Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 2, April 2025, Page. 539-556
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4195

Geo-Sentiment Analysis of Public Opinion of X Users towards the Documentary Film Dirty Vote using the Bidirectional Long Short-Term Memory Method

Syifa Salsabila*[1], Yuliant Sibaroni[2], Sri Suryani Prasetiyowati[3]

[1,2,3]School of Computing, Telkom University, Indonesia

Email: [1]SyifaSalsabilaa@student.telkomuniversity.ac.id

Received : Dec 9, 2024; Revised : Jan 5, 2025; Accepted : Jan 6, 2025; Published : Apr 26, 2025

Abstract

Presidential elections held every five years, often generates significant public discourse. The 2024 presidential election saw the release of the documentary Dirty Vote, which raised allegations of electoral fraud and sparked polarized opinions on social media, especially on X. This study aims to analyze public sentiment toward Dirty Vote using geo-sentiment analysis and the Bidirectional Long Short-Term Memory (Bi-LSTM) model. Data were collected from geotagged tweets, with sentiment classified as positive, negative, or neutral. The research explored various data processing techniques, including TF-IDF for feature extraction, FastText for feature expansion, and balancing methods like SMOTE and class weighting to address data imbalance. Results showed that the baseline Bi-LSTM model achieved an accuracy of 71.57% and an F1-Score of 74.05%. When enhanced with TF-IDF and FastText, accuracy increased to 77.07%, though the F1-Score dropped slightly to 72.95%. Applying SMOTE resulted in a decrease in accuracy to 76.45%, but significantly improved the F1-Score to 74.93%. Exploratory data analysis revealed that negative sentiment was most concentrated in Java Island, particularly Jakarta, and peaked during February 2024, coinciding with the documentary's release and the election period. This study significantly contributes to understanding how geographic locations influence public opinion on sensitive political issues. A lack of understanding of geographically-based sentiment patterns can hinder identifying regional needs, leading to poorly targeted policies. By integrating data analysis methods with geographical approaches, this research provides deep insights for designing more effective, data-driven public intervention strategies and supports policymaking that is more responsive to the dynamics of public opinion.

Keywords : *Bi-LSTM, Class Weight, Dirty Vote, FastText, Geo-sentiment analysis, SMOTE, TF-IDF*

1. INTRODUCTION

Indonesia has adopted a democratic system since its independence in 1945. One of the key procedures in this democracy is the General Election (Pemilu), which holds significant meaning. The purpose of holding elections is to ensure that the transfer of power occurs safely and orderly, to uphold the sovereignty of the people, and to protect the human rights of every citizen [1], [2]. The presidential and vice-presidential elections are held every five years, making them a popular public event that is always widely discussed. Various perspectives emerge from different segments of society regarding the presidential and vice-presidential candidates. Amidst the festive atmosphere of this democratic celebration, the release of the documentary Dirty Vote on February 11, 2024, during the election's quiet period, was like a bomb exploding in the middle of the celebration. This film raises issues regarding alleged fraud in the 2024 Indonesian Presidential Election media Indonesia. The documentary's release sparked both support and opposition, generating diverse public opinions on social media, especially on X, which plays a significant role in shaping political views and broadening public discussions on political issues [3].

To understand the reviews of the documentary Dirty Vote on social media X, a sentiment analysis approach can be used to analyze the opinions and emotions contained in the sentences [4] and to uncover the sentiment polarity in the text, whether it is positive, neutral, or negative [5].

There has been extensive research on sentiment analysis [6], [7], [8], which has become a rapidly developing research topic, with many studies exploring its various aspects [9], [10]. One such study [11] by Sinanto et al. compared feature selection algorithms in sentiment analysis of movie reviews. Another study [12] analyzed sentiment regarding the 2024 election using Twitter data, employing the Naïve Bayes algorithm. Additionally, several previous studies have applied sentiment analysis using deep learning methods, which utilize layered architectures to handle complex and multidimensional data [13], [14]. For example, research by [15], [16] and [17] successfully demonstrated high accuracy using deep learning models such as Bi-LSTM. Overall, Bi-LSTM has the advantage of combining information from both the past and the future to understand context more comprehensively, thereby improving accuracy in natural language processing tasks [18], [19].

Another study [20] also conducted sentiment analysis on the performance of the General Election Commission (KPU) ahead of the 2024 election. Most previous studies tend to focus on sentiment analysis without incorporating the geographical dimension. To gain a more comprehensive understanding of public sentiment, there is a need to expand sentiment analysis by considering the geographic aspect, known as geo-sentiment analysis. This analysis can provide deep insights into how geographic location influences public sentiment [21].

Several researchers have demonstrated the potential of geo-sentiment analysis in providing deeper insights into public opinion [22], [23]. For example, in a study by Mushofy et al. [24], sentiment analysis was focused on COVID-19 vaccination using geo-tagged Twitter data, and the results showed that the dominant sentiment was positive, concentrated in the Karawang region. Another study by Tau Hu et al. [25] data to analyze sentiment using various techniques, including spatial and temporal analysis, as well as the Local Indicators of Spatial Association (LISA) method to identify spatial clusters. They also applied the Latent Dirichlet Allocation (LDA) model to classify geo-tweets. Through geo-sentiment analysis, insights can be gained regarding the differences in sentiment distribution patterns based on users' geographic locations [25]. This also helps in understanding the geographical impact on users' emotions in a specific context [26].

This study addresses a gap in sentiment analysis research, which has predominantly focused on non-geographical aspects. While most previous studies have utilized machine learning and deep learning algorithms to analyze public opinion on social media, the geographical dimension, which significantly influences public perception, remains underexplored. By implementing geo-sentiment analysis, this research provides insights into how geographic location shapes sentiment toward the documentary "Dirty Vote" and identifies specific sentiment patterns across various regions. Additionally, this study enhances sentiment analysis methodologies by integrating the Bi-LSTM model with data imbalance handling techniques, such as class weight and SMOTE, to improve sentiment analysis performance on complex and imbalanced datasets.

The problem formulation of this study is to evaluate the performance of the Bi-LSTM model in analyzing geo-sentiment from Twitter data related to the Dirty Vote documentary, and to understand the geographical distribution of public sentiment toward the film across various regions in Indonesia. This study is limited to sentiment analysis of public opinion on Twitter regarding Dirty Vote using the Bi-LSTM method. The analysis focuses on the collection of geolocation data from Twitter users, extraction of relevant tweets about Dirty Vote, and sentiment classification (positive, negative, or neutral) using the Bi-LSTM model.

The objective of this study is to analyze the geo-sentiment of X users toward the Dirty Vote documentary using the Bi-LSTM model, and to evaluate the model's performance in sentiment

classification. This study also aims to compare the effectiveness of various data imbalance handling techniques, such as the use of class weight and SMOTE, in improving the performance of the Bi-LSTM model on geo-sentiment data. Furthermore, this research seeks to deepen the understanding of how geographic location influences public sentiment, contributing to the development of geo-sentiment analysis methodologies. The findings of this study are expected to provide practical insights for political communication strategies and the analysis of public opinion on socially sensitive issues.

