

Geospatial Sentiment Analysis Using Twitter Data on Natural Disasters in Indonesia with *Support Vector Machine*

Muhamad Agung Nulhakim ^{1*}, Yuliant Sibaroni ², Ku Muhammad Naim Ku Khalif ³

^{1,2}*School Of Computing, Telkom University
Jl. Telekomunikasi No.1 Terusan Buah Batu, Bandung, Jawa Barat, Indonesia, 40257*

³*Centre of Excellence for Artificial Intelligence & Data Science, Universiti Malaysia Pahang
Kuantan 26300, Pahang, Malaysia*

*agunghakim@student.telkomuniversity.ac.id

Abstract

Twitter serves as a crucial platform for expressing public sentiment during natural disasters, yet understanding and addressing these sentiments remains challenging due to data volume, imbalance, and regional disparities in response. This study aims to bridge this gap by conducting geospatial sentiment analysis on 988 labeled tweets related to the eruption of Mount Marapi, categorized into four aspects which are Basic Needs, Impact and Damage, Response and Action, and Weather and Nature. The preprocessing stage includes data cleaning, case folding, tokenization, normalization, stopword removal, and stemming. Feature extraction uses TF-IDF, while class imbalance is addressed with SMOTE. Each aspect is modeled separately using Support Vector Machine (SVM) with linear, polynomial, and RBF kernels, evaluated through 10-fold cross-validation. Results show that the linear kernel performed best across most aspects, achieving 92.42% accuracy for Impact and Damage, 80.38% for Response and Action, and 94.22% for Weather and Nature. Meanwhile, the RBF kernel showed competitive performance with 89.54% accuracy for Basic Needs. Geospatial visualization highlights regional sentiment distribution patterns, offering insights into public responses across Indonesian regions. This study contributes to improving disaster response strategies by providing insights into public sentiment, enabling authorities to better allocate resources and address community concerns effectively.

Keywords: Sentiment Analysis, Geospatial Analysis, Natural Disasters, SVM, Twitter.

I. INTRODUCTION

Natural disasters such as volcanic eruptions, earthquakes, and floods are recurring challenges in Indonesia, a country situated on the Pacific Ring of Fire [1]. These disasters cause not only physical damage but also significant emotional and psychological distress to affected communities. During such crises, social media platforms, particularly Twitter, serve as critical tools for individuals to express opinions, share experiences, and disseminate information in real time [2]. Public responses on Twitter range from expressions of distress and dissatisfaction to gratitude and resilience. However, the sheer volume and unstructured nature of this data pose significant challenges for effective analysis. Extracting meaningful insights from this data is essential to understanding public sentiment and informing disaster response strategies [3]. Despite the potential of this data,

sentiment patterns across different regions remain underexplored, especially in terms of how sentiments vary between directly affected regions and those observing from afar.

Sentiment analysis is a field of study which strives to classify human opinions and sentiments into positive and negative classes [4][5]. Textual data is a rich source of information. This is why textual data analysis has gained more and more importance in recent years, given the powerful mainstream algorithms expected to extract meaningful insights from the data. Various approaches hold promising utility in sentiment analysis, with machine learning methods, particularly Support Vector Machine (SVM), emerging as a popular choice owing to the significant advantage of their ability to classify correctly in high-dimensional data [6][7][8]. In addition, SVM is effective for both linear and non-linear classification tasks and excels at handling small to moderately sized datasets efficiently [9].

Previous studies have shown that SVM outperforms other machine learning algorithms in text classification tasks, particularly when paired with feature extraction techniques like TF-IDF [10]. Additionally, SVM allows the flexibility of using different kernel functions, including linear, polynomial, and Radial Basis Function (RBF), to capture diverse data relationships. However, selecting the most suitable kernel remains a critical challenge in achieving optimal performance. While existing studies often focus solely on sentiment classification performance, the integration of geospatial analysis with sentiment analysis remains limited, particularly in disaster-related contexts. Geospatial analysis provides a powerful way to visualize and interpret sentiment distribution across specific locations [11]. By combining sentiment analysis with geospatial mapping, it becomes possible to identify how sentiments vary regionally, particularly between directly affected regions and those observing from a distance [12]. This spatial understanding is crucial for government agencies, disaster management authorities, and humanitarian organizations in prioritizing resource allocation, directing aid efforts, and addressing public concerns more effectively. By mapping sentiment distributions, geospatial analysis helps uncover localized reactions, such as differing levels of concern, support, or opposition across regions [13][14]. Despite its potential, there is a noticeable gap in research that combines SVM-based sentiment analysis with geospatial visualization to comprehensively analyze disaster-related sentiments.

This study aims to address these gaps by analyzing Twitter data related to the eruption of Mount Marapi through sentiment classification and geospatial visualization. The objectives of this research are threefold. First, to evaluate the performance of SVM kernels (linear, polynomial, and RBF) across different sentiment aspects. Second, to identify the most effective kernel configuration for each aspect using evaluation metrics such as accuracy, precision, recall, and F1-score. And lastly, to map sentiment distributions geospatially, offering insights into how sentiments vary across affected and unaffected regions. Data preprocessing includes data cleaning, case folding, tokenization, normalization, stopword removal, and stemming to make a high-quality input to the model [15]. Also, TF-IDF is used for feature extraction and class balance is handled with the help of SMOTE to improve model reliability [16]. All SVM models are evaluated using 10-fold cross-validation and compared based on metrics such as accuracy, precision, recall, and F1 score. SVM is particularly suitable for this type of analysis due to its effectiveness in distinguishing between sentiment classes, even in high-dimensional spaces, ensuring robustness and reliability in the results.

The main contribution of this research lies in integrating sentiment analysis using SVM with geospatial analysis to provide a comprehensive understanding of public sentiment during natural disasters. By identifying regional sentiment patterns, this study offers actionable insights for policymakers, disaster response agencies, and local governments to improve their communication strategies, optimize resource distribution, and enhance public trust. Furthermore, this study serves as a foundation for future research exploring the integration of advanced machine learning models with geospatial techniques for sentiment analysis in disaster contexts. This paper is organized into several sections, with Section 2 discussing related works, Section 3 describing the proposed methods, Section 4 presenting results and analysis, and Section 5 concluding the findings.

