural Network (RNN) (Prasetyo, 2024). This Deep Learning-based algorithm is mathematically designed to address the problem of vanishing gradients encountered in conventional models (Prasetyo, 2024). LSTM offers computational advantages in retaining and processing information with long-term sequential dependencies within sentence structures (Prasetyo, 2024). In the context of natural language processing, neural network-based algorithms are frequently employed as primary comparative models to evaluate the reliability of conventional Machine Learning architectures (Suryani et al., 2023). Through this Deep Learning approach, the system is expected to automatically identify more complex patterns in natural language without requiring manual feature extraction (Prasetyo, 2024).

Data Imbalance and Model Performance Evaluation

A major challenge frequently encountered when collecting real-world opinion data is the uneven distribution of data across sentiment classes, commonly referred to as imbalanced data (Huzna et al., 2024). An imbalanced distribution of sentiment samples across positive and negative classes can introduce bias and reduce the reliability of model predictions (Polgan et al., 2024). To assess the effectiveness and validity of the classification algorithms being evaluated, the study requires an evaluation instrument in the form of a Confusion Matrix (Santoso et al., 2022). The Confusion Matrix provides a visualization of the prediction results, which are subsequently quantified into various evaluation metrics (Tarigan & Yusupa, 2024). Model reliability is comprehensively evaluated based on accuracy, precision, recall, and F1-score (Tarigan & Yusupa, 2024).

3. METHOD

Research Stages

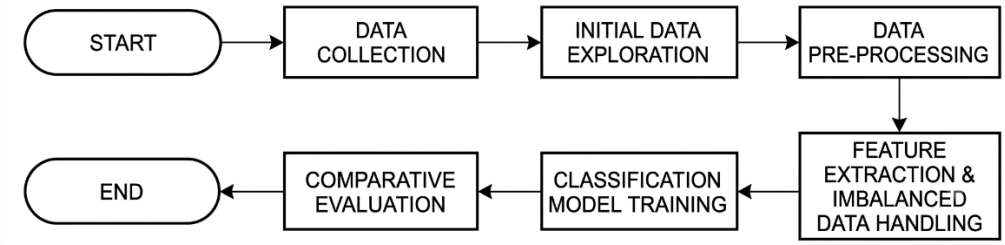


Figure 1. Research Stages Flowchart

This study was conducted using a systematic workflow based on Knowledge Discovery in Databases (KDD), adopting the standard stages of the Cross-Industry Standard Process for Data Mining (CRISP-DM) (Polgan et al., 2024). The research workflow consists of six main stages: dataset collection, data preprocessing, feature extraction/word representation, model training using Naïve Bayes and LSTM, and comparative evaluation using a Confusion Matrix. The computational processing pipeline for sentiment analysis was developed using the Python programming language to transform unstructured data into valid and meaningful information (Ernawati et al., 2023). This experimental approach focuses on comparing multiple classification algorithms to identify the most optimal model for mapping public opinions expressed on social media (Tarigan & Yusupa, 2024).

Data Imbalance Handling Using SMOTE

The initial stage of the study involved collecting a raw dataset consisting of public opinions or comments regarding the presence of electric vehicles in Indonesia (Bayu & Wurijanto, 2024). The dataset is publicly available on Kaggle at https://www.kaggle.com/datasets/billycemerson/analisis-sentimen-terkait-intensif-mobil-listrik. The dataset was subsequently explored visually using sentiment distribution graphs and WordCloud to identify the frequency of keyword occurrences and the proportion of opinion classes (Huzna et al., 2024). This data exploration is an important step for detecting potential class imbalance before the data are used in the classification algorithms (Polgan et al., 2024).

The exploration results showed that the dataset consists of three sentiment classes: negative, positive, and neutral, with 868, 503, and 141 comments, respectively. These class labels were already provided in the dataset according to the version submitted by the contributor. In this experiment, the dataset was divided into 80% training data and 20% testing data. Accordingly, the training set contained 694 negative, 402 positive, and 113 neutral samples, while the testing set contained 174 negative, 101 positive, and 28 neutral samples.