2. METHOD

Figure 1 shows the system design in this study, which aims to provide a clear depiction of the system developed.

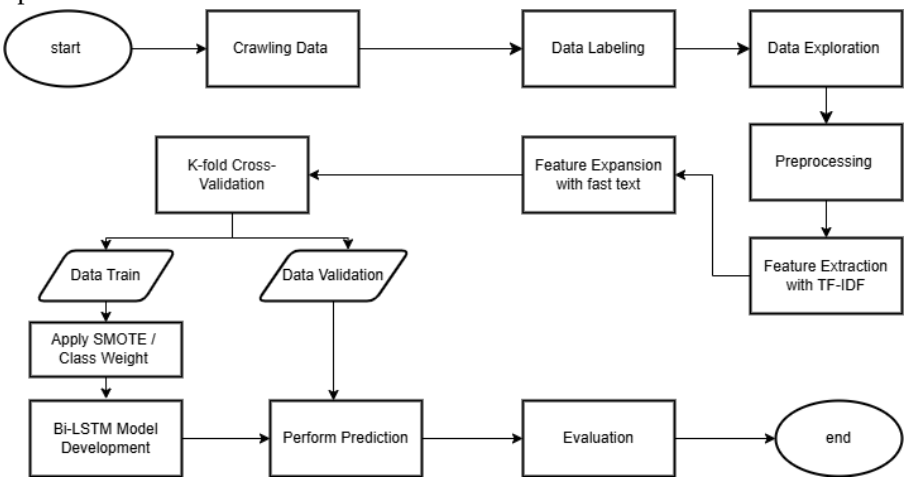

Figure 1. System Design Flowchart

2.1. Crawling Data

The data for this study were collected from tweets by users of the X application. The crawling process was carried out using the open-source Application Programming Interface (API) provided by the X application, with Python programming language. The collected tweets were from X users containing geolocation information from various regions, along with keywords or topics related to Dirty Vote.

2.2. Data Labeling

The collected data will then undergo labeling. Data labeling will be performed on each tweet and classified into several sentiment label categories: "Positive", "Negative", and "Neutral", with the aim of identifying the sentiment of each text in the dataset [24]. Data Exploration

The data collected in this study can be further analyzed to gain a better understanding of the sentiment and the geographic distribution of public opinion regarding the Dirty Vote documentary. The data exploration involves calculating the percentage of tweets with positive, negative, and neutral sentiments, as well as comparing sentiment across geographic regions. A histogram will be used to visualize the number of tweets with positive, negative, and neutral sentiments. As in the study by [27], a distribution map helps to understand the spread of sentiment across different regions. This visualization identifies sentiment distribution patterns and allows for comparison of sentiment across geographic areas.

2.3. Preprocessing

The collected data is then processed through the preprocessing or data preparation stage. This stage aims to clean and improve the quality of the data to ensure it is ready for subsequent analysis. With clean and high-quality data, the analysis results will be more accurate, efficient, and easier to interpret. The following steps are performed:

2.3.1. Data Cleaning

In this stage, the data is cleaned by removing punctuation, symbols, numbers, and unnecessary emojis that do not contribute to the analysis.

2.3.2. Case Folding

In this stage, all words in the data are converted to lowercase to make the data uniform, making it easier to compare and speeding up the analysis process.

2.3.3. Tokenization

This stage involves the process of splitting sentences into tokens or individual words that are separated by spaces.

2.3.4. Normalization

The normalization stage is carried out to ensure that the text data follows a consistent format. This process includes converting non-standard words, abbreviations, or terms into standard words according to the general spelling rules of the Indonesian language (PUEBI), using an Indonesian lexicon dictionary as a tool to improve the quality of the processed text data.

2.3.5. Stopword Removal

This stage involves the removal of words considered to have little significant impact or non-specific meaning. The process uses the NLTK (Natural Language Toolkit) library from Python. The NLTK stopword list for the Indonesian language serves as the base, and it is extended with additional words such as slang, informal expressions, or frequently occurring words that do not contribute to the analysis.

2.3.6. Stemming

In this stage, the process of transforming a word with affixes into its root form is performed. The stemming process uses the StemmerFactory from the Sastrawi library, a Python library for natural language processing that specifically supports stemming for the Indonesian language.

2.4. Feature Extraction with TF-IDF

After completing the preprocessing stage, the next process is the weighting of sentences using the TF-IDF (Term Frequency - Inverse Document Frequency) method. TF-IDF is a commonly used method in natural language processing and information retrieval that assigns weights (W_{ki}) to a word (k) in a document (i). Term Frequency (tf_{ki}) is a metric that measures how frequently a word appears in a document, while Inverse Document Frequency (IDF) indicates the importance of a word within a collection of documents, calculated using the logarithm of the total number of documents (N) divided by the number of documents containing the word (n_k). The main focus of the TF-IDF method is to determine how important a word (term) is in a document within the context of a larger document [28]. The higher the frequency of a word in a document, the higher its TF value, which reflects the relevance of that word to the document being analyzed. The word's weight is calculated based on two main factors: the term frequency (TF) and the inverse document frequency (IDF) [29]. The rarer a word appears across documents, the higher its IDF value, which helps emphasize words that are more specific and relevant to the particular document [30]. Below is the formula for calculating TF-IDF [31]:

$$W_{ki} = tf_{ki} * log\left(\frac{N}{n_k}\right) \quad (1)$$

This feature extraction process generates a numerical matrix that represents the documents, with each element indicating the weight of a term. With the parameter max_features set to 5000, only the 5000 terms with the highest weights are retained. This matrix can then be used as input for machine learning models to understand patterns and information in the text.

2.5. Feature Expansion with FastText

FAIR, an AI research group from Facebook, developed the FastText library to assist in word representation and text classification tasks [32]. In this study, the application of FastText for feature expansion is used to address challenges in classifying tweets. FastText has exceptional capabilities in generating word representations and classifying sentences, particularly for rare words [31]. This capability is achieved by leveraging character-level information. FastText works by generating vector representations for each word in a text corpus [27]. The model is then trained to use previous words to predict the next word in the text's context. The advantage of FastText lies in its ability to understand words. This is made possible by breaking words into subwords or n-grams and treating each subword as a separate unit [33]. In this way, FastText can recognize unknown or rare words by matching small parts with words already present in the model's database. The result of this process is a fixed-dimensional numeric vector that encompasses semantic information and contextual relationships between words, providing richer feature representations for use in classification models.

2.6. K-fold Cross Validation

Cross-validation is a common method in testing classification algorithms, which divides the training and testing data into several groups. As explained by [34], one form of cross-validation is k-fold cross-validation. In k-fold cross-validation, the data is divided into k equally sized subsets (folds). The architecture of K-Fold Cross-validation is shown in Figure 2, as described in [35].