II. LITERATURE REVIEW

Sentiment analysis is widely used to classify opinions expressed in textual data into categories such as positive, negative, or neutral sentiments. Twitter's data has been extensively analyzed due to its accessibility

and real-time nature. A study by Ainun Zumarniansyah et al compared the performance of Support Vector Machine (SVM) and Naïve Bayes methods for sentiment classification [17]. The results demonstrated that SVM achieved higher accuracy in handling high-dimensional textual data, making it a suitable method for complex datasets like tweets. Similarly, research by Kharisma Wiati Gusti explored the use of SVM and Logistic Regression for classifying disaster-related tweets [18]. The study concluded that SVM consistently outperformed Logistic Regression in accuracy, with SVM achieving 80.41% accuracy compared to Logistic Regression’s 70.74% with SMOTE, showcasing SVM’s robustness in handling complex and imbalanced datasets [18]. Another notable study by Mera Kartika Delimayanti applied a multiclass SVM approach to classify flood-related tweets into eyewitness, non-eyewitness, and unknown categories [19]. The research employed various preprocessing steps, feature extraction techniques like TF-IDF, and word weighting methods. The findings revealed that the RBF kernel produced the best classification performance with an accuracy of 87.03%, highlighting the potential of SVM in disaster-related sentiment analysis. However, these studies focused on textual sentiment classification without incorporating geospatial dimensions.

Geospatial analysis involves examining the spatial distribution of data to uncover patterns and trends associated with specific locations. A study by Tao Hu et al utilized the Local Indicators of Spatial Association (LISA) method to identify spatial clusters of sentiments [20]. It demonstrated the value of combining geospatial and temporal dimensions to provide insights into the relationship between user sentiment and geographical regions. In another study, Olga Buchel and Diane Rasmussen Pennington emphasized the importance of spatial and temporal contexts in understanding communication patterns and human behavior [11]. Despite their contributions, these studies focused primarily on general geospatial trends and lacked a detailed exploration of sentiment analysis combined with geospatial insights specific to disaster contexts. Furthermore, research by Tao Hu et al applied spatial clustering techniques to analyze sentiment variations in urban settings, uncovering associations between sentiments and specific points of interest [20]. These approaches demonstrated the potential of combining sentiment and geospatial analysis to provide actionable insights but often overlooked the optimization of sentiment classification models.

TABLE I
 SUMMARY OF RELATED WORKS IN SENTIMENT AND GEOSPATIAL ANALYSIS

Author(s)	Methodology	Key Findings	Limitations	Research Gap
Ainun Zumarniansyah et al	SVM vs Naïve Bayes	SVM achieved higher accuracy in text classification	No integration of geospatial analysis	Lack of spatial sentiment insights
Kharisma Wiati Gusti	SVM vs Logistic Regression with SMOTE	SVM outperformed Logistic Regression with 80.41% accuracy	No geospatial sentiment mapping	Missing spatial sentiment visualization
Mera Kartika Delimayanti et al	Multiclass SVM with TF-IDF and RBF Kernel	RBF kernel showed 87.03% accuracy in disaster sentiment classification	No geospatial integration	Absence of spatial analysis insights
Tao Hu et al	LISA Method (Spatial Clustering)	Identified spatial clusters of sentiments	No integration with SVM classification models	Limited sentiment model optimization
Olga Buchel et al	Spatial and Temporal Analysis	Contextual insights on communication patterns	No focus on disaster-related sentiment analysis	Limited focus on regional variations

The findings from the key studies discussed above are summarized in Table I, highlighting the methods, primary results, and identified gaps. While previous studies have significantly advanced sentiment analysis and

geospatial analysis, several gaps remain. First, there is limited research systematically evaluating the performance of different SVM kernels for sentiment classification in disaster-related datasets. Second, most studies have not fully integrated geospatial analysis to enhance sentiment analysis results, particularly in the context of Indonesian natural disasters. Third, challenges such as data noise, imbalanced datasets, and the absence of robust preprocessing pipelines often hinder the reliability of results. This study aims to address these gaps by combining sentiment analysis and geospatial analysis to provide a comprehensive understanding of public sentiment during natural disasters in Indonesia. By systematically evaluating various SVM kernels, including linear, polynomial, and RBF, and incorporating geospatial visualization techniques using Nominatim and Folium libraries, this research contributes to the growing body of literature on sentiment analysis and geospatial insights.

Furthermore, this study differentiates itself from previous research through a structured comparison of SVM kernel configurations to identify the most effective model for sentiment classification across multiple aspects. Unlike earlier works that focused primarily on textual analysis, this research integrates geospatial mapping tools to visualize sentiment distribution across different regions. This integration allows for a more thorough understanding of sentiment patterns and their spatial context, offering actionable insights that can aid in disaster response planning and policy-making. Through this combined sentiment-geospatial analysis framework, the study bridges the gap between sentiment classification and spatial analysis, delivering a nuanced perspective on public sentiment during natural disasters.

III. RESEARCH METHOD

A. System Design

The system implemented in this research is designed to perform sentiment analysis with geospatial insights using Support Vector Machine (SVM). Fig. 1 illustrates the system overview, which consists of several interconnected stages. The process begins with data crawling, where tweets related to natural disasters, such as the eruption of Mount Marapi, are collected from Twitter. These tweets form the primary dataset for analysis. Following this, a manual labeling stage categorizes the data into positive and negative sentiments to ensure the dataset's reliability and suitability for sentiment analysis.

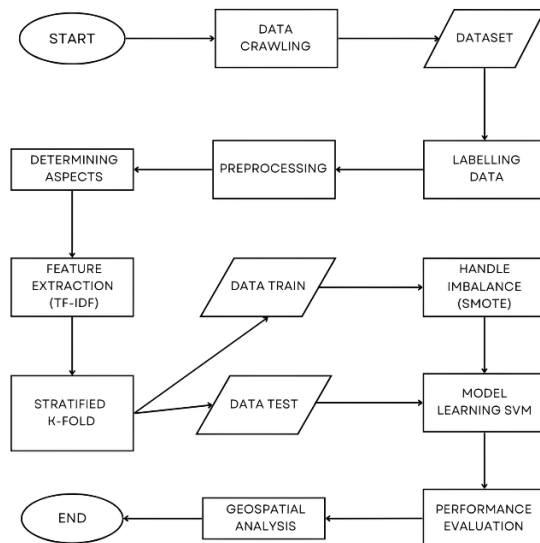


Fig. 1. System Overview

Preprocessing is then performed to clean and prepare the data. This stage includes essential tasks such as data cleaning, case folding, tokenization, normalization, stopword removal, and stemming. These steps enhance the

quality of the textual data, making it ready for feature extraction. Additionally, determining aspects based on predefined keywords is carried out both before TF-IDF feature extraction and after preprocessing, ensuring accurate categorization of tweets into specific aspects (Basic Needs, Impact and Damage, Response and Action, Weather and Nature). The numerical representation of the text is achieved through Term Frequency-Inverse Document Frequency (TF-IDF), which measures the significance of words in the dataset. Additionally, the Synthetic Minority Over-sampling Technique (SMOTE) is applied to address imbalanced data by generating synthetic samples for minority classes, improving the model's robustness in handling skewed datasets.