Classification Modeling

Naïve Bayes Classifier (NBC): This study conducts a comparative analysis by implementing two classification methods based on different architectural paradigms. Naïve Bayes is a probabilistic, statistical-based classification algorithm grounded in Bayes' theorem, with a strong assumption of independence among word features (Suryani et al., 2023). This algorithm is employed because it is known for its high efficiency, fast computational performance, and effectiveness in medium-scale text classification tasks (Ernawati et al., 2023). The basic Bayes theorem formula for text classification is as follows:

$$P(C|X) = \frac{\left(P(X|C)\cdot P(C)\right)}{P(X)} \qquad (1)$$

P(C|X): The probability of sentiment class C (positive/negative/neutral) given the occurrence of word X.
P(X|C): The probability of word X occurring in sentiment class C.
P(C): The prior probability of sentiment class C.
P(X): The overall probability of word X occurring in the dataset.

Long Short-Term Memory (LSTM): LSTM is a Deep Learning-based architecture developed from the Recurrent Neural Network (RNN) to process sequential or textual data with long-term dependencies (Sri Widagdo Adika Nuresa Qodri et al., 2023). Unlike Naïve Bayes, which treats words independently, LSTM can capture the semantic context of sequential sentence structures through three main gates Forget Gate, Input Gate, and Output Gate that regulate memory storage. During the training process, the LSTM model is trained iteratively over multiple epochs, while the progression of the loss and accuracy values is monitored to ensure optimal model convergence.

Performance Evaluation Metrics

The modeling process in this study compares the performance of different algorithmic architectures to objectively identify patterns in opinion classification (Tarigan & Yusupa, 2024). The first model implemented is the Naïve Bayes Classifier, a computational probabilistic approach known for its high efficiency in processing and extracting information from textual data (Ernawati et al., 2023). As a comparative model, the second model employs the Long Short-Term Memory (LSTM) artificial neural network, which is specifically designed to process complex and sequential patterns in textual data (Prasetyo, 2024). The comparative experiments using these algorithms were conducted to examine the differences in the capabilities of standard machine learning and deep learning models in handling variations in opinion sentences (Suryani et al., 2023).

Model Performance Evaluation

The final performance of the Naïve Bayes and LSTM models in predicting sentiment classes (positive, negative, and neutral) on the test data was objectively evaluated using a Confusion Matrix (Tarigan & Yusupa, 2024). Based on the results of the Confusion Matrix, the comparative effectiveness of the two models was assessed using four standard evaluation metrics (Polgan et al., 2024), as follows:

Accuracy refers to the percentage of correctly classified sentiment predictions (positive, negative, and neutral) relative to the total number of test data instances.

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \quad (2)$$

Precision refers to the degree of accuracy of the model in predicting a particular sentiment class compared with the actual data.

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (3)$$

Recall refers to the model's ability to successfully identify and retrieve all relevant instances belonging to the actual sentiment class.

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (4)$$

F1-Score refers to the harmonic mean of Precision and Recall, providing a balanced measure that represents the overall stability of the model's performance.

$$\text{F1} - \text{Score} = 2 \cdot \frac{(\text{Precision} \cdot \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (5)$$

TP (True Positive): Positive instances that are correctly predicted as positive.
TN (True Negative): Negative or neutral instances that are correctly classified by the model.
FP (False Positive): Instances that are not actually positive but are incorrectly predicted as positive by the model.
FN (False Negative): Instances that are actually positive but are incorrectly predicted as belonging to another class by the model.

The calculation of these comparative evaluation metrics ultimately provides a scientific basis for determining the highest level of reliability among the algorithms evaluated (Suryani et al., 2023).