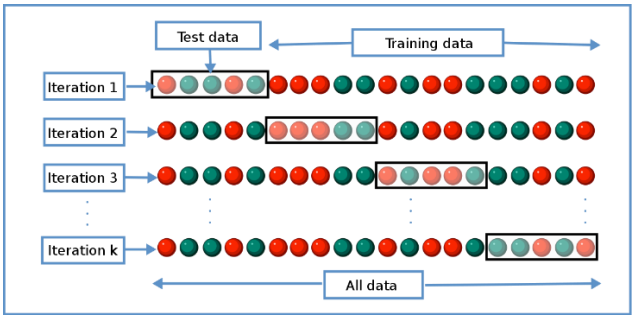

Figure 2. K-Fold Cross Validation Architecture [21].

In this study, the value of k used is 5. Therefore, the data will be divided into 5 subsets, and the training and testing processes will be performed five times [29]. In each iteration, one subset is used as the test data, while the remaining four subsets are used as the training data. This approach ensures that the model not only learns from specific data but also generalizes to new data, avoiding overfitting, and providing a more consistent performance estimate.

2.7. Bi-LSTM Model Development

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 2, April 2025, Page. 539-556
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4195

Bidirectional Long Short Term Memory (Bi-LSTM) is a derivative of LSTM. Bi-LSTM is a type of recurrent neural network architecture consisting of two LSTM components, one moving forward and the other moving backward. The forward LSTM utilizes contextual information from the past to read the input text in chronological order. Conversely, the backward LSTM reads the text in reverse order, storing contextual information from the future [18]. The Bi-LSTM architecture as shown in Figure 3, as described in [36].

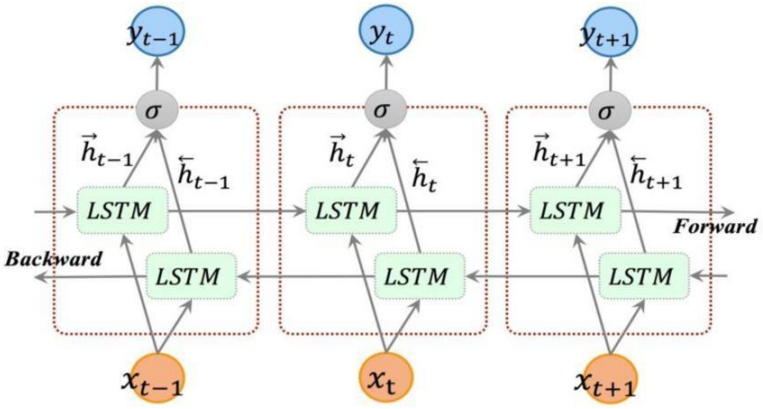

Figure 3. BiLSTM Architecture [36]

The forward LSTM output sequence h $\vec{}$ is obtained through a forward flow. On the other hand, the backward LSTM output sequence h $\overleftarrow{}$ is computed using a reverse flow [37]. The output $\boldsymbol{y_t}$, obtained from both the ForwardLSTM and BackwardLSTM, is then combined using the function σ to provide a more comprehensive understanding of the entire text, allowing Bi-LSTM to achieve higher accuracy in various NLP tasks. The BiLSTM model development in this study begins with the creation of an embedding layer to represent words as numerical vectors. The input parameter is set to 10,000 to limit the maximum number of words processed, and the feature vector size is set to 128. The maximum sequence length is set to 50 to ensure input consistency. This model includes a BiLSTM layer with 64 memory units in each direction to process information forward and backward. Additionally, a dropout layer with a rate of 0.5 is applied to reduce the risk of overfitting, and a dense layer with 2 units using the softmax activation function is used to generate class probabilities. The model is optimized using the Adam algorithm, with the categorical_crossentropy loss function for handling multiclass classification. With this parameter configuration, the BiLSTM model is expected to accurately generate sentiment labels based on patterns learned from the training data.

2.8. SMOOTE

One technique for balancing the data distribution in the minority class is the SMOTE (Synthetic Minority Over-sampling Technique) method. This method works by selecting samples from the minority class until the number of samples becomes balanced with the majority class samples [38]. The technique involves randomly selecting samples from the minority class and then finding the k-nearest neighbors using the k-Nearest Neighbors (k-NN) algorithm. After identifying the nearest neighbors, SMOTE generates new data by performing linear interpolation between the minority sample and its neighbors. This process creates new, more varied, and more representative examples in the feature space of the minority class, allowing the model to learn more patterns from that class. SMOTE not only increases the number of minority class data but also improves the representation of that class in the feature space, enabling the model to more effectively identify the characteristics of the minority class. Research has shown that SMOTE produces better results compared to other

oversampling and undersampling methods in improving model accuracy on imbalanced datasets, particularly in the context of educational data mining [39]. Furthermore, another study [40] demonstrated that SMOTE can increase recall for the minority class, improve the balance between precision and recall, and enhance the F1-score, providing a more representative evaluation of the model's performance on imbalanced datasets. Figure 4 illustrates the mechanism of the SMOTE algorithm as referenced in [41].

Algorithm 1 Description of the SMOTE Over-Sampling Algorithm

Repeat

Select a minority sample, denoted as X_o, randomly.

Identify the K-nearest neighbors of X_o from the minority class samples.

Randomly choose one of the K-nearest neighbors of X_o from the previous step, and label the selected sample as X_k, where k is the rank of the chosen neighbor.

Perform linear interpolation between X_o and the selected neighbor X_k to generate a new synthetic sample, z, using the formula:

$z = X_o + w (X_k - X_o)$, where w is a uniformly distributed random variable between [0,1].

until M synthetic samples are created.

Figure 4. SMOTE Algorithm [41]

2.9. Class Weight

This study explores methods for handling imbalanced data. One machine learning technique to address imbalanced data is class weight. This method is applied by assigning greater weight to samples from the less dominant class [42]. With this weighting approach, the bias in the model's predictions toward the dominant class can be minimized.

2.10. Classification Performance Metrics

One way to evaluate the performance of a classification model is by using evaluation metrics based on the model's prediction results. Below is an explanation of the evaluation metrics:

2.10.1. Accuracy

Accuracy is the percentage of correct predictions overall, based on the total classification data.

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

2.10.2. Precision

Precision is the ratio of correctly classified positive instances to the total number of instances expected to be positive.

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

2.10.3. Recall

Recall is the ratio of correctly classified positive instances to the total number of actual positive instances.

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

2.10.4. F1 Score

F1 Score Score is the harmonic mean of precision and recall, used to measure the balance between the two.

$$F1\ Score = \frac{2 \times (precision \times recall)}{(precision + recall)} \quad (5)$$

To calculate the formulas above, a confusion matrix is used. A confusion matrix is a table that maps the predicted results of a classification model into four categories: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). These four categories can be represented as shown in Table 1 below.

Table 1. Confusion Matrix

Confusion Matrix		Actual Values	
		Positif	Negatif
Predicted Values	Positive	TP	FP
	Negative	FN	TN

3. RESULT

3.1. Data Crawling

The data collected consists of 2,603 tweets from users of the X application, including geolocation information and keywords related to Dirty Vote. The data collection process was conducted using an open-source API implemented with the Python programming language.

3.2. Data Labeling

Based on the labeling results, the number of entries in each class is 1,799 for negative, 417 for neutral, and 387 for positive. Below, Table 2 shows an example of the labeled data.