After preprocessing and feature extraction, the dataset is split into train and test sets using Stratified K-Fold cross-validation with 10 folds to ensure balanced representation of sentiment classes across multiple iterations. The SVM model is then trained on the processed data, leveraging three kernel configurations, which are linear, polynomial, and RBF, each offering unique capabilities in handling linear and non-linear data patterns. Hyperparameters such as C equal to one and gamma set to scale ensure optimal performance across varying data structures. The pipeline architecture integrates preprocessing, SMOTE balancing, and SVM modeling, streamlining the entire sentiment classification workflow. Compared to previous studies, this method provides a more comprehensive perspective by combining advanced text classification with geospatial analysis, enabling stakeholders, including governments and disaster response teams, to identify priority regions requiring immediate attention.

B. Data Crawling

Finding data from a specific source is referred to as "crawling." In this study, 988 tweets were collected from the Twitter social media platform using the Twitter API and Python programming language. The keywords used for the crawling process were 'marapi meletus', 'meletus marapi', 'erupsi marapi', and 'marapi erupsi', focusing on tweets related to the eruption of Mount Marapi. The retrieved data includes important features such as tweet text and location information, which are essential for sentiment and geospatial analysis. The data was automatically stored in a CSV file format for further analysis. This dataset serves as the foundation for sentiment and geospatial analysis in this study.

TABLE II
 COUNT OF SENTIMENT LABELS

Label	Amount	Ratio (%)
Positive	220	22.26
Negative	768	77.74
Total Data	22	22

TABLE III
 EXAMPLE OF LABELING

Data	Label
“Sebanyak 38 mahasiswa Politeknik Negeri Padang (PNP) mendapatkan penghargaan khusus. Penghargaan itu diberikan berkat jasa mereka yang ikut terjun di dalam proses evakuasi korban erupsi Marapi 3 Desember 2023 silam”	Positive
“Salah satu korban erupsi Gunung Marapi Zhafirah Zahrim Febrina atau yang dikenal Ife kini dinyatakan meninggal dunia setelah berjuang selama 13 hari dirawat di Rumah Sakit Simak Selengkapnya Disini”	Negative

C. Labeling Data

After completing the data collection process and obtaining the dataset, the next step in this research is data labeling, which was carried out manually to ensure accuracy and relevance. Each tweet was assigned a sentiment label, categorized into positive or negative sentiments, based on its content related to the eruption of Mount Marapi. To enhance validity, the dataset was thoroughly reviewed multiple times, with ambiguous tweets being carefully discussed to ensure consistency and correctness. The overall data and sample data are presented in Table II and Table III.

D. Preprocessing Data

Preprocessing is one of the most critical steps in preparing data for analysis or modeling, ensuring that raw data is cleaned, simplified, and optimized for better performance in sentiment analysis [21]. In this research, preprocessing plays a significant role in improving the quality and accuracy of the sentiment and geospatial analysis conducted on tweets related to the eruption of Mount Marapi. The preprocessing stages include Data Cleansing, Case Folding, Tokenizing, Normalization, Stopword Removal, and Stemming to reduce words to their base forms. These steps collectively refine the dataset, addressing the challenges posed by social media data and ensuring its readiness for feature extraction and model training.

1) *Data Cleansing*: This step removes unnecessary elements such as URLs, numbers, punctuation, symbols, hashtags, and emoticons from the dataset. Simplifying the text in this way helps eliminate noise and retain only relevant content for analysis.

2) *Case Folding*: Text is converted to lowercase to ensure consistency. By treating words with different cases as the same word, case folding reduces complexity and enhances the uniformity of the dataset.

3) *Tokenizing*: This step breaks down the text into smaller units or tokens, typically individual words. Tokenization simplifies the text, making subsequent processing steps more efficient.

4) *Normalization*: Normalization in this study was performed manually by standardizing slang words and informal language into their formal equivalents. Each word or phrase was carefully reviewed and mapped to its proper form. For example, *gk* was replaced with *nggak*, *sy* with *saya*, and *trs* with *terus*. This manual normalization process ensures that context-specific slang words commonly found in social media data are accurately interpreted and standardized.

5) *Stopword Removal*: Stopwords, such as *dan*, *di*, *ke*, which do not contribute significant meaning to the sentiment analysis, are removed from the text. In this study, NLTK's stopwords library was used to filter out unnecessary words from the dataset, ensuring that only meaningful tokens remain for further analysis.

6) *Stemming*: This step reduces words to their root forms to eliminate morphological variations. For example, words like *berlari* and *lari* are reduced to a single root word, which is *lari*. In this research, Sastrawi, a widely-used stemming library for the Indonesian language, was employed to perform this step efficiently. perform this step efficiently.

E. Determining Aspects

After the preprocessing stage, the next step in this research is determining the aspects of the tweets to categorize them into meaningful groups. The categorization is based on predefined keywords associated with four primary aspects which are *Basic Needs*, *Response and Action*, *Impact and Damage*, and *Weather and Nature*. Each aspect is defined by a set of relevant keywords extracted from the textual data. For example, the *Basic Needs* aspect includes keywords like *makan*, *air*, and *obat*, while the *Response and Action* aspect contains keywords such as *bantuan*, *evakuasi*, and *penyelamatan*. Similarly, the *Impact and Damage* aspect focuses on terms like *kerugian*, *korban*, and *luka*, and the *Weather and Nature* aspect captures words such as *hujan*, *erupsi*, and *longsor*. During this step, each tweet is analyzed for the presence of these keywords to assign it to the most

appropriate aspect category. This classification ensures that sentiment analysis and subsequent evaluations are conducted with a clear focus on specific thematic dimensions, thereby enhancing the granularity and relevance of the results.

