Journal of
Information Systems Engineering
and Business Intelligence

Vol.10, No.3, October 2024
Available online at: http://e-journal.unair.ac.id/index.php/JISEBI

Leveraging Social Media Data for Forest Fires Sentiment Classification: A Data-Driven Method

Warih Maharani [1)*], Hanita Daud [2)], Noryanti Muhammad [3)], Evizal Abdul Kadir [4)]

[1)]*Telkom University, Bandung, Indonesia*
[1)]wmaharani@telkomuniversity.ac.id

[2)]*Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak Darul Ridzuan, Malaysia*
[2)]hanita_daud@utp.edu.my

[3)]*Universiti Malaysia Pahang, Lebuh Persiaran Tun Khalil Yaakob, Kuantan, Pahang, Malaysia*
[3)]noryanti@ump.edu.my

[4)]*Universitas Islam Riau, Pekanbaru, Riau, Indonesia*
[4)]evizal@eng.uir.ac.id

Abstract

Background: The rise in forest fires over the last two years, which is due to rise in dry weather conditions and human activities, have greatly impacted an area of 1.6 million hectares, leading to significant ecological, economic, and health issues, hence the need to improve disaster response strategies. Previous research determined the lack of coverage regarding public response during forest fires with conventional methods such as satellite images and sensor data. However, social media platforms provide real-time information generated by users, along with location information of disaster events. Sentiment analysis helps in understanding the public reactions and responses to natural disasters, thereby increasing awareness about forest fires.
Objective: The purpose of this research is to assess the efficiency of Long Short-Term Memory (LSTM) method in classifying sentiment for social networks in regard to forest fires. This research aims to examine the effect of TF-IDF, unigram, and the FastText features on the effectiveness of the classification of sentiment.
Methods: The precision, recall, and F1 score of 2, 3, and 4 determined in the LSTM models with commonly available sentiment analysis tools, such as the Vader Sentiment Analysis and SentiWordNet was used to evaluate the performance of the model.
Results: With an improvement of roughly 10%, the four layers of the LSTM model generated the best performance for the evaluation of sentiments about forest fires. The LSTM model with FastText achieved F1, recall and precision scores of 0.649, 0.641, and 0.659, which exceeds lexicon-based method including SentiWordNet and Vader.
Conclusion: The experimental results showed that the LSTM model outperformed lexicon-based methods when used to analyse the tweets related to forest fire. Additional research is required to integrate rule-based models and LSTM models to develop a more robust model for dynamic data.

Keywords: Forest Fire, Disaster, Long Short-Term Memory, LSTM, Vader, SentiWordnet

Article history: Received 27 January 2024, first decision 19 April 2024, accepted 4 October 2024, available online 28 October 2024

I. INTRODUCTION

Forest fires are large destructive outbursts that had escalated significantly, affecting nearly twice the damage to these areas compared to two decades ago. The outbursts should be accompanied with preventive and control initiatives to reduce international conflicts arising from this natural disaster [1], [2]. Previous research stated that forest fires currently destroyed three million hectares of land and forest annually, in comparison to 2001. Additionally, over 25% of agricultural land in the past two decades had been destroyed. Attaining extremely high intensity since the inception of the twenty-first century, forest fires had burned 9.3 million hectares of land in 2021 [3], [4]. Despite the decline in 2022, 6.6 million hectares were also damaged [5]. Wildfires tend to be caused by several factors, including human activities and climate change. Land clearing activities for agriculture and mining purposes led prolonged dry season,

[*] Corresponding author

ISSN 2443-2555 (online) 2598-6333 (print) © 2024 The Authors. Published by Universitas Airlangga.

doi: http://dx.doi.org/10.20473/jisebi.10.3.392-407

and a decrease in rainfall intensity. Prior research adopted traditional methods, including satellite imagery and sensor data, to detect forest fires [6], [7]. However, several challenges were presented, specifically the lack of coverage regarding public response in the event of the natural disaster.