Table 2. Dataset

Date	Text	Location	Latitude	Longitude	Sentimen
Tue Feb 13 22:11:46	Bangke bener ..!! Masih mau bilang film dirty vote fitnah ??	Jakarta Selatan, DKI Jakarta	-628381815	1068048633	Negatif
Tue Feb 13 21:59:34 +0000 2024	Dirty vote bukan fBerikut qilm imajiner jika tidak ada fakta dan dokumen nyata mana bisa film dibuat	Bogor Barat, Indonesia	-657662075	106769581	Positif
Mon Feb 12 23:24:41 +0000 2024	@CNNIndonesia Kalo iya dirty vote hanya hoax dan framing kotor. Ganti aja dg dirty heart	Jakarta	-6175247	1068270488	Positif
Sun Feb 11 23:21:21 +0000 2024	@mamagemoybiasa Orang ToloL bin Bodoh saja yg ga paham isi film Dirty Vote!	Jakarta	-6175247	1068270488	Negatif
Wed Mar 20 10:43:24 +0000 2024	KPU hanya ikuti DIRTY VOTE Jokowi... BEJAT. https://t.co/qgS4g9JVNl	Bintaro Jaya	-627270985	106754634	Negatif

3.3. Data Exploration

The results of the data exploration can be seen in Figure 5, which shows the word cloud visualization depicting the words associated with sentiment in the dataset. From this visualization, key

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 2, April 2025, Page. 539-556
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4195

terms related to the sentiment in the dataset, particularly regarding the Dirty Vote documentary, can be identified.

Figure 5. Word Cloud Visualization

Next, the sentiment distribution map focusing on the Java Island region provides valuable insights into how public opinion on the documentary Dirty Vote is geographically spread. Java Island was selected as the main area of analysis due to the larger and more representative dataset gathered from this region. As seen in Figure 6, it is clear that negative sentiment dominates compared to positive sentiment, which may indicate dissatisfaction or critical reactions towards the Dirty Vote documentary. Furthermore, although there is a high level of negative sentiment across Java, large cities like Jakarta show a significant proportion of positive sentiment. This suggests that certain elements of the film, such as effective storytelling techniques or the addressing of social issues relevant to urban society, have been appreciated by certain groups in larger cities. Jakarta, as the center of government and media, may be more open to controversial political issues, leading the film to be received more positively by segments of the population with a more critical understanding of political dynamics.

Specifically, as seen in Figure 7, Jakarta appears as the area with the highest concentration of negative sentiment. This phenomenon could be attributed to more complex factors, including greater political tension in the capital, which may worsen the negative perception of the film's content. Jakarta, as the political center and home to a diverse population with varying social and political backgrounds, may be more polarized in responding to the messages conveyed in this film. As the hub of political and media activity, Jakarta is also more exposed to criticism or analysis of films related to election and democracy issues.

Figure 6. Sentiment Distribution Map of Java Island

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 2, April 2025, Page. 539-556
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4195

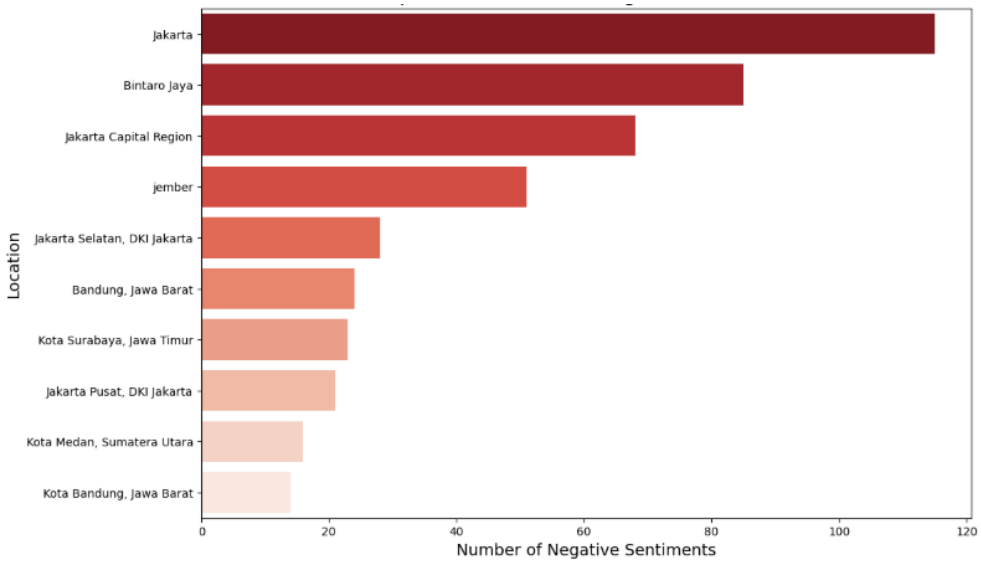

Figure 7. Ten Cities with the Highest Negative Sentiment

These results reflect a high concentration of negative sentiment towards the Dirty Vote documentary in areas with large populations and complex social and political dynamics. Jakarta, as the capital city, is not only the political center but also the hub of intense media, economic, and social activity. Outside Jakarta, surrounding areas such as Bogor, Depok, and Bekasi also show significant levels of negative sentiment. This may reflect the influence of the nearby capital, both in terms of information flow and the spread of political opinions. As regions directly bordering Jakarta, the populations in these satellite cities are often exposed to the political dynamics happening in the capital, which may worsen their negative response to the documentary. Overall, this sentiment distribution map provides a clearer picture of how various segments of society in Java Island respond to the documentary. The high concentration of negative sentiment in Jakarta and its surrounding areas is influenced not only by the film's content but also by external factors related to the intensity of politics, social diversity, and the strong media influence in the region. This illustrates how public sentiment can be shaped in a highly dynamic social and political context, especially in areas with a high concentration of political and social activity.

Another result of the data exploration is the temporal sentiment analysis, shown in Figure 8. This visualization provides an overview of the dynamic fluctuations in sentiment, which can be used to identify patterns in the changes of public opinion.

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 2, April 2025, Page. 539-556
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4195

