

1. Introduction

Information and communication technology development has allowed users to connect and share information through social media platforms. One popular platform is Twitter, a social network often used for social media interactions. Founded on July 13, 2006 [1], Twitter is a social networking service that prioritizes short updates or microblogging. Twitter has a very large number of users, including in Indonesia, with 11.8 million Twitter social media users [2], based on data (We Are Social & Hootsuite, 2020).

Twitter allows users to tweet about certain topics and participate in discussions. In the context of information dissemination, Twitter can be an effective tool because shared tweets can be analyzed to find out sentiments related to the topic. One of the most frequently discussed topics on Twitter is forest fires. Forest fires have become a major environmental problem in Southeast Asia in recent decades, with significant impacts on the atmosphere, carbon cycle, and various ecosystem services [3]. Forest fires can broadly impact the population, economy, business, and health in the short and long term [3][4]. Indonesia has an amount of carbon stored in peatlands of 28.1 gigatonnes [5]. For example, forest fires in Indonesia in 2015 resulted in the release of 0.89 gigatonnes of carbon dioxide [6].

Community sentiment towards forest fires also varies. In the Indonesian context, disseminating electronic information that harms one's reputation can be a criminal act following the 2008 ITE Law Article 27 Paragraph (3) [7]. Therefore, analyzing sentiment related to forest fires in more detail and accuracy is important. Some previous studies have used deep learning methods, especially the LSTM model, to perform sentiment analysis with good results. For example, in Yuliana Romadhoni's research entitled "Sentiment Analysis of PERMENDIKBUD No.30 on Twitter Social Media Using Naive Bayes and LSTM Methods", the LSTM method achieved an accuracy of 77% [9]. Another study by Lilyani Asri Utami used PSO-based SVM with 87.11% accuracy in sentiment analysis related to forest fires on the news [7]. Usha Devi Gandhi, Priyan Malarvizhi Kumar, Gokulnath Chandra Babu, and Gayathri Karthick also conducted research using Convolutional Neural Networks (CNN) and LSTM with an accuracy of 87.74% in sentiment analysis on Twitter data [8]. Previous research of Winda Kurnia Sari [10] The multilabel text classification task involves categorizing text into categories. In this study, the Deep Learning method, specifically LSTM with GloVe and Word2Vec features, achieved good performance with the highest accuracy of 95.17% and 95.38%, respectively, as well as an average precision, recall, and F1-score of 95%. Previous research [11] showed that the proposed approach in Arabic sentiment analysis outperformed the basic and competitive methods with better results. Using bidirectional LSTM with feature embedding improves the extraction of aspectual opinion target expressions by 39% compared to the baseline method. In addition, using attention-based features by linking aspect embedding to corresponding word embedding in sentiment polarity classification (AB-LSTM PC) improves aspect sentiment polarity identification by 6%. Research [12] uses deep learning techniques with Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) models to identify sentiment patterns in tweet data. Features like word2vec, stopwords, and tweet words are used in model training and integration. The dataset is 50,000 movie reviews from IMDB, and large Twitter data is processed for sentiment classification. The proposed methodology distinguishes the samples from the real-time environment well and improves the system's effectiveness. The Deep Learning algorithms achieved a % testing accuracy of 87.74% for CNN and 88.02% for LSTM in scoring review tweets and identifying movie reviews. Research [13] Muh Amin Nurrohmat, Azhari SN Classify Indonesian novel reviews based on positive, neutral, and negative sentiments using the LSTM method. The test results show that the LSTM method has better accuracy, precision, recall, and f-measure than the Naïve Bayes method. These studies show significant developments in text mining and text processing. However, there are challenges in feature extraction on unstructured text.

The study conducted by Arliyanti Nurdin, Bernadus Anggo Seno Aji, Anugrayani Bustamin, and Zaenal Abidin aims to compare the performance of insertion of words such as Word2Vec, GloVe, and FastText with classification using the Convolutional Neural Network algorithm. The study results reached an F-Measure value of 0.979 for the 20 Newsgroup dataset and 0.715 for Reuters [14]. Therefore, in this study, we will use the LSTM method and the extended GloVe and FastText features to carry out a sentiment analysis of English tweets related to forest fires. The purpose of this study is to measure the performance of the LSTM model in classifying public sentiment regarding forest fires and to analyze the results of sentiment analysis built using this method. This research will also build two corpus, namely the tweet corpus for the GloVe extension feature and the common crawl corpus for the FastText extension feature, to increase the model's accuracy. Classification of reviews through this process will make it easier for users to categorize opinions into negative or positive more precisely. With a deep learning approach, especially the LSTM method, this research is expected to contribute to developing sentiment analysis in the context of forest fires. In this study, the authors will evaluate the performance of the LSTM model and the expansion features used. In addition, the sentiment analysis will provide a deeper understanding of public sentiment regarding forest fires.