The challenges were delays in data processing, complexity of capturing instant changes, particularly in remote regions, and obstacles directly associated with acquisition of information by the local communities. Vani et al. [8] and Kang et al. [9] carried out a research on the use of satellite data to identify the regions with higher probabilities of occurrence. However, the investigation showed an effective result, outlining the importance of rapid data sources. Pham [10], Preeti et al. [11], Wu et al. [12], and Yang et al. [13] conducted a similar research by adopting machine learning techniques to examine ground report data for the use of fire prediction. Considering this perspective, certain issues were associated with the need for real-time data to improve reliability and performance. In the past decade, various social media platforms such as X, Facebook, YouTube, and Instagram had generated a immeasurable user-generated data. This provided up-to-date information that enhanced the reliability and current situation of forest fires.

Conventional monitoring systems are used by social media users to frequently respond to and notify other about local incidents, such as smoke, fire, etc. This user-generated insight provided current information along with location-related details, essential for responding to wildfires in real-time. Several preliminary research had been conducted on the use of social media for disaster management [14], [15], [16], [17]. These showed that the tweets shared by users on various platforms provided useful and real-time information during disaster. Meanwhile, numerous sentiment analysis methods had proven exceptional in classifying sentiment and public responses to disasters. Lever J. et al. consistently explored the potential of integrating Twitter data for disaster management [18], [19], [20], [21]. The research aimed to measure the effectiveness of social media as a source of information, particularly regarding forest fires.

The classification of sentiment in respect to social media served as an early warning system for disasters, raising awareness about forest fire incidents. Previous research adopted various machine learning techniques to analyze sentiment related to forest fire. This also included the integration of social media and geophysical data to formulate effective predictive models. Boulton et al. [22] investigated the relationship between social media engagement related to forest fires and the temporal and spatial dimensions of these occurrences in real time. The research established a basis for the collection of data regarding social impacts of forest fires through social media. It enriched conventional observational data by incorporating information from social media to assist fire departments and government agencies in first-response efforts during the occurrence of a disaster. Additionally, the research focused on the ability of social media to identify natural disasters, and the challenges encountered in analyzing the acquired data. Aramendia et al. [23] evaluated data acquired through Twitter on wildfire incidents recorded in Artenara and Valleseco, Canary Islands, Spain, during summer in 2019. Fitriany et al. [24] conducted similar research, by using data acquired through Twitter to identify forest fires in Riau, Indonesia. The analysis found that the obtained information was effectively used to accurately detect forest fires. Meanwhile, Arief Rachman et al. [25] investigated Forest Management Unit (KPH) of West Kalimantan province, reviewing the effects of the disaster and effective reduction strategies.

The National Disaster Management Agency (BNPB) designed the PetaBencana.id, which relied on the use of social media platforms, including the influence on data collection. This was aimed at enabling the accurate reporting of disaster incidents, such as forest fire, flood, haze, volcanic eruption, or strong winds. People can quickly notify PetaBencana.id of any incidents in respective areas using various social media platforms including Instagram, Facebook, X, and Telegram. Furthermore, the general public uses this app to report potential forest fire incidents [26]. Social media analysis improves satellite imagery and field reports by providing additional information through human observations and experiences, that tends to be omitted in traditional data. Combining these various data sources enabled more comprehensive and suitable models, critical in reducing the impact of wildfires.

Several research assessed the performance of transformer models such as BERT and IndoBERT in analyzing sentiment related to natural disasters. Madichetty et al. [27] compared Convolutional Neural Networks (CNN) and Artificial Neural Networks (ANN) in respect to classifying related tweets as either informative or not informative. Subsequently, Devlin et al. [28] introduced BERT (Bidirectional Encoder Representations from Transformers), a transformer technique aimed at improving the accuracy of sentiment analysis. BERT solved problems encountered during the application of previous methods by using the capacity to understand sentence context bidirectionally. This resulted in a more sophisticated interpretation of sentiment in tweets. Prior research developed IndoBERT, a modification of BERT specifically designed in Indonesian language. Characterized by the ability to effectively capture local language features and contextual relevance, IndoBERT had been proven to enhance sentiment analysis performance in the Indonesian context. Koto et al. [29] focused on investigating the use for sentiment classification in the Indonesian language. The results of the test showed that IndoBERT outperformed CNN. Meanwhile, Long Short-Term Memory (LSTM) offered other benefits compared to the BERT and IndoBERT transformer models, thereby making the method suitable for sentiment analysis tasks, specifically in the dynamic domain. LSTM is more appropriate for situations where sentiment changes dynamically over time, such as in natural disaster data, because it retains information in long sequences. This enabled the modelling of changes in sentiment that occur in the event of natural disasters. Furthermore, LSTM models are more efficient than other methods due to the distinct stable training process, suitable for real-time cases with limited resources. BERT and IndoBERT transformers provide deep contextual understanding, with LSTM balancing performance with practical utility in disaster situations.