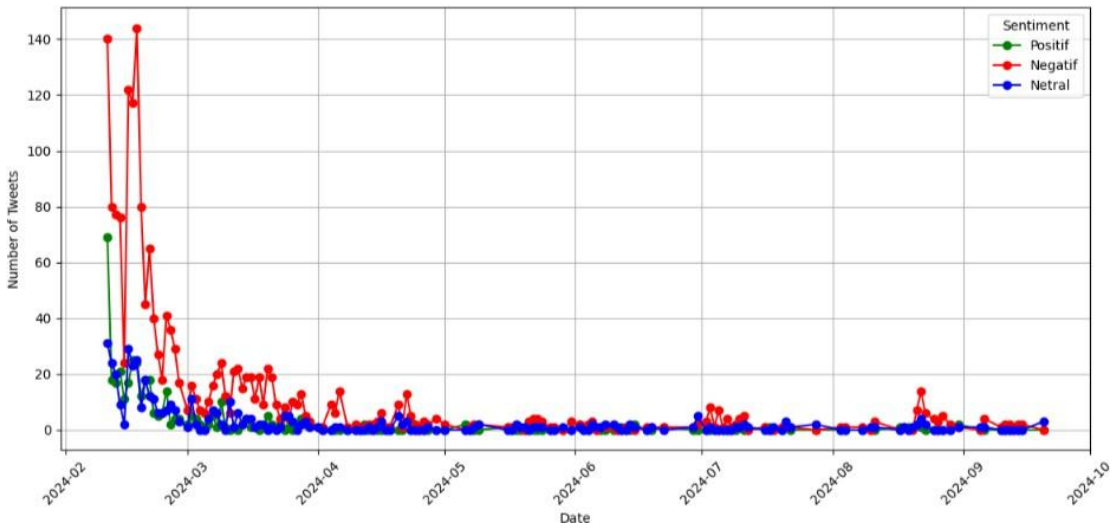

Figure 8. Sentiment Trends Over Time

One of the key findings from this graph is the significant spike in negative sentiment in February, coinciding with several important events, namely the release of the Dirty Vote film and the start of the 2024 Indonesian Presidential Election. This sharp rise in negative sentiment is likely caused by the proximity of the film's release to the beginning of the election, which is a major event with broad political implications. Dirty Vote, which addresses controversial issues surrounding electoral practices and politics in Indonesia, has likely triggered emotional responses from a society already divided in intense political debates. In February, the increasing political tension associated with the 2024 election likely worsened perceptions of the film, with many people feeling that the film was overly critical or misaligned with their political views.

3.4. Preprocessing

The preprocessing process was performed to improve data quality and produce data that is ready for further analysis. Table 3 below shows a comparison of the data before and after preprocessing.

Table 3. Data Before and After Preprocessing

Before Preprocessing	After Preprocessing
Bangke bener ..!! Masih mau bilang film dirty vote fitnah ??	bangkai film dirty vote fitnah
KPU hanya ikuti DIRTY VOTE Jokowi...BEJAT. https://t.co/qgS4g9JVNl	komisi pilih umum ikut dirty vote jokowi bejat
Dokumenter Eksil & Dirty Vote Masuk 50 Film Global dengan Rating Tertinggi Versi Letterboxd â€¼ï¸ https://t.co/xMCdpZX3vH	dokumenter eksil dirty vote masuk 50 film global rating tinggi versi letterboxd
bajingan ga kuat nonton dirty vote gw jijik bgt	bajing kuat tonton dirty vote gue jijik banget
Teringat Judul Novel : *RUBUHNYA NEGERI KAMI* NEGERI PARA BEDEBAH NEGERI PARA BANDIT *DIRTYVOTE* *GENERASI	Ingat judul novel rubuh negeri negeri bedebah negeri bandit dirty vote generasi cundang

PECUNDANG* https://t.co/DWEqi5egcF

3.5. Testing Schemes

This study applies three testing schemes with the aim of determining the optimization of the BiLSTM model's accuracy with respect to TF-IDF feature extraction, FastText feature expansion, and comparing data imbalance handling methods using Class Weight and SMOTE techniques. Table 4 outlines the explanations and objectives of each testing scheme applied.

Table 4. Testing Scheme

Stage	Testing Scheme	Explanation	Objective
1	Baseline BiLSTM	Testing the BiLSTM model with the baseline scheme, without additional feature extraction or data balancing techniques.	To evaluate the accuracy performance of the baseline BiLSTM model as the primary comparison.
2	BiLSTM-Enhanced (TF-IDF + FastText)	Testing the BiLSTM model with TF-IDF feature extraction and FastText feature expansion.	Evaluating the performance improvement of the BiLSTM model with the combination of TF-IDF and FastText techniques.
3	BiLSTM-Enhanced Handling Imbalance	Testing the BiLSTM model with TF-IDF feature extraction, FastText expansion, and data balancing using Class Weight and SMOTE.	Measuring the impact of data balancing techniques in handling class imbalance in the sentiment dataset.

3.5.1. **Baseline BiLSTM**

In this first scheme, the baseline performance of the BiLSTM model is tested. Based on the results of the 5-fold cross-validation, shown in Table 5, the model's performance demonstrates a good variation across each fold.

Table 5. Baseline BiLSTM Testing Results

Fold	Akurasi	Presisi	Recall	F1-Score
Fold 1	0.7869	0.8098	0.7869	0.7962
Fold 2	0.6967	0.8153	0.6967	0.7304
Fold 3	0.7716	0.7961	0.7716	0.7815
Fold 4	0.6558	0.8327	0.6558	0.6888
Fold 5	0.6673	0.8447	0.6673	0.7056
Avararage	0.7157	0.8197	0.7157	0.7405

The average accuracy of the model reached 71.57%, with the highest accuracy in fold 1 (78.69%) and the lowest accuracy in fold 4 (65.58%). The model demonstrated stable performance in correctly classifying tweets, as indicated by the average precision of 81.97%. Meanwhile, the average F1-Score was recorded at 74.05%, with the highest value in fold 1 (79.62%), showing a balance between precision and recall. On the other hand, the average recall was 71.57%, with the lowest value in fold 4 (65.58%), indicating the model's difficulty in handling data imbalance.

3.5.2. **Result of TF-IDF Extraction and FastText Feature Expansion Application**

In the second testing scheme, TF-IDF feature extraction is applied in combination with feature expansion using the FastText model. The results of this test scheme can be seen in Table 6, which

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4195

compares the performance of the baseline model from scheme 1 with the model that applies TF-IDF feature extraction and FastText feature expansion (Scheme 2).

Table 6. Baseline BiLSTM vs BiLSTM-Enhanced Comparison

Nama	Akurasi	F1-Score
Baseline Model	71.57%	**74.05%**
BiLSTM-Enhanced (TF-IDF + FastText)	**77.07%**	72.95%

Based on the results obtained in Table 5, the accuracy in Scheme 2 reached 77.07%, which is higher than the accuracy of 71.57% in Scheme 1. This represents an accuracy increase of 5.50%. However, the F1-Score decreased, with Scheme 2 showing a result of 72.95%, which is lower than the 74.05% recorded in Scheme 1. The accuracy improvement in Scheme 2 occurred because the model was able to correctly classify only one class, particularly the majority class.

3.5.3. **Result of TF-IDF Extraction & FastText Expansion with Data Balancing Methods**

In Scheme 3, testing is conducted using data balancing methods, namely class weight and SMOTE. This scheme aims to compare the two data imbalance handling methods against the performance of the BiLSTM model with TF-IDF feature extraction and FastText feature expansion. The results of Scheme 3 testing can be seen in Table 7.

Table 7. Comparison of Class Weight and SMOTE

Model	**Class Weight**		**SMOTE**	
	Accuracy	F1 Score	Accuracy	F1 Score
Baseline Model	71.96%	72.33%	55.36%	58.14%
BiLSTM+ Tf-IDF + Fasstext	74.49%	75.15%	76.45%	74.93%

The SMOTE method achieved an accuracy of 76.45% and an F1-Score of 74.93%. Compared to the Class Weight method, which resulted in an accuracy of 74.49% and an F1-Score of 75.15%, SMOTE shows a significant improvement in accuracy, although the F1-Score for Class Weight is slightly higher. Overall, both methods resulted in improvements, but SMOTE provided a more significant increase in accuracy, while Class Weight performed slightly better in terms of F1-Score. This difference is due to the way the two techniques work. SMOTE generates synthetic data for the minority class, allowing the model to learn better from the underrepresented data, which has been shown to contribute to the accuracy improvement. In contrast, the Class Weight method adjusts the class weights based on the existing data distribution, helping the model pay more attention to the underrepresented class, resulting in a more balanced performance in terms of precision and recall, which is reflected in the slightly higher F1-Score compared to SMOTE.