4. RESULT AND DISCUSSION

Dataset Collection and Exploration

This study uses a public opinion dataset on electric vehicles collected from the YouTube platform. The distribution of the dataset is presented in Figure 2. An initial data exploration was conducted to examine the general trends in public sentiment. Based on the visualization, the data distribution shows varying proportions across positive, negative, and neutral sentiment classes.

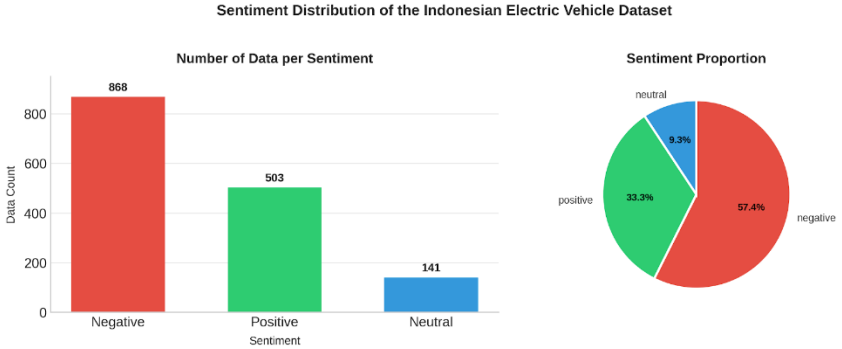


Figure 2. Sentiment distribution of the electric vehicle dataset in Indonesia

This data distribution is consistent with the dataset information presented in Subsection 3.2. The numbers of instances in the positive, negative, and neutral classes are 868, 503, and 141 comments, respectively. Each instance has been assigned a class label according to the sentiment expressed in its content. The number of instances across the three classes is imbalanced, reflecting the original class distribution provided by the dataset contributor. This distribution is consistent with previous findings indicating that public opinion toward electric vehicles tends to be negative due to factors such as vehicle prices and charging infrastructure. As shown in the WordCloud visualization in Figure 3, the most dominant keywords in public discussions include "harga", "baterai", "subsidi", and "kualitas". The prominence of these keywords indicates that Indonesian public perceptions of electric vehicles are not limited to environmental considerations but also encompass economic and technical aspects of electric vehicle use.

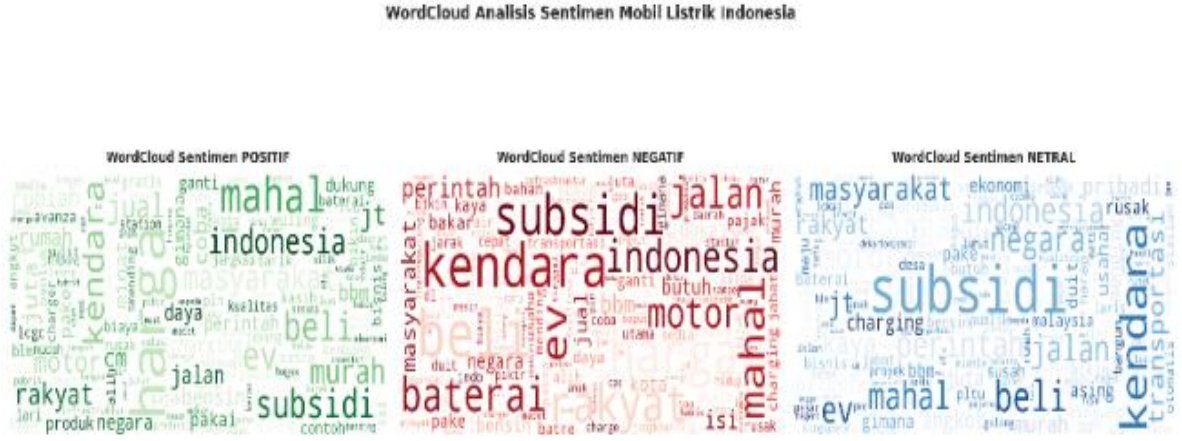


Figure 3. WordCloud visualization of sentiment analysis of electric vehicles in Indonesia

Text Preprocessing Results

The raw dataset underwent a series of preprocessing steps to improve data quality. The processes of case folding, tokenization, and stopword removal successfully transformed unstructured sentences into cleaner textual representations. For example, the sentence “subsidi kualitas diturunin harga dinaikin usaha cari cuan”, which initially contained textual noise, was normalized to enable more accurate keyword extraction. This preprocessing stage is consistent with the methodology applied to optimize the performance of the classification models.