The present research does not rely solely on static data sources, namely satellite images or sensor data commonly used in previous analyses. In addition, it adopted the dynamic characteristics of social media data. The public opinion on social media regarding forest fires was examined, with a focus on the enhancement of several features used in the LSTM models. This method broadened perspectives, enhancing accuracy in classifying public sentiment during fire events. The current research mainly contributed to the thorough investigation of sentiment analysis methods using LSTM models. These deftly handled the complex linguistic characteristics of Indonesian language in respect to forest fires data. A detailed investigation of regional idioms, daily language usage, and culturally specific expressions often overlooked by conventional sentiment analysis methods was performed. The feature extraction process was enhanced by combining TF-IDF, unigrams, and FastText, thereby optimizing the performance of sentiment analysis methods. These contributions proved that the incorporation of social media data enabled accurate and optimal results.

The remaining aspect of this research was organized as follows Section 2 described related works on forest fire sentiment analysis methods. Furthermore, Section 3 focused on the detailed exploration of LSTM model. Sections 4 and 5 explained the experiments conducted, as well as a detailed explanation of the results obtained. Finally, Section 6 comprised the conclusions and research limitations.

II. RELATED WORKS

Wildfires had caused environmental problems in various parts of the continents, damaging millions of hectares of land annually. Apart from compromising public health and safety, it influenced emissions, biodiversity, and species survival as well. Some human activities also contributed to the increasing intensity, namely land clearing for agricultural practices, prolonged drought, and climate change. During Australia Black Summer within 2019 and 2020, 18 million hectares of land were destroyed, resulting in significant damage to wildlife and the release of a large amount of CO_2 [30]. In California, there was an intense wildfire that caused severe damages, prompting the government to evacuate a large number of residents [31]. The incident was caused by heatwave and severe drought conditions experienced in several Mediterranean countries, including Greece and Turkey. This forest fires seriously destroyed the environment, affecting the surrounding businesses [32]. Meanwhile, in Indonesia, the following regions Sumatra and Kalimantan characterized by massive forest and peatlands are vulnerable to forest fires during the dry season. According to Forest Watch Indonesia, 1.6 million hectares of land were burned in 2022, compared to the 1.4 million hectares recorded last year (FWI [33]).

The widespread threat of forest fires over extensive areas, required anticipated efforts to reduce the risk of disaster. Traditional methods presented several challenges, including difficulty in obtaining real-time data related to similar events and spatial resolution needed for evaluating fire dynamics. The effective functioning of satellite imagery and sensor data are not reliable enough to address the complexities of public response to forest fire incidents. As a result, emergency services and fire departments had to work extra hard to pull alongside due to the lengthy response time. In addition, the issues increased due to lack of readiness and inefficient allocation of firefighting resources, frequently causing fires to escalate spontaneously when it could had been prevented or reduced earlier. Apart from aggravating the negative effects on the economy and environment, the situation compromised human welfare and security.

Kang et al. [9] stated that satellite data was highly valuable in offering comprehensive coverage. However, it frequently suffered delays in reporting, which tend to be particularly challenging in situations where fires are quick to spread. Delayed response and identification have major effects given that there is less chance for quick and efficient fire control. Pham et al. [10] and Preeti et al. [11] focused on how machine learning improved prediction accuracy. The research relied on historical and field report data, making it unable to provide real-time responses during critical disaster periods [2]. Additionally, current models often did not consider the precise information obtained from sources at ground level. Several research reported the significance of user-generated content during disasters, without specifically applying the method to forest fire scenarios, leaving a gap in using data for prediction purposes [15], [34], [35], [36].