4. DISCUSSIONS

From a geo-sentiment analysis perspective, the results of this study reveal key insights into how sentiment varies across different regions of Indonesia, particularly in response to the documentary Dirty Vote. The sentiment distribution map shows a notable concentration of negative sentiment in urban areas, especially Jakarta, suggesting that urban centers are more likely to be affected by political content. The temporal sentiment analysis, which tracks sentiment over time, indicates a significant spike in negative sentiment during February 2024, coinciding with the release of Dirty Vote and the 2024 presidential election.

Based on the results of all the testing schemes, several findings can be analyzed and compared. In the first scheme, the application of the BiLSTM model without additional data balancing techniques showed relatively good performance with an accuracy of 71.57% and an F1-Score of 74.05%. Although this model is effective for tweet classification in general, the data imbalance issue seems to affect the results, particularly in terms of recall, which tends to be lower in some folds. This indicates that while BiLSTM has the ability to process sequential data, it still struggles with handling significant class imbalance in this dataset.

In the second scheme, where TF-IDF is used for feature extraction and FastText for feature expansion, the accuracy increased from 71.57% to 77.07%. However, despite the accuracy improvement, the F1-Score decreased (from 74.05% to 72.95%). This suggests that while the addition of FastText features can improve word representation in the vector space, it does not necessarily improve the balance between precision and recall. The use of additional features often introduces extra complexity, which may not always improve the balance of predictions across classes. The decrease in F1-Score could be due to the fact that a more complex model risks overfitting or generating more false positives or false negatives, affecting the balance between precision and recall.

In the third scheme, as shown in Figure 9, the SMOTE method demonstrated a more significant improvement in both accuracy and F1-Score. The use of SMOTE to balance class distribution provided better results compared to the Class Weight technique, with SMOTE achieving an accuracy of 76.45% and an F1-Score of 74.93%, compared to the Class Weight results, which reached only 74.49% and 75.15%, respectively.

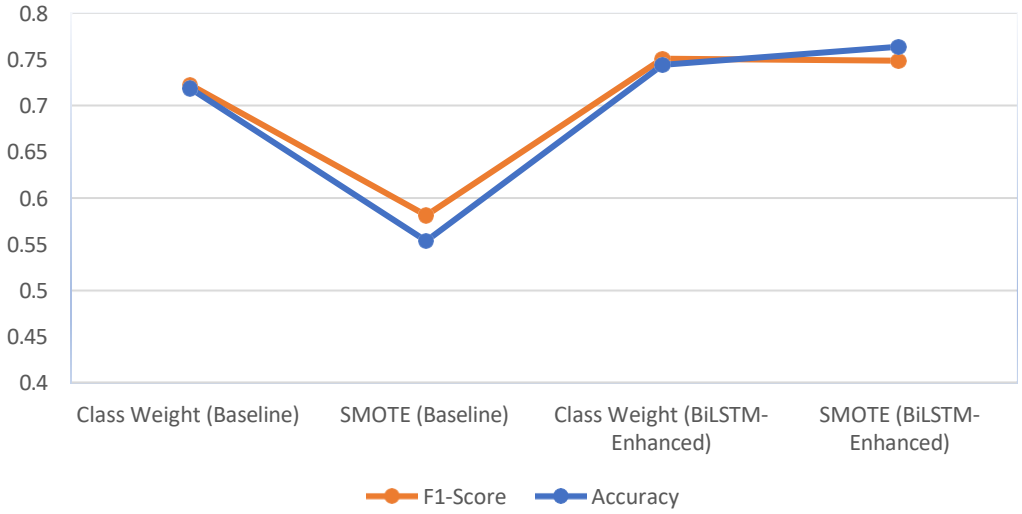

Figure 9. Comparison of Class Imbalance Handling Methods with Class Weight vs SMOTE

This shows that the SMOTE technique is more effective in improving model performance on imbalanced datasets, as SMOTE generates synthetic samples for the minority class, allowing the model to learn better at distinguishing between the majority and minority classes. Meanwhile, Class Weight, which works by adjusting the class weights based on the existing data distribution, also shows an improvement compared to the baseline, but the results are not as significant as those obtained with SMOTE. This reflects that SMOTE provides a stronger approach to addressing data imbalance issues. Additionally, the application of feature expansion techniques on BiLSTM, using TF-IDF and FastText, demonstrates that while embedding-based features can provide richer word representations, the results are highly dependent on the balancing technique used. In other words, although TF-IDF and FastText provide better word representations, the application of data balancing techniques such as SMOTE has a much greater impact on improving classification performance.

Previous research on geo-sentiment analysis has largely focused on specific social issues, such as responses to government programs or community dynamics [24], [25]. However, this study expands that approach by integrating BiLSTM techniques with data balancing methods to develop a model for location-based sentiment analysis on complex datasets. The improvements in accuracy and F1-Score achieved in this study demonstrate the effectiveness of the applied approach, particularly in addressing the challenge of data imbalance, which is often a limitation in sentiment analysis. These findings contribute to advancing methodologies in sentiment analysis, emphasizing the importance of integrating natural language processing (NLP) techniques with geographical dimensions.

5. CONCLUSION

This research highlights the importance of geo-sentiment analysis in understanding public opinion, especially regarding politically sensitive topics like the documentary film Dirty Vote. By incorporating geographic information, we can observe how sentiment varies across different regions, with a notable concentration of negative sentiment in urban areas such as Jakarta. This geographic perspective is crucial for gaining deeper insights into how public sentiment is influenced by regional factors and how it evolves over time, particularly during significant events such as the documentary release and presidential elections.

Based on the results of the three testing scenarios conducted, it can be concluded that the application of data balancing techniques has a significant impact on the performance of the classification model. The use of the SMOTE technique has proven to provide better results compared to Class Weight in terms of accuracy and F1-Score, particularly in imbalanced datasets. While TF-IDF and FastText provide richer and deeper word representations, the application of data balancing techniques such as SMOTE has a greater overall impact on improving classification performance. The results of this study indicate that when addressing class imbalance, data balancing techniques are a key factor that supports the effectiveness of the classification model, even more significantly than enriching embedding feature representations. Therefore, to improve model performance on imbalanced datasets, a combination of techniques such as SMOTE and FastText can be an optimal approach.

This research contributes to the field of informatics by showing that the combination of the Bi-LSTM model with data balancing techniques can significantly improve the accuracy of sentiment analysis on geotagged data. The integration of geographic information with sentiment analysis offers a new approach that can be applied to various fields. Future research can explore other deep learning models, such as transformers or hybrid models that can improve classification performance.