F. Feature Extraction TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) is an algorithm used to calculate the weight of each term within a document, assessing the significance of the term relative to the entire corpus. This method combines two metrics: Term Frequency (TF) and Inverse Document Frequency (IDF). Term Frequency (TF) measures how often a term appears in a specific document and is calculated using the formula 1.

$$TF = \frac{\text{The number of words in the document}}{\text{Total number of words in the document}} \tag{1}$$

This calculation highlights the importance of a term within a single document, assigning a higher weight to frequently occurring words. Inverse Document Frequency (IDF), on the other hand, evaluates the rarity of a term across the entire corpus of documents. It is defined as formula 2.

$$IDF = \log \frac{\text{Total number of documents}}{\text{Number of documents containing words}} \tag{2}$$

This component reduces the weight of commonly occurring terms, ensuring that terms frequently present in many documents (e.g., stopwords) do not dominate the analysis. The TF-IDF score for a term is calculated by multiplying the TF and IDF values as in formula 3.

$$TF - IDF = TF \times IDF \tag{3}$$

G. Stratified K-Fold Cross Validation

Stratified K-Fold Cross Validation is a statistical technique used to evaluate the performance of machine learning models by splitting the dataset into subsets, or folds, while preserving the proportion of each class in the data [22]. This method ensures that each fold maintains the same distribution of class labels as the original dataset, making it particularly effective for imbalanced datasets. The process begins by dividing the data into folds. For each iteration, one fold is used as the test set, while the remaining folds serve as the training set. This process is repeated times, with each fold being used as the test set exactly once. The final performance metrics, such as accuracy, precision, recall, and F1-score, are computed as the average of the results from all iterations.



Fig. 2. Schematic Diagram of Stratified K-fold Cross Validation

Stratified K-Fold Cross Validation offers several advantages over traditional K-Fold Cross Validation. By preserving the class distribution in each fold, it ensures that the model is tested on data that represents the true class proportions, which is critical for tasks involving imbalanced datasets. This method also reduces the risk of bias in performance evaluation, providing a more reliable estimate of the model’s generalization ability. The schematic diagram of Stratified K-fold Cross Validation is shown in Fig. 2.

H. Handle Imbalance (SMOTE)

Synthetic Minority Over-sampling Technique (SMOTE) is a method used to address class imbalance by creating new synthetic samples for the minority class, thereby balancing the dataset [23]. This technique generates synthetic data points by interpolating between existing samples of the minority class. The newly synthesized examples are then added to the training data, ensuring that the classifier is trained on a balanced dataset. Class imbalance can significantly affect the performance of machine learning models, as models tend to favor the majority class when the dataset is skewed [24]. By incorporating SMOTE, the training dataset becomes more representative, enabling the classifier to learn from a more equitable distribution of classes. This approach improves the model's ability to generalize and correctly classify instances of the minority class. An illustration of how SMOTE works can be seen in Fig. 3.

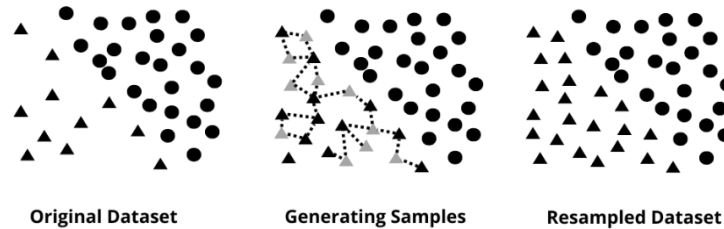


Fig. 3. SMOTE Illustration

I. Model Learning SVM

Support Vector Machine (SVM) is an algorithm used for data classification, capable of handling both linear and non-linear patterns. The strength of SVM in managing non-linear patterns lies in its ability to utilize kernel functions to project data into a higher-dimensional space, enabling better separation of classes [25]. SVM effectively groups different types of objects into distinct categories. As a supervised learning model, SVM uses pre-labeled training data to predict the classes of new, unseen samples. SVM is fundamentally a binary classification algorithm, meaning it is designed to differentiate between two classes. However, for tasks requiring the classification of more than two classes, SVM can be extended into a multi-class classifier using strategies such as One-vs-Rest or One-vs-One. These strategies allow SVM to handle multi-class problems effectively by decomposing them into multiple binary classification tasks.

SVM kernels play a pivotal role in defining the decision boundary between classes. Each kernel function provides a unique way to calculate the similarity between data points:

1) *Linear Kernel:* The linear kernel calculates the dot product of two vectors to create a hyperplane for linearly separable data. The formula is expressed as formula 4. where x and y represent input vectors.

$$K(x, y) = (x \cdot y) \tag{4}$$

2) *Polynomial Kernel:* The polynomial kernel extends the concept of the linear kernel by raising the dot product to a polynomial degree d, enabling the separation of more complex relationships. The formula is expressed as formula 5. where c is a constant and d denotes the degree of the polynomial.

$$K(x, y) = (x \cdot y + c)^d \tag{5}$$

3) *Radial Basis Function (RBF) Kernel*: The RBF kernel, also known as the Gaussian kernel, calculates the similarity between two points using a Gaussian function. This kernel is suitable for high-dimensional and irregular data. The formula is expressed as formula 6. where γ controls the influence of individual data points.

$$K(x, y) = \exp(-\gamma(x - y)^2) \tag{6}$$

In this study, SVM was applied across four predefined aspects which are Basic Needs, Impact and Damage, Response and Action, and Weather and Nature. For each aspect, the dataset was processed using TF-IDF for feature extraction, and SMOTE was employed to address class imbalance by generating synthetic samples for minority classes. The model was validated using Stratified K-Fold Cross-Validation with 10 folds, ensuring balanced representation of sentiment classes across training and testing datasets. Each kernel (linear, polynomial, and RBF) was tested independently on every aspect to evaluate their effectiveness. The C parameter was set to *one* to control the trade-off between achieving a low error margin and minimizing misclassification, while the gamma parameter was set to *scale*, allowing the model to automatically adjust gamma based on the number of features. Additionally, a pipeline approach was implemented to streamline the modeling process by combining preprocessing, SMOTE balancing, and SVM training in a single workflow. The illustration of the SVM method is shown in Fig. 4.

In this study, SVM was applied across four predefined aspects which are Basic Needs, Impact and Damage, Response and Action, and Weather and Nature. For each aspect, the dataset was processed using TF-IDF for feature extraction, and SMOTE was employed to address class imbalance by generating synthetic samples for minority classes.

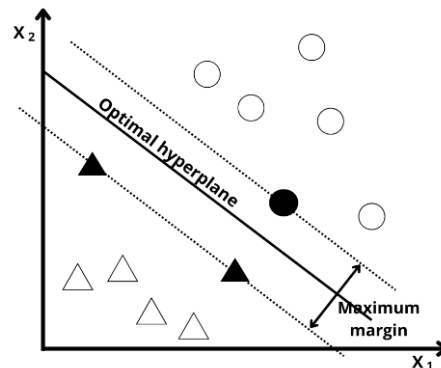


Fig. 4. Illustration of the SVM Method

J. Performance Evaluation

Performance evaluation is a critical step in determining the effectiveness of the developed classification model. In this study, evaluation metrics such as accuracy, precision, recall, and F1-score were used to assess the model’s performance. These metrics were calculated using a confusion matrix, which provides a detailed comparison between the actual and predicted values of the dataset.