To facilitate a clearer understanding of the preprocessing stages, the following example illustrates the transformation of data from a raw comment into a numerical (vector) representation ready to be processed by the Naïve Bayes and LSTM algorithms.

1. Raw Data
   The original YouTube comment before any preprocessing is applied: "Mobil listrik itu MAHAL bangettt!!! 😡 Belum lagi ngecasnya susah, mending pake bensin aja bro" #MobilListrik
2. Cleansing and Case Folding
   The system removes irrelevant elements, such as links, emojis, hashtags, and punctuation, and converts all letters to lowercase. The resulting text is: “mobil listrik itu mahal bangettt belum lagi ngecasnya susah mending pake bensin aja bro”
3. Normalization (Word Normalization)
   Informal words, abbreviations, and typographical errors (typos) are converted into standardized forms to ensure consistent interpretation by the system. The resulting text is:“mobil listrik itu mahal banget belum lagi mengecasnya susah lebih baik pakai bensin saja saudara”
4. Tokenization

The complete sentence is divided into individual words or tokens. The resulting tokens are:['mobil', 'listrik', 'itu', 'mahal', 'banget', 'belum', 'lagi', 'mengecasnya', 'susah', 'lebih', 'baik', 'pakai', 'bensin', 'saja', 'saudara']

5. Stopword Removal and Stemming
   Words with limited relevance to sentiment classification, such as itu, belum, lagi, lebih, baik, and saja, are removed, while affixed words are reduced to their root forms. The final cleaned text is: ['mobil', 'listrik', 'mahal', 'banget', 'cas', 'susah', 'pakai', 'bensin', 'saudara']
6. Feature Extraction (Numerical Representation)
   At this stage, the cleaned text is converted into numerical representations because machine learning algorithms cannot directly process raw text. Since this study compares two models with different architectures, the data are represented differently for each model:
   - Naïve Bayes Model (TF-IDF Weighting): The system calculates the importance of each word within the entire document collection. Words such as “mahal” and “susah” may receive relatively high weights because they strongly represent the negative sentiment class. The model input is represented as a TF-IDF weight vector, for example: [0.0, 0.0, 0.45, 0.62, 0.0, 0.58, ...].
   - LSTM Model (Word Embedding/Sequence): Unlike TF-IDF, which does not preserve word order, LSTM models explicitly account for the sequential structure of sentences. Each word is converted into a numerical index according to its position in the vocabulary (e.g., mobil = 12, listrik = 8, mahal = 45). The model input is represented as a sequence vector, for example: [12, 8, 45, 102, 33, 88, 54, 9, 71], with sequences standardized to a fixed length through padding.

Model Training Process

The training process of the Long Short-Term Memory (LSTM) model was conducted by monitoring the loss and accuracy curves across each epoch, as shown in Figure 4.