In order to develop a more comprehensive method in analyzing this disaster, there is need to integrate various data sources to complement each other. Current methodology focused on the importance of integrating existing real-time data sources, such as social media, to enrich conventional methods. Furthermore, the development of more flexible and responsive analytical models was essential in terms of handling several difficulties. These models must efficiently use social media data in real-time to solve the issues outlined in prior research.

Several investigations had been conducted on the use of social media data to analyze forest fire events in Indonesia. In addition, all these analyses adopted differing methods, reviewing the issue from diverse viewpoints. For example, Budiharto and Meiliana [37] predicted political events using data acquired from Twitter. The research showed the potential of social media as a forecast specifically in the political sphere. Similarly, Mustaqim et al. [38] focused on public opinion regarding government reaction to forest fires, through information acquired on Twitter. The research showed that the platform was an effective tool for measuring public opinion on environmental matters [38]. Irawanto et al. [39] classified forest fire tweets using Naïve Bayes, Random Forest, and Support Vector Machine (SVM). Moreover, Negara et al. [40] analyzed forest fires in Riau by applying Decision Trees and Bayesian Networks based on weather data. Yang et al. [13] used remote sensing data to predict forest fires in Indonesia by comparing several machine learning methods. The research proved the adopted method had better accuracy compared to baseline models.

Syarifudin et al. [42] used 1D CNN to predict forest fire hotspots, outlining the significance of data analysis in addressing environmental issues. Furthermore, Nurkholis et al. [7] showed that the C5.0 algorithm was effectively used to classify hotspots, making substantial contribution to the advancement of forest fire predictive models based on regional characteristics. Evelina et al. [41] investigated the effects of a public awareness campaign on forest fires carried out by the Indonesian Ministry of Environment using data obtained from Instagram. The main focus was on how social media influenced environmental communication. Meanwhile, Muslikhin et al. [43] examined data on forest fires during the COVID-19 pandemic obtained through Twitter. The perspective adopted was relatively different because it focused on how people communicated during an environmental crisis.

Nikonovas et al. [44] reported the novelty in the invention of ProbFire, by incorporating climate forecasts into an early fire detection system, thereby enhancing management strategies. As an extra source of information for the management of peatland fires and haze events, Kibanov et al. [45] reviewed the relevance of social media, particularly Twitter. Viktor Slavkovikj et al. [46] further conducted extensive research on several methods for identifying forest fires, focusing on the importance of social media as a tool for real-time public monitoring and participation in related events. This research analyzed the speed and simplicity of information availability from social media by investigating the shift from traditional static data sources such field reports and satellite imaging. Gorka Zamarreño-Aramendia et al. [23] examined the proactive social media use to help prevent disasters in the Canary Islands, where forest fires are a continual threat. The investigation reported how effectively social media enabled public-emergency response lines of communication to be improved, thereby increasing situational awareness and reactions.

This research outlined the relevance of social media to provide real-time informatin, in order to improve conventional emergency strategies in the framework of disaster management. However, the main focus remained on examining the content shared on social media including the rapid information distribution advantages. User-generated content on these platforms was closely related to the timing of disasters and the geographical location of occurrence. Prior research did not examine social media posts containing sentiment, which tended to provide a comprehensive insight into public opinion and emotional reactions during natural disasters. Therefore, it is necessary to improve sentiment analysis methods that can interpret human expression with optimal performance.

Considering the improved performance of sentiment analysis methods, this research combined the linguistic traits of the Indonesian language including local idioms, slang, and culturally specific expressions. The aim was to expand the features used for modeling as a novelty distinguishing it from previous research. Additionally, social media data on forest fires were used to analyze and monitor the disaster. This also included developing and validating sentiment analysis method based on LSTM designed to suit the linguistic and cultural context of the Indonesian language.

III. METHODS

The research objective focused on the development of sentiment analysis method using a LSTM model, and forest fire dataset as shown in Fig.1.