REFERENCES

[1] F. Esfandiari and N. Hidayah, “General Elections in Indonesia : Between Human Rights and Constitutional Rights,” European Alliance for Innovation n.o., Feb. 2021. doi: 10.4108/eai.1-7-2020.2303622.

[2] Inayatulloh, I. K. Hartono, and D. L. Kusumastuti, “The Blockchain Conceptual Model to Prevents Fraud and Increase Transparency in the Presidential Election in Indonesia,” *2023 IEEE 9th International Conference on Computing, Engineering and Design, ICCED 2023*, 2023, doi: 10.1109/ICCED60214.2023.10425635.

[3] K. Kurnia, “Measuring the Concept of Deliberative Democracy in the Indonesian Election Supervision System,” Dec. 2020, doi: 10.4108/EAI.26-9-2020.2302602.

[4] F. Agung, J. Ayomi, and K. E. Dewi, “Analisis Emosi pada Media Sosial Twitter Menggunakan Metode Multinomial Naive Bayes dan Synthetic Minority Oversampling Technique,” *Komputa : Jurnal Ilmiah Komputer dan Informatika*, vol. 12, no. 2, pp. 9–19, Sep. 2023, doi: 10.34010/KOMPUTA.V12I2.9454.

[5] R. DAHIYA, A. Dhankhar, and A. DHANKHAR, "Sentimental Analysis Of Social Networks :A Comprehensive review(2018-2023)," Nov. 20, 2023. doi: 10.21203/rs.3.rs-3625172/v1.

[6] D. F. Putra and Y. Sibaroni, "Sentiment Analysis For The 2024 Presidential Election (Pilpres) Using BERT CNN," *Eduvest - Journal of Universal Studies*, vol. 4, no. 11, pp. 11042–11056, Nov. 2024, doi: 10.59188/EDUVEST.V4I11.49961.

[7] L. Chaudhary, N. Girdhar, D. Sharma, J. Andreu-Perez, A. Doucet, and M. Renz, "A Review of Deep Learning Models for Twitter Sentiment Analysis: Challenges and Opportunities," *IEEE Trans Comput Soc Syst*, vol. 11, no. 3, pp. 3550–3579, Jun. 2024, doi: 10.1109/TCSS.2023.3322002.

[8] R. Cahyani, I. S. Rozas, and N. Yalina, "Analisis Sentimen Pada Media Sosial Twitter Terhadap Tokoh Publik Peserta Pilpres 2019," *MATICS: Jurnal Ilmu Komputer dan Teknologi Informasi (Journal of Computer Science and Information Technology)*, vol. 12, no. 1, pp. 79–86, Apr. 2020, doi: 10.18860/MAT.V12I1.8356.

[9] R. F. Ananda, A. Syahri, and F. N. Hasan, "SENTIMENT ANALYSIS OF CUSTOMER SATISFACTION IN GOJEK AND GRAB APPLICATION REVIEWS USING THE NAIVE BAYES ALGORITHM," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 1, pp. 233–241, Feb. 2024, doi: 10.52436/1.JUTIF.2024.5.1.1680.

[10] G. Radiena and A. Nugroho, "ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN APLIKASI KAI ACCESS MENGGUNAKAN METODE SUPPORT VECTOR MACHINE," *Jurnal Pendidikan Teknologi Informasi (JUKANTI)*, vol. 6, no. 1, pp. 1–10, Apr. 2023, doi: 10.37792/JUKANTI.V6I1.836.

[11] I. Sinanto Ate, A. Nuraminah, and P. Studi, "Komparasi Algoritma Feature Selection Pada Analisis Sentimen Review Film," *Jurnal Ilmiah Teknik Informatika dan Komunikasi*, vol. 2, no. 2, pp. 96–102, Jul. 2022, doi: 10.55606/JUITIK.V2I2.326.

[12] S. Puad, G. Garno, and A. S. Y. Irawan, "ANALISIS SENTIMEN MASYARAKAT PADA TWITTER TERHADAP PEMILIHAN UMUM 2024 MENGGUNAKAN ALGORITMA NAÏVE BAYES," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 3, pp. 1560–1566, Oct. 2023, doi: 10.36040/JATI.V7I3.6920.

[13] A. Raup, W. Ridwan, Y. Khoeriyah, S. Supiana, and Q. Y. Zaqiah, "Deep Learning dan Penerapannya dalam Pembelajaran," *JIIP - Jurnal Ilmiah Ilmu Pendidikan*, vol. 5, no. 9, pp. 3258–3267, Sep. 2022, doi: 10.54371/JIIP.V5I9.805.

[14] Sudianto, P. Wahyuningtias, H. W. Utami, U. A. Raihan, H. N. Hanifah, and Y. N. Adanson, "COMPARISON OF RANDOM FOREST AND SUPPORT VECTOR MACHINE METHODS ON TWITTER SENTIMENT ANALYSIS (CASE STUDY: INTERNET SELEBGRAM RACHEL VENNYA ESCAPE FROM QUARANTINE)," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 1, pp. 141–145, Feb. 2022, doi: 10.20884/1.JUTIF.2022.3.1.168.

[15] L. Xiaoyan and R. C. Raga, "BiLSTM Model With Attention Mechanism for Sentiment Classification on Chinese Mixed Text Comments," *IEEE Access*, vol. 11, pp. 26199–26210, 2023, doi: 10.1109/ACCESS.2023.3255990.

[16] A. I. Safitri and T. B. Sasongko, "SENTIMENT ANALYSIS OF CYBERBULLYING USING BIDIRECTIONAL LONG SHORT TERM MEMORY ALGORITHM ON TWITTER," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 2, pp. 615–620, Apr. 2024, doi: 10.52436/1.JUTIF.2024.5.2.1922.

[17] E. Subowo, F. A. Artanto, I. Putri, and W. Umaedi, "Algoritma Bidirectional Long Short Term Memory untuk Analisis Sentimen Berbasis Aspek pada Aplikasi Belanja Online dengan Cicilan," *JURNAL FASILKOM*, vol. 12, no. 2, pp. 132–140, Aug. 2022, doi: 10.37859/JF.V12I2.3759.

[18] A. Muzakir and U. Suriani, "Model Deteksi Berita Palsu Menggunakan Pendekatan Bidirectional Long Short-Term Memory (BiLSTM)," *Journal of Computer and Information Systems Ampera*, vol. 4, no. 2, 2023, doi: 10.51519/journalcisa.v4i2.397.

[19] H. Akbar, D. Aryani, M. Kadhim, M. Al-Shammari, and M. B. Ulum, "SENTIMENT ANALYSIS FOR E-COMMERCE PRODUCT REVIEWS BASED ON FEATURE FUSION AND BIDIRECTIONAL LONG SHORT-TERM MEMORY," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 5, pp. 1385–1391, Oct. 2024, doi: 10.52436/1.JUTIF.2024.5.5.2675.

[20] Fuad Amirullah, Syariful Alam, and M.Imam Sulistyo S, "Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naïve Bayes," *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 2, no. 3, pp. 69–76, Aug. 2023, doi: 10.55123/STORAGE.V2I3.2293.