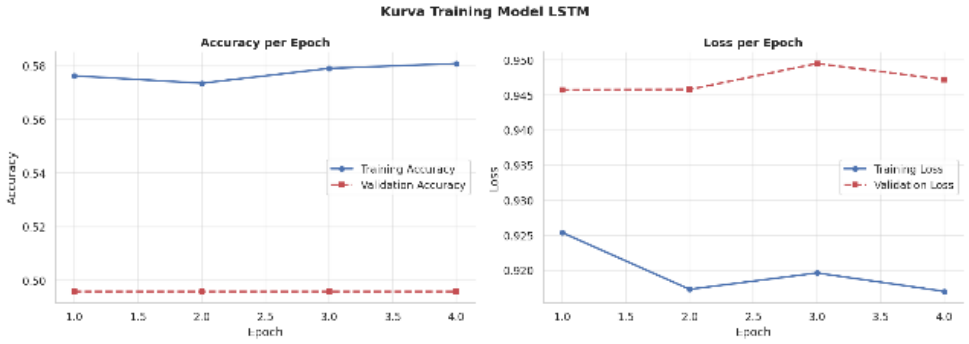
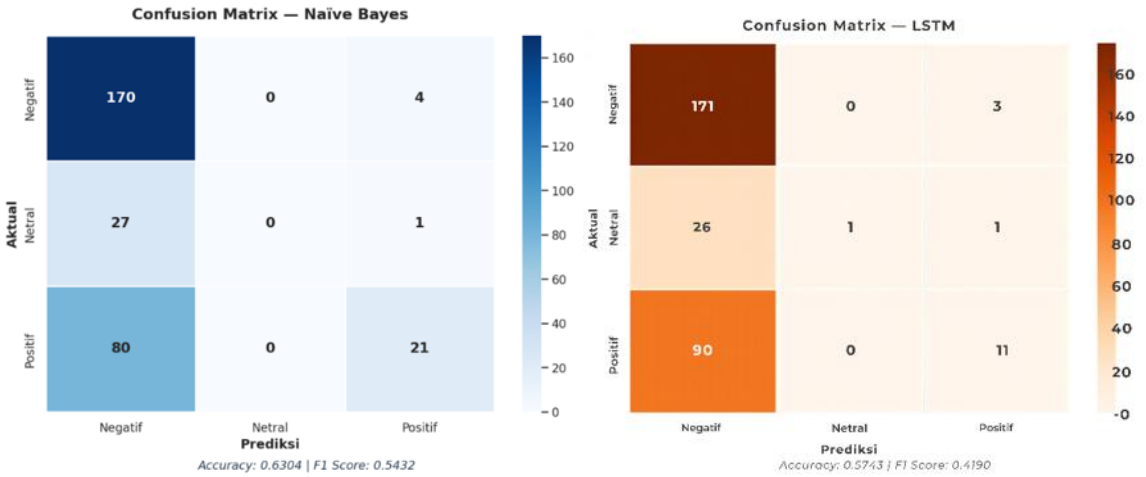


Figure 4. LSTM Model Training Curves

The curves indicate that the model achieved convergence after several iterations, suggesting that the LSTM architecture successfully learned sequential patterns in public opinion without exhibiting substantial overfitting. In contrast, the Naïve Bayes model employs a more computationally efficient probabilistic approach for processing word distributions in the TF-IDF dataset.

Model Performance Evaluation Using the Confusion Matrix

Model performance was evaluated by predicting all test instances using the trained models. The prediction results were then mapped onto a Confusion Matrix, as presented in Figure 5. The same test set consisting of 303 instances was used for both methods, comprising 174 negative, 101 positive, and 28 neutral instances.



a. Naïve Bayes b. LSTM

Figure 5. Confusion Matrix results for the Naïve Bayes and LSTM models

As shown in Figure 5, both models demonstrated relatively good performance in distinguishing the negative sentiment class but performed less effectively on the positive and neutral classes. The confusion matrices provide a clear overview of the misclassification patterns observed in each model.

Comparative Evaluation Metrics of the Models

The performance of the two models was compared using the Accuracy, Precision, Recall, and F1-Score metrics. The visualization of this comparison is presented in Figure 6.