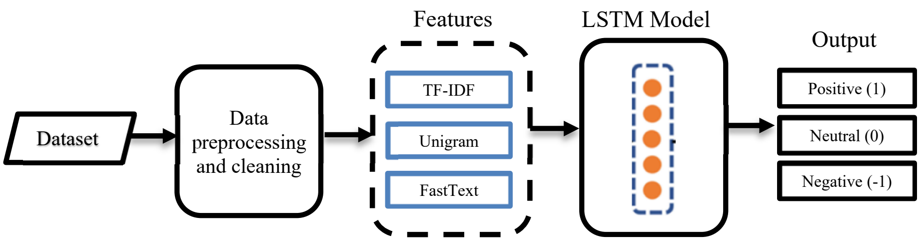

Fig. 1 Sentiment Analysis Model using LSTM

A. Data Collection

In Indonesia, the selection of social media platforms depended on popularity, high number of active users, completeness and richness of content shared in real-time, as well as data accessibility. X or Twitter, Facebook, and Instagram were selected because of the widespread use for the dynamic real-time updates, community discussions, and reviewing visual content. This research adopted Application Programming Interfaces (APIs) and web scraping tools. APIs provided structured data directly from social media platforms, while web scraping was used due to limited access to API access. Subsequently, keywords and hashtags related to the disaster, such as kebakaran hutan, karhutla, kabut asap, and forest fires, were used to filter relevant posts. Consultations with environmental experts enabled the development of data, refined through initial validation to ensure completeness. Due to the occurrence of several forest fire events in the past two years, data was gathered within 2019 and 2021. In the middle of 2019 and towards the end of 2020, Sumatra and Kalimantan experienced widespread forest fires that gained considerable public attention due to the adverse effects on air quality and biodiversity. According to Drone Emprit social media observers, X or Twitter had the most conversations about forest and land fires (karhutla) in 2019, with 269,141 tweets. Additionally, the number of conversations regarding these incidents increased significantly in Riau. This period included several seasons and fire intensities, presenting an entire picture of social media activity related to forest fire.

A dataset consisting of 7,515 tweets in Indonesian related to forest fires was used, with three annotators interpreting dataset manually. All annotators have given their consent after being informed of all relevant aspects necessary for their decision to participate, including the objectives of this research and the methods used. Additionally, they were also informed that their information would be kept confidential and used only for academic research purposes. The annotators were undergraduates from different backgrounds who communicated in Indonesian fluently, thereby exhibiting exceptional knowledge during the annotating process. The annotations were conducted according to established guidelines. Each tweet was individually interpreted by the three annotators to improve dependability, and minimize personal prejudice during sentiment classification process. A consensus discussion was held when the annotators disagreed about the classification. Cohen kappa statistic allowed several research to measure the annotator degree of agreement. Considering the possibility of random agreement, Cohen kappa was a reliable measure between annotators. Moreover, the value of 0.91 obtained showed the annotators had nearly perfect agreement. The training and testing dataset comprised exclusively of manually annotated data provided by three undergraduates.

The polarity column in Table 1 represents sentiment classification of the analyzed textual data, with the corresponding numeric value. In addition, a value of -1 denoted a quite negative sentiment marked by pessimistic emotions, unhappiness, anger, or sadness. However, a value of 0 showed a neutral attitude without any tendency toward either positive or negative sentiment. This also showed textual data reflecting objective facts or neutral information lacking positive or negative emotions. A rating of 1 showed a favorable attitude including happy feelings, hope, satisfaction, or pleasure.

Table 1 shows an illustration of dataset, partitioned into training and testing sets based on an 80:20 ratio. Consequently, 80% of data was allocated for training the proposed models, while the remaining 20% was set for testing. It was guaranteed that the training and testing datasets fairly reflected the entire information by using random sampling methods. These ensured random data point selection for each dataset, thereby minimizing the potential for non-random selection process-induced bias.