TABLE IV
 CONFUSION MATRIX

Confusion Matrix	Positive	Negative
Positive	True Positive	False Negative
Negative	False Positive	False Negative

The confusion matrix encompasses four key terms that aid in understanding model predictions [26]. True Positive (TP), which refers to correctly predicted positive instances. True Negative (TN), indicating correctly

predicted negative instances. False Positive (FP), representing incorrect predictions where the true label is negative but predicted as positive. False Negative (FN), which describes incorrect predictions where the true label is positive but predicted as negative. The following formulas were used to calculate the performance metrics:

1) *Accuracy*: The ratio of correctly predicted instances (both positive and negative) to the total number of instances. The following formula can be used to calculate accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

2) *Precision*: The ratio of correctly predicted positive instances to all instances predicted as positive. The following formula can be used to calculate precision:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

3) *Recall*: The ratio of correctly predicted positive instances to all actual positive instances. The following formula can be used to calculate recall:

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

4) *F1-Score*: A harmonic mean of precision and recall, providing a balanced measure of both metrics. The following formula can be used to calculate F1-Score:

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \quad (10)$$

These metrics ensure a comprehensive evaluation of the model, particularly in datasets with imbalanced classes. By analyzing the confusion matrix and these performance indicators, the robustness and reliability of the classification model can be effectively measured.

K. Geospatial Analysis

Geospatial analysis involves the visualization and interpretation of data based on geographic or spatial locations [27]. In this study, geospatial analysis is applied to understand the distribution of sentiments related to the eruption of Mount Marapi across different regions in Indonesia. By combining sentiment analysis with geographic information, this method enables the identification of patterns, trends, and regional disparities in public sentiment. The analysis focuses on four predefined aspects which are basic needs, response and action, impact and damage, and weather and nature, offering a multi-dimensional view of how sentiment varies geographically.

The geospatial analysis process begins with extracting location data during the initial data crawling stage, where tweet metadata such as user-provided locations are collected. These raw location data are then processed using the Nominatim library to convert textual location names into geographical coordinates (latitude and longitude) via OpenStreetMap's geocoding service. Invalid or unrecognized locations are handled systematically by assigning default coordinate values to maintain data consistency. Once the coordinates are obtained, data aggregation is performed to calculate sentiment dominance at each location for every aspect. Using Folium, an interactive geospatial map is generated for each aspect. Each map includes markers that represent locations with dominant sentiment classifications (positive or negative) color-coded for clarity. Specifically, blue markers represent positive sentiment, while red markers represent negative sentiment. The

size and opacity of the markers are adjusted to improve readability and emphasize regions with significant sentiment trends.

IV. RESULTS AND DISCUSSION

In this research, various steps and methods were systematically applied to evaluate the performance of the sentiment analysis model using Support Vector Machine (SVM) across four predefined aspects which are *Basic Needs*, *Impact and Damage*, *Response and Action*, and *Weather and Nature*. Each aspect was analyzed independently using SVM with three kernel configurations—linear, polynomial, and RBF—to identify the optimal kernel for sentiment classification in each specific context. The evaluation process involved several critical stages, including preprocessing with techniques such as data cleaning, case folding, tokenization, normalization, stopword removal, and stemming. Feature extraction was performed using Term Frequency-Inverse Document Frequency (TF-IDF), while class imbalance was addressed using Synthetic Minority Over-sampling Technique (SMOTE). Model validation was carried out using Stratified K-Fold Cross Validation with 10 folds to ensure robust and reliable evaluation results. Additionally, geospatial analysis was conducted after performance evaluation to visualize the sentiment distribution across Indonesian regions for each aspect, offering insights into regional sentiment patterns and public responses.

A. Evaluation Results

This section presents the evaluation results of the SVM model for each aspect which are *Basic Needs*, *Impact and Damage*, *Response and Action*, and *Weather and Nature*. Each aspect was processed separately using the same methodology, including data preprocessing, feature extraction with TF-IDF, and class balancing with SMOTE. The evaluation was performed using three SVM kernels—linear, polynomial, and RBF—with Stratified K-Fold Cross Validation set to 10 folds to ensure a robust and reliable performance evaluation. The model was trained with C equal to one and gamma set to scale hyperparameters. The performance of each kernel was assessed using four key metrics which are accuracy, precision, recall, and F1-score. In the *Basic Needs* aspect, the dataset was preprocessed using cleaning, tokenization, normalization, stopword removal, and stemming. TF-IDF was applied to extract relevant features, and SMOTE was used to address class imbalance. Stratified K-Fold Cross Validation with 10 folds ensured an even distribution of data across training and testing sets. The SVM model was tested with three kernels which are linear, polynomial, and RBF. The linear kernel performed the best, achieving the highest scores across all metrics. The evaluation results for the *Basic Needs* aspect are presented in Table V.

TABLE V
 METRIC EVALUATION FOR BASIC NEEDS ASPECT

Metric	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Linear	89.11	90.77	89.11	87.97
Polynomial	79.22	73.72	79.22	73.88
RBF	81.44	78.32	81.44	77.37

TABLE VI
 METRIC EVALUATION FOR IMPACT AND DAMAGE ASPECT

Metric	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Linear	92.42	92.38	92.42	92.09
Polynomial	87.97	88.94	87.97	85.80
RBF	89.54	89.98	89.54	88.12

For the *Impact and Damage* aspect, similar preprocessing and feature extraction steps were performed, followed by class balancing using SMOTE. Data was split into training and testing sets using 10-fold Stratified K-Fold Cross Validation. The evaluation results revealed that the linear kernel achieved the best performance

across all metrics, followed closely by the RBF kernel, while the polynomial kernel performed the lowest. The evaluation results for the *Impact and Damage* aspect are presented in Table VI.

In the *Response and Action* aspect, the dataset was preprocessed, balanced with SMOTE, and split into 10 folds using Stratified K-Fold Cross Validation. The linear kernel demonstrated the best performance, while the polynomial and RBF kernels performed relatively lower, with the RBF kernel showing the weakest performance across all metrics. The evaluation results for the *Response and Action* aspect are shown in Table VII.