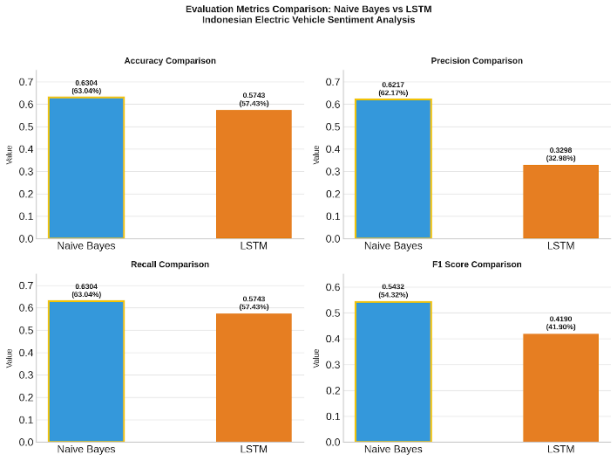


Figure 6. Comparison of Evaluation Metrics for the Naïve Bayes and LSTM Models

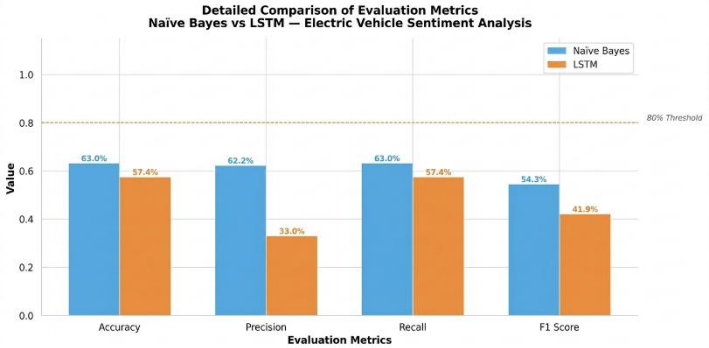


Figure 7. Comparison of Evaluation Metrics for the Naïve Bayes and LSTM Models

The data presented in Table 1 summarize the performance of the two models.

Table 1. Summary of Comparative Evaluation Results

Metode	Accuracy	Precision	Recall	F1 Score
Naïve Bayes (TF-IDF)	0.6304	0.6217	0.6304	0.5432
LSTM (Deep Learning)	0.5743	0.3298	0.5743	0.4190

The results show that Naïve Bayes (TF-IDF) outperformed LSTM in this study, achieving an accuracy of 63.04%, compared with 57.43% for LSTM. These findings indicate that, for the dataset used in this study, the Naïve Bayes statistical approach provided more stable performance than the LSTM architecture. The lower performance of LSTM may be attributed to the relatively limited dataset size and the need for further hyperparameter tuning to achieve optimal performance.

Sentiment Analysis Discussion

The results indicate that Naïve Bayes was the most effective model for classifying sentiment in public opinions regarding electric vehicles within the dataset. Overall, the sentiment analysis suggests that public enthusiasm for environmentally friendly technologies is accompanied by critical concerns regarding cost efficiency, particularly in relation to battery components. Therefore, the findings may provide valuable insights for government agencies and industry stakeholders in improving public communication strategies and developing supporting infrastructure to enhance public confidence in electric vehicles in Indonesia.

5. CONCLUSION

Based on the evaluation results, the Naïve Bayes algorithm with TF-IDF weighting outperformed the LSTM model in classifying public sentiment toward electric vehicles, achieving an accuracy of 63.04%. Although the public demonstrates enthusiasm for this environmentally friendly innovation, negative sentiments remain due to concerns regarding vehicle prices, limited charging infrastructure, and battery sustainability issues. Further text mining analysis also identified battery component prices and subsidy policies as the dominant factors most frequently discussed by the public on social media.

For future research, the dataset should be expanded by incorporating data from multiple social media platforms and balanced using the SMOTE (Synthetic Minority Over-sampling Technique) method to address the imbalanced data problem and potentially improve model precision and recall. In addition, more extensive optimization of deep learning models is recommended, including the exploration of advanced architectures such as BERT. The application of Aspect-Based Sentiment Analysis (ABSA) is also recommended to enable more specific sentiment mapping across particular aspects, such as battery quality and pricing policies.