[21] A. Almaslukh, A. Almaalwy, N. Allheeib, A. Alajaji, M. Almukaynizi, and Y. Alabdulkarim, "Top-k sentiment analysis over spatio-temporal data," *PeerJ Comput Sci*, vol. 10, p. e2297, Sep. 2024, doi: 10.7717/PEERJ-CS.2297/FIG-5.

[22] Z. Gong, T. Cai, J. C. Thill, S. Hale, and M. Graham, "Measuring relative opinion from location-based social media: A case study of the 2016 U.S. presidential election," *PLoS One*, vol. 15, no. 5, p. e0233660, May 2020, doi: 10.1371/JOURNAL.PONE.0233660.

[23] D. Tas and R. P. Sanatani, "Geo-located Aspect Based Sentiment Analysis (ABSA) for Crowdsourced Evaluation of Urban Environments," *arXiv preprint*, Dec. 2023, Accessed: Jan. 05, 2025. [Online]. Available: https://www.alphaxiv.org/abs/2312.12253v1

[24] A. Mushofy Anwary, A. ID Hadiana, and P. Nurul Sabrina, "ANALISIS SENTIMENT PENGGUNAAN VAKSIN COVID-19 MENGGUNAKAN GEO-TAGGED TWEETS DAN ALGORITMA NAIVE BAYES," *Informatics and Digital Expert (INDEX)*, vol. 3, no. 2, pp. 75–83, Nov. 2021, doi: 10.36423/INDEX.V3I2.876.

[25] T. Hu, B. She, L. Duan, H. Yue, and J. Clunis, "A systematic spatial and temporal sentiment analysis on geo-tweets," *IEEE Access*, vol. 8, pp. 8658–8667, 2020, doi: 10.1109/ACCESS.2019.2961100.

[26] E. Veltmeijer and C. Gerritsen, "SentiMap: Domain-Adaptive Geo-Spatial Sentiment Analysis," *Proceedings - 17th IEEE International Conference on Semantic Computing, ICSC 2023*, pp. 17–24, 2023, doi: 10.1109/ICSC56153.2023.00010.

[27] I. F. Ramadhy and Y. Sibaroni, "Analisis Trending Topik Twitter dengan Fitur Ekspansi FastText Menggunakan Metode Logistic Regression," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 1, p. 1, Feb. 2022, doi: 10.30865/jurikom.v9i1.3791.

[28] N. Silalahi, J. Josh, and G. Leonarde Ginting, "Analisa Sentimen Masyarakat Dalam Penggunaan Vaksin Sinovac Dengan Menerapkan Algoritma Term Frequence – Inverse Document Frequence (TF-IDF) dan Metode Deskripsi," *Journal of Information System Research (JOSH)*, vol. 3, no. 3, pp. 206–217, Apr. 2022, doi: 10.47065/JOSH.V3I3.1441.

[29] K. Pamungkas, K. T. Pamungkas, L. Aridinanti, and W. Wibowo, "Analisis Sentimen Pelaporan Masyarakat di Situs Media Centre Surabaya dengan Naïve Bayes Classifier," *Jurnal Sains dan Seni ITS*, vol. 11, no. 2, pp. D197–D203, Jul. 2022, doi: 10.12962/j23373520.v11i2.72566.

[30] M. Das, S. Kamalanathan, and P. Alphonse, "A Comparative Study on TF-IDF feature Weighting Method and its Analysis using Unstructured Dataset," *CEUR Workshop Proc*, vol. 2870, pp. 98–107, Aug. 2023, doi: 10.48550/arXiv.2308.04037.

[31] M. F. D. Putra and E. B. Setiawan, "Influence of Sentiment on Mandiri Bank Stocks (BMRI) Using Feature Expansion with FastText and Logistic Regression Classification," *ICACNIS 2022 - 2022 International Conference on Advanced Creative Networks and Intelligent Systems: Blockchain Technology, Intelligent Systems, and the Applications for Human Life, Proceeding*, 2022, doi: 10.1109/ICACNIS57039.2022.10055450.

[32] M. Liebenlito, A. A. Yesinta, and M. I. S. Musti, "Deteksi Clickbait pada Judul Berita Online Berbahasa Indonesia Menggunakan FastText," *Journal of Applied Computer Science and Technology*, vol. 5, no. 1, pp. 56–62, Mar. 2024, doi: 10.52158/jacost.v5i1.655.

[33] M. Rahman, M. D. Rahman, A. Djunaidy, and F. Mahananto, "Penerapan Weighted Word Embedding pada Pengklasifikasian Teks Berbasis Recurrent Neural Network untuk Layanan Pengaduan Perusahaan Transportasi," *Jurnal Sains dan Seni ITS*, vol. 10, no. 1, pp. A1–A6, Aug. 2021, doi: 10.12962/j23373520.v10i1.56145.

[34] B. G. Marcot and A. M. Hanea, "What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis?," *Comput Stat*, vol. 36, no. 3, pp. 2009–2031, Sep. 2021, doi: 10.1007/S00180-020-00999-9/METRICS.

[35] M. Jaiswal, H. Tiwari, and S. Kumar, "Diabetes Prediction using Optimization Techniques with Machine Learning Algorithms," *Int J Electron Healthc*, vol. 13, no. 1, p. 1, 2023, doi: 10.1504/ijeh.2023.10054229.

[36] D. I. Puteri, "Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah," *Euler : Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 11, no. 1, pp. 35–43, May 2023, doi: 10.34312/euler.v11i1.19791.

[37] R. Annas Sholehat, E. B. Setiawan, and Y. Sibaroni, "Aspect-Based Sentiment Analysis on Twitter Using Bidirectional Long Short-Term Memory," *JURNAL MEDIA INFORMATIKA BUDIDARMA* , vol. 7, pp. 999–999, 2023, doi: 10.30865/mib.v5i1.2293.

[38] E. Erlin, Y. Desnelita, N. Nasution, L. Suryati, and F. Zoromi, "Dampak SMOTE terhadap Kinerja Random Forest Classifier berdasarkan Data Tidak seimbang," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 3, pp. 677–690, Jul. 2022, doi: 10.30812/MATRIK.V21I3.1726.

[39] T. Wongvorachan, S. He, and O. Bulut, "A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining," *Information (Switzerland)*, vol. 14, no. 1, Jan. 2023, doi: 10.3390/info14010054.

[40] E. Sutoyo and M. A. Fadlurrahman, "Penerapan SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Television Advertisement Performance Rating Menggunakan Artificial Neural Network," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 6, no. 3, p. 379, Dec. 2020, doi: 10.26418/JP.V6I3.42896.

[41] D. Elreedy, A. F. Atiya, and F. Kamalov, "A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Mach Learn*, vol. 113, no. 7, pp. 4903–4923, Jul. 2024, doi: 10.1007/s10994-022-06296-4.

[42] H. Akbar and W. K. Sanjaya, "Kajian Performa Metode Class Weight Random Forest pada Klasifikasi Imbalance Data Kelas Curah Hujan," *Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi*, vol. 3, no. 1, Dec. 2023, doi: 10.20885/SNATI.V3I1.30.