TABLE 1
THE EXAMPLE OF DATASET

Sentences	Polarity
Forest fires make tourist areas in Australia smoky	-1
Finally, get food in peace without people around who do not care about forest smoke	1
I saw the weather in Jakarta and why it suddenly rained. Smoke. I forgot that today is a solar eclipse.	0
My city is full of smoke because of forest fires, after a while, I'm annoyed by cases like this	-1
Forest fires including in Indonesia cause smoke	-1

B. Data Preprocessing and Cleaning

Forest fire sentiment dataset may contain irrelevant information affecting model performance. Therefore, data preparation and cleaning stage is quite important to guarantee the development of an exceptional model. The process included transforming unstructured data into a structured format, to ensure it is ready for analysis [47], [48]. The initial stage required thorough data filtering to remove irrelevant content, such as hashtags, advertisements, spams, and off-topic discussions. A contextual filtering process was the removal of tweets not related to forest fires. For example, although the phrase api cinta contains the keyword api, it is not related to forest fires. In addition, there are several unique challenges in sentiment analysis that included regional languages. The research filtered several regional languages shared on Twitter using automatic language detection tools in conjunction with thorough manual categorization. For example, a tweet in Javanese about the eruption of Mount Merapi was translated into Indonesian to maintain dataset consistency. Another Javanese tweet related to forest fire Gunung Merapi mbledos, akeh asap Tebal, was translated into Indonesian. The preprocessing stage comprised the elimination of personal identifiers to guarantee privacy, while complying with regulations regarding data protection. This also included the preservation of ethical standards enabling the conduct of research based on related principles.

This section provided a comprehensive description of the steps implemented during the preparation and preprocessing of dataset. Text normalization was adopted as an initial preprocessing step to handle informal words in Indonesian. This method was used to standardize data in order to reduce variance. Meanwhile, the preprocessing phases consisted of the following

1) Lowercasing

Lowercasing refers to the process of converting all text to lowercase, thereby promoting consistency and allowing the model to focus on the meaning of words rather than respective capitalization, as shown in Table 2.

2) Noise Removal

Eliminating punctuation, emojis, URLs, HTML, as well as non-essential characters including numbers and special fonts comprised of noise removal. However, this research acknowledged emoticons and expresses sentiment. For example, after the lowercasing stage, the following tweet: `Fire in Jakarta!!! 😱` can be transformed into: `fire in jakarta 😱`.

Punctuation removal stage focused on eliminating commas (,), periods (.), exclamation (!), and question marks (?). Although, punctuation does not affect the analysis process, it needs to be removed to enable the model focus on representing sentiment in words. Table 2 shows that eliminating punctuation marks helped the model to concentrate on sentiment.

Remove URLs and HTML tags frequently contained in social media posts and web scraped data. These elements are unnecessary in sentiment analysis and should be removed to improve data quality.

3) Text Normalization

Text normalization refers to the process of addressing informal words and slang. This process includes substituting informal words and slang with formal counterparts using a dictionary specially designed to suit Indonesian Internet language. For example, the abbreviation yg was substituted with the word yang, and bgt with banget. Expanding the occurrence of repeated letters frequently used in informal Indonesian to accentuate specific words, such as transforming baguuusss into bagus. Numerous sources of linguistic resources including informal-formal corpus, alongside alay dictionaries, resulted in a total of 17,039 terms. Furthermore, the system examined data identifying matching terms in the informal-to-formal or alay dictionaries. The scanning process adopted pattern-matching algorithms that identified various word forms, including misspellings, alays, and abbreviations. An example of text normalization is shown in Table 2.