6. REFERENCES

Bayu, L. K., & Wurijanto, T. (2024). Analisis sentimen mobil listrik menggunakan metode Naïve Bayes Classifier. INFOTECH : Jurnal Informatika & Teknologi, 5(2), 328–335. https://doi.org/10.37373/infotech.v5i2.1465

Ernawati, N. W., Satya Kumara, I. N., & Setiawan, W. (2023). Perbandingan Metode Klasifikasi Support Vector Machine Dan Naïve Bayes Pada Analisis Sentimen Kendaraan Listrik. Jurnal SPEKTRUM, 10(3), 106. https://doi.org/10.24843/spektrum.2023.v10.i03.p12

Fatah, D. A., Kamil, F. I., & Soesilo, B. (2024). Analisis Sentimen Dan Pemodelan Topik Terhadap Pembangunan Kereta Cepat Jakarta-bandung Menggunakan Naive Bayes Dan Latent Dirichlet Allocation (LDA). Jurnal Simantec, 12(2), 53–62. https://doi.org/10.21107/SIMANTEC.V12I2.25440

Huzna, A. N., Nurhayati, I., Saputri, A. E., & Qomarul Huda, M. (2024). Analisis Sentimen Terhadap Mobil Listrik Di Indonesia Pada Twitter: Penerapan Naïve Bayes Classifier Untuk Memahami Opini Publik. JustIT:Jurnal Sistem Informasi, Teknologi Informasi Dan Komputer, 14(2), 80–149.

Persada Pulungan, M., Purnomo, A., & Kurniasih, A. (2024). Penerapan SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Kepribadian MBTI Menggunakan Naive Bayes Classifier. Jurnal Teknologi Informasi Dan Ilmu Komputer, 11(5), 1033–1042. https://doi.org/10.25126/JTIIK.2024117989

Polgan, J. M., Kumalasari, J. T., Puspitorini, I., Bina, U., Informatika, S., & Listrik, M. (2024). Perbandingan Metode Klasifikasi dan SMOTE. 13(September), 2257–2268.

Prasetyo, Y. (2024). Analisis Sentimen Perspektif Publik Terhadap Mobil Listrik di Indonesia, Jurnal Informatika dan Sistem Informasi, vol. 10, no. 2, pp. 45-56, 2024.

Rezki, M. K., Mazdadi, M. I., Indriani, F., Muliadi, M., Saragih, T. H., & Athavale, V. A. (2024). Application Of SMOTE To Address Class Imbalance In Diabetes Disease Classification Utilizing C5.0, Random Forest, And SVM. Journal of Electronics, Electromedical Engineering, and Medical Informatics, 6(4), 343–354. https://doi.org/10.35882/JEEEMI.V6I4.434

Santoso, A., Nugroho, A., & Sunge, A. S. (2022). Analisis Sentimen Tentang Mobil Listrik Dengan Metode Support Vector Machine Dan Feature Selection Particle Swarm Optimization. Journal of Practical Computer Science, 2(1), 24–31. https://doi.org/10.37366/jpcs.v2i1.1084

Sri Widagdo Adika Nuresa Qodri, Krisna Edi Nugroho, Fachruddin S, Akbar Rizky, & Nisrina P. (2023). Analisis Sentimen Mobil Listrik di Indonesia Menggunakan Long-Short Term Memory (LSTM). Jurnal Fasilkom, 13(3), 416–423.

Suryani, S., Fayyad, M. F., Savra, D. T., Kurniawan, V., & Estanto, B. H. (2023). Sentiment Analysis of Towards Electric Cars using Naive Bayes Classifier and Support Vector Machine Algorithm. Public Research Journal of Engineering, Data Technology and Computer Science, 1(1), 1–9. https://doi.org/10.57152/predatecs.v1i1.814

Tarigan, V. T., & Yusupa, A. (2024). Perbandingan Algoritma Maching Learning dalam Analisis Sentimen Mobil Listrik di Indonesia pada Media Sosial Twitter/X. Jurnal Informatika Polinema, 10(4), 479–490. https://doi.org/10.33795/jip.v10i4.5130