TABLE VII
METRIC EVALUATION FOR RESPONSE AND ACTION ASPECT

Metric	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Linear	80.38	79.31	80.38	78.15
Polynomial	75.11	70.31	75.11	67.92
RBF	71.43	59.11	71.43	61.62

For the *Weather and Nature* aspect, data preprocessing and balancing steps were performed similarly, followed by 10-fold Stratified K-Fold Cross Validation for robust evaluation. The linear kernel achieved the best performance, with the polynomial and RBF kernels following closely behind. However, all three kernels demonstrated strong performance across the metrics. The evaluation results for the *Weather and Nature* aspect are presented in Table VIII.

TABLE VIII
METRIC EVALUATION FOR WEATHER AND NATURE ASPECT

Metric	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Linear	94.22	92.40	94.22	92.63
Polynomial	93.70	89.61	93.70	91.26
RBF	93.18	87.59	93.18	90.24

B. Geospatial Results

The geospatial analysis conducted in this study aims to visualize the distribution of sentiments across various regions based on four predefined aspects which are *Basic Needs*, *Impact and Damage*, *Response and Action*, and *Weather and Nature*. These aspects were determined based on predefined keywords that categorize the tweets into specific themes. For example, keywords such as *makanan*, *air*, and *obat* were used for the *Basic Needs* aspect, while words like *kerugian* and *korban* were used for the *Impact and Damage* aspect. Similarly, words like *bantuan* and *evakuasi* defined the *Response and Action* aspect, and terms like *hujan*, *erupsi*, and *longsor* were associated with the *Weather and Nature* aspect. Sentiment classification is divided into positive and negative categories. The data consists of 26 positive and 355 negative sentiments for the *Weather and Nature* aspect, 73 positive and 309 negative sentiments for the *Impact and Damage* aspect, 28 positive and 64 negative sentiments for the *Basic Needs* aspect, and 93 positive and 40 negative sentiments for the *Response and Action* aspect. These datasets serve as the foundation for geospatial visualization and subsequent analysis to understand regional sentiment trends.

The geospatial analysis for the *Basic Needs* aspect shows a distribution of 28 positive and 64 negative sentiments across multiple regions. Dominant negative sentiments are observed in regions such as Aceh, Bandung, Batusangkar, Depok, and Jakarta, indicating dissatisfaction with the availability of essential resources like food, water, and shelter. Conversely, positive sentiments are seen in Bogor, Buleleng, Gorontalo, and Jayapura, reflecting satisfaction with the accessibility of basic necessities. These results highlight regional disparities in the fulfillment of basic needs during disaster response efforts. The geospatial visualization for the *Basic Needs* aspect can be seen in Fig. 5.

The results for the *Impact and Damage* aspect reveal 73 positive and 309 negative sentiments. Negative sentiments dominate regions such as Aceh, Agam, Bandung, Denpasar, Jakarta, and Medan, where significant destruction, casualties, and economic losses are frequently highlighted. In contrast, positive sentiments emerge

in regions like Jawa Tengah, Payakumbuh, and Riau, suggesting perceptions of effective recovery efforts and resilience in handling disaster impacts. These findings emphasize the varying severity of disaster impacts and the importance of recovery initiatives across different regions. The geospatial visualization for the *Impact and Damage* aspect can be seen in Fig. 6.

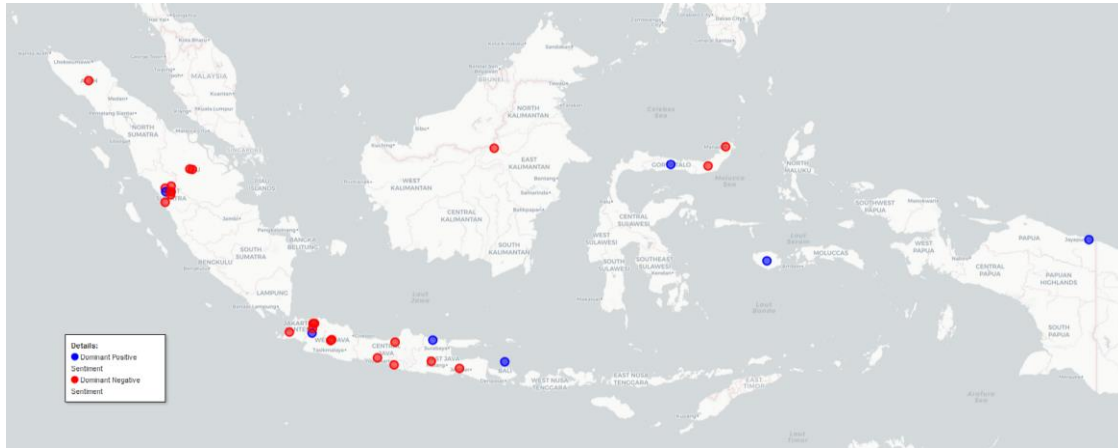


Fig. 5. Dominant Sentiments for the Basic Needs Aspect

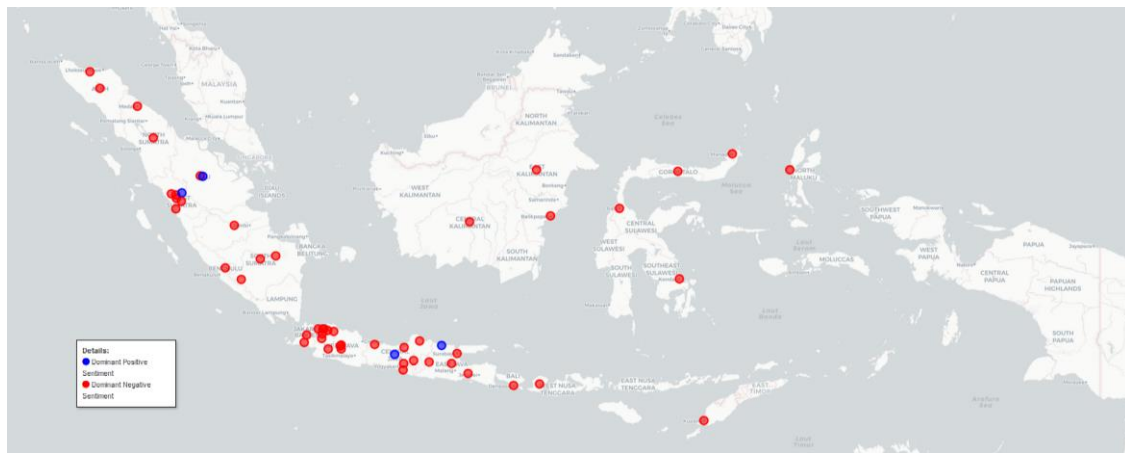


Fig. 6. Dominant Sentiments for the Impact and Damage Aspect

For the *Response and Action* aspect, the analysis identifies 93 positive and 40 negative sentiments across several regions. Positive sentiments are concentrated in areas such as Agam, Bali, Bekasi, Depok, Padang Panjang, and Yogyakarta, where respondents acknowledged effective aid distribution and swift rescue efforts. Conversely, negative sentiments are predominantly observed in Aceh, Bogor, Jambi, Medan, and Pekanbaru, reflecting dissatisfaction with the timeliness and adequacy of disaster response measures. These results highlight the variability in public perception regarding disaster response effectiveness. The geospatial visualization for the *Response and Action* aspect can be seen in Fig. 7.

