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, Received on 25 Dec 2024. Revised on 1 Jan 2025. Accepted and Published on 23 Jan 2025. 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 highdimensional 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.