The analysis of the *Weather and Nature* aspect reveals 26 positive and 355 negative sentiments. Negative sentiments are heavily concentrated in regions such as Aceh, Bandung, Jakarta Timur, Jawa Barat, Jawa Tengah, and Padang, where extreme weather conditions, landslides, volcanic eruptions, and floods are frequent concerns. Positive sentiments, on the other hand, appear in regions like Maluku, Jayapura, and Bogor, indicating perceptions of improvements in weather conditions or resilience to natural disasters. These findings underscore the ongoing challenges posed by environmental conditions during and after disasters. The geospatial visualization for the *Weather and Nature* aspect can be seen in Fig. 8.



Fig. 7. Dominant Sentiments for the Response and Action Aspect



Fig. 8. Dominant Sentiments for the Weather and Nature Aspect

C. Discussion

The evaluation results of the SVM model reveal distinct patterns in kernel performance across various aspects, demonstrating the adaptability of the linear kernel in sentiment classification tasks. For Basic Needs, Impact and Damage, Response and Action, and Weather and Nature, the linear kernel consistently delivered the best performance, showcasing its efficiency in datasets with relatively well-defined patterns. The linear kernel's robustness is evident in its ability to handle diverse datasets, capturing underlying relationships within text data while maintaining computational simplicity. The polynomial and RBF kernels, while effective in capturing non-linear patterns, showed slightly lower performance due to the nature of the datasets, which appear to favor simpler, linear relationships. These results are consistent with studies like Ainun Zumarniansyah et al. and Kharisma Wiati Gusti, which highlighted SVM's strength in high-dimensional and imbalanced datasets. The integration of SMOTE and TF-IDF, combined with Stratified K-Fold Cross Validation, ensured balanced datasets and reliable model evaluation. This framework emphasizes the adaptability of the linear kernel for diverse sentiment classification tasks, making it suitable for practical applications in disaster-related contexts.

The geospatial analysis complements the evaluation results by providing actionable insights into the regional distribution of sentiments. Regions directly affected by the eruption of Mount Marapi, such as Payakumbuh, Bukittinggi, Padang Panjang, and Batusangkar, exhibit dominant negative sentiments. These sentiments are largely driven by firsthand experiences of physical damage, limited access to basic needs, and delays in aid

distribution. However, certain directly affected regions, such as Payakumbuh and Bukittinggi, also display clusters of positive sentiment, likely attributed to successful community-driven recovery initiatives and effective local government responses. Conversely, regions not directly impacted, such as Jakarta and Bandung, show negative sentiments that are more policy-driven, focusing on criticisms of government policies and perceived inefficiencies in aid distribution. The stark contrast in sentiment patterns between directly and indirectly affected regions underscores the importance of considering geographical proximity and contextual factors in sentiment analysis.

A deeper exploration of geospatial sentiment patterns reveals a flow of sentiment that transitions from tangible concerns in disaster-hit regions to broader public concerns shaped by media narratives and perceptions in indirectly affected areas. For instance, the clustering of negative sentiments in Sumatera Barat emphasizes the concentrated impact of the disaster, while the dispersed sentiment patterns in Java and Bali highlight the influence of secondary factors like media coverage and public discourse. This analysis demonstrates the importance of integrating geospatial insights into sentiment analysis, as it bridges the gap between textual sentiment and spatial understanding, offering a more nuanced perspective on public responses to disasters.

Despite the strengths of the proposed approach, threats to validity must be acknowledged. The reliance on geo-tagged Twitter data introduces biases, as such data may not fully represent the affected populations or capture the entire sentiment landscape. Additionally, noise in the dataset, uneven geographical distribution of tweets, and the use of predefined aspects may limit the generalizability of the findings. These limitations align with observations in previous studies, such as those by Tao Hu et al., which emphasized the need for more balanced datasets and dynamic categorizations. Future research could address these limitations by incorporating larger, more diverse datasets and exploring advanced geospatial clustering techniques to enhance the robustness of sentiment analysis.

Overall, this study demonstrates the effectiveness of combining SVM-based sentiment classification with geospatial analysis to understand public sentiment during natural disasters. The linear kernel's consistent performance across all aspects highlights its robustness and efficiency, while the integration of geospatial insights provides valuable data for prioritizing resource allocation and refining disaster communication strategies. By bridging sentiment classification and geospatial mapping, this research contributes to advancing methodologies in disaster management and offers a comprehensive framework for understanding public sentiment at both the textual and spatial levels.

V. CONCLUSION

This research successfully applied sentiment analysis and geospatial visualization to understand public responses during the Mount Marapi eruption, focusing on four aspects, which are Basic Needs, Impact and Damage, Response and Action, and Weather and Nature. By employing the Support Vector Machine (SVM) algorithm with linear, polynomial, and RBF kernels, the linear kernel demonstrated consistent superiority across all aspects due to its robustness in capturing well-defined patterns. The integration of preprocessing techniques, such as TF-IDF for feature extraction and SMOTE for class balancing, ensured the reliability of results, while geospatial mapping revealed distinct sentiment patterns, highlighting the challenges and successes in disaster response across various regions. The findings show that sentiment distribution patterns differ significantly between directly and indirectly affected regions, providing actionable insights for disaster management. This study advances existing methodologies by combining SVM-based sentiment analysis with geospatial insights, filling critical gaps in understanding regional disparities in public sentiment during disasters. Key advantages of this approach include its ability to identify areas requiring prioritized intervention and its adaptability to diverse datasets. Future research should address identified limitations, such as dataset size, the reliance on geo-tagged Twitter data, and simplified visualization techniques, by exploring richer datasets and advanced geospatial clustering methods. These improvements can further enhance the applicability of sentiment and geospatial analysis in diverse contexts. Overall, this study underscores the value of integrating textual and spatial analysis for more effective disaster response and planning strategies.

DATA AND COMPUTER PROGRAM AVAILABILITY

Data and program used in this paper can be accessed in the following site [github.com: https://github.com/AgungHakim/Geospatial-Sentiment-Analysis-Using-Twitter-Data-on-Natural-Disasters-in-Indonesia-with-SVM](https://github.com/AgungHakim/Geospatial-Sentiment-Analysis-Using-Twitter-Data-on-Natural-Disasters-in-Indonesia-with-SVM).

ACKNOWLEDGMENT

The authors would like to express their gratitude to Allah SWT for granting them the opportunity to write this journal. The authors would also like to extend their heartfelt thanks to Mr. Yuliant Sibaroni for his patience and guidance throughout the writing process. The authors are deeply grateful to their families for their unwavering support during the completion of this work. The authors would also like to thank their friends for the assistance and encouragement provided during the writing of this journal. Lastly, the author would like to thank everyone who has been involved and helped in writing this journal.

REFERENCES

- [1] Wei, H. L. (2021). Natural Hazards: Volcanic Eruptions. In *Encyclopedia of Security and Emergency Management* (pp. 697-700). Cham: Springer International Publishing.
- [2] Mirbabaie, M., Bunker, D., Stieglitz, S., Marx, J., & Ehnis, C. (2020). Social media in times of crisis: Learning from Hurricane Harvey for the coronavirus disease 2019 pandemic response. *Journal of Information Technology*, 35(3), 195-213.
- [3] Duc, H. M., & Ha, N. T. T. (2024). Enhancing Disaster Preparedness, Emergency Response and Resource Allocation using Big Data Integration and Predictive Analytic. *Journal of Big-Data Analytics and Cloud Computing*, 9(1), 1-13.
- [4] Liu, B. (2022). *Sentiment analysis and opinion mining*. Springer Nature.
- [5] Dervenis, C., Kanakis, G., & Fitsilis, P. (2024). Sentiment analysis of student feedback: A comparative study employing lexicon and machine learning techniques. *Studies in Educational Evaluation*, 83, 101406.
- [6] Pisner, D. A., & Schnyer, D. M. (2020). Support vector machine. In *Machine learning* (pp. 101-121). Academic Press.
- [7] Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. *Information*, 15(4), 235.
- [8] Jahan, I., Islam, M. N., Hasan, M. M., & Siddiky, M. R. (2024). Comparative analysis of machine learning algorithms for sentiment classification in social media text. *World J. Adv. Res. Rev*, 23(3), 2842-2852.
- [9] Abdullah, D. M., & Abdulazeez, A. M. (2021). Machine learning applications based on SVM classification a review. *Qubahan Academic Journal*, 1(2), 81-90.
- [10] Nafis, N. S. M., & Awang, S. (2021). An enhanced hybrid feature selection technique using term frequency-inverse document frequency and support vector machine-recursive feature elimination for sentiment classification. *Ieee Access*, 9, 52177-52192.
- [11] Buchel, O., & Pennington, D. R. (2022). Geospatial analysis. *The SAGE handbook of social media research methods*, 255.
- [12] Jabalameli, S., Xu, Y., & Shetty, S. (2022). Spatial and sentiment analysis of public opinion toward COVID-19 pandemic using twitter data: At the early stage of vaccination. *International Journal of Disaster Risk Reduction*, 80, 103204.

- [13] Beigi, G., Hu, X., Maciejewski, R., & Liu, H. (2016). An overview of sentiment analysis in social media and its applications in disaster relief. *Sentiment analysis and ontology engineering: An environment of computational intelligence*, 313-340.
- [14] Lock, O., & Pettit, C. (2020). Social media as passive geo-participation in transportation planning—how effective are topic modeling & sentiment analysis in comparison with citizen surveys?. *Geo-Spatial Information Science*, 23(4), 275-292.
- [15] Mursi, J. K. A., Subramaniam, P. R., & Govender, I. (2023, September). Exploring the Influence of Pre-Processing Techniques in Obtaining Labelled Data from Twitter Data. In *2023 IEEE AFRICON* (pp. 1-6). IEEE.
- [16] Mujahid, M., Kina, E. R. O. L., Rustam, F., Villar, M. G., Alvarado, E. S., De La Torre Diez, I., & Ashraf, I. (2024). Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering. *Journal of Big Data*, 11(1), 87.
- [17] Zumarniansyah, A., Pebrianto, R., Normah, N., & Gata, W. (2020). Twitter Sentiment Analysis of Post Natural Disasters Using Comparative Classification Algorithm Support Vector Machine and Naive Bayes. *Jurnal Pilar Nusa Mandiri*, 16(2), 169-174.
- [18] Gusti, K. W. (2023). Perbandingan Metode Support Vector Machine dan Logistic Regression untuk Klasifikasi Bencana Alam. *Informatik: Jurnal Ilmu Komputer*, 19(2), 134-140.
- [19] Delimayanti, M. K., Sari, R., Laya, M., Faisal, M. R., & Pahrul, P. (2021). Pemanfaatan Metode Multiclass-SVM pada Model Klasifikasi Pesan Bencana Banjir di Twitter. *Edu Komputika Journal*, 8(1), 39-47.
- [20] Hu, T., She, B., Duan, L., Yue, H., & Clunis, J. (2019). A systematic spatial and temporal sentiment analysis on geo-tweets. *Ieee Access*, 8, 8658-8667.
- [21] Dube, L., & Verster, T. (2023). Enhancing classification performance in imbalanced datasets: A comparative analysis of machine learning models. *Data Science in Finance and Economics*, 3(4), 354-379.
- [22] Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences*, 13(7), 4550.
- [23] Soltanzadeh, P., & Hashemzadeh, M. (2021). RCSTMOTE: Range-Controlled synthetic minority over-sampling technique for handling the class imbalance problem. *Information Sciences*, 542, 92-111.
- [24] Ünalan, S., Günay, O., Akkurt, I., Gunoglu, K., & Tekin, H. O. (2024). A comparative study on breast cancer classification with stratified shuffle split and K-fold cross validation via ensembled machine learning. *Journal of Radiation Research and Applied Sciences*, 17(4), 101080.
- [25] Saravanan, K., Prakash, R. B., Balakrishnan, C., Kumar, G. V. P., Subramanian, R. S., & Anita, M. (2023, December). Support Vector Machines: Unveiling the Power and Versatility of SVMs in Modern Machine Learning. In *2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 680-687). IEEE.
- [26] Parishani, M., & Rasti-Barzoki, M. (2024). CWBCM method to determine the importance of classification performance evaluation criteria in machine learning: Case studies of COVID-19, Diabetes, and Thyroid Disease. *Omega*, 127, 103096.
- [27] Selmy, S. A., Kuche, D. E., & Yang, Y. (2024). Geospatial Data: Acquisition, Applications and Challenges.