TABLE 2
THE EXAMPLE OF TEXT NORMALIZATION

Original Text	Lowercased Text	Noise Removal	Text Normalization
PEMADAMAN KEBAKARAN HUTAN DI KALIMANTAN MEMERLUKAN BANTUAN! kebakaran tidak kunjung padam. parah bgt. kebakaran hebat di hutan, ga ada yang peduli. Info terbaru terkait kebakaran hutan bisa dilihat di https://www.kompas.com/tag/kebakaran-hutan-dan-lahan! #prayforIndonesia #prayforsumatra (Forest fire extinguishing in Kalimantan needs help! The fire won't die out. very bad. a severe fire in the forest, nobody cares #prayforIndonesia #prayforsumatra)	*pemadaman kebakaran hutan di kalimantan memerlukan bantuan! kebakaran tidak kunjung padam. parah bgt. kebakaran hebat di hutan, tidak ada yang peduli. info terbaru terkait kebakaran hutan bisa dilihat di https://www.kompas.com/tag/kebakaran-hutan-dan-lahan! #prayforindonesia #prayforsumatra* (forest fire extinguishing in kalimantan needs help! the fire won't be die out. very bad. a severe fire in the forest, nobody cares #prayforindonesia #prayforsumatra)	*pemadaman kebakaran hutan di kalimantan memerlukan bantuan kebakaran tidak kunjung padam parah sekali kebakaran hebat di hutan tidak ada yang peduli info terbaru terkait kebakaran hutan bisa dilihat di prayforindonesia prayforsumatra* (forest fire extinguishing in kalimantan needs help the fire won't be die out very bad a severe fire in the forest, nobody cares prayforindonesia prayforsumatra)	*pemadaman kebakaran hutan di kalimantan memerlukan bantuan kebakaran tidak kunjung padam parah banget kebakaran hebat di hutan tidak ada yang peduli info terbaru terkait kebakaran hutan bisa dilihat di prayforindonesia prayforsumatra* (forest fire extinguishing in kalimantan needs help the fire won't die out very bad. a severe fire in the forest, nobody cares prayforindonesia prayforsumatra)

4) Tokenization

Tokenization is the process of dividing a sequence of text or segmentation of similar strings into smaller units called tokens. Meanwhile, NLTK tools were used in the actualization of this objectives. After the process of tokenization, the phrase kebakaran hebat di hutan (*a severe fire in forest*) was transformed into a list of words [kebakaran, hebat, di, hutan].

5) Stop words removal

Stop words removal refers to the elimination of frequently used words that lack substantial meaning. Several words in Indonesian, such as yang (*which*), dan (*and*), and di (*in*) were eliminated to prioritize more significant terms. For example, in the following list of words [kebakaran, hebat, di, hutan], stop-word removal would eliminate the term di in order to outline the importance of [kebakaran, hebat, hutan]. Elimination of stop words enabled the model focus on sentiment-specific terms, thereby reducing the size of data. Indonesian stop word list can be obtained from online sources or generated by analyzing the word frequency in the acquired dataset.

6) Stemming

Stemming aims to improve computational efficiency and lower vocabulary size by transforming words into basic forms. For example, the word kebakaran (*fire*) can be transformed into the base form bakar (*fire*).

C. Features Extraction

Feature extraction is a critical phase responsible for converting raw text data on forest fires in Indonesia into a format acceptable to the LSTM model used for sentiment analysis. However, unigrams and Term Frequency-Inverse Document Frequency (TF-IDF) were used as feature sets for the LSTM model. TF-IDF is a metric adopted to evaluate the importance of words within the corpus relative to specific documents [49]. This metric allocates weights to terms according to the frequency across documents. Unigrams, which denote singular isolated words, were adopted to encapsulate the essential lexical content of social media posts [50]. In order to optimize the performance of the LSTM model by providing a comprehensive and contextual representation of textual material, TF-IDF and unigrams were used for feature extraction. This method enabled the model to understand the sequential structure of text, including the significance and relevance of certain words and phrases in forest fire framework. The current research examined the ramifications of the FastText feature, an advanced word representation model that encapsulated the semantic characteristics of words [51]. FastText encoded each word as a character n-gram, for example the term forest was analyzed into the following n-grams fo, or, re, es, st. It subsequently encoded each of these n-grams as a vector within a high-dimensional space. The final word vector focused on the culmination of these n-gram vectors. Moreover, FastText effectively captured the semantic significance of individual words as well as the nuances and contexts arising from the amalgamation of these n-grams [51].

This subsection described the methods used in feature extraction, specifically focusing on the application of TF-IDF, unigrams, and FastText embeddings.

1) TF-IDF

TF-IDF is a statistical metric adopted to assess the significance of a word in a set of documents. This statistical metric is a widely used method for converting text data into a matrix of TF-IDF features. TF quantified the frequency of terms occurring in documents, which vary in length. Therefore, some terms might appear frequently in longer documents than in shorter ones. This led to the normalization process focused on dividing the TF by the document length, denoting the total number of terms in the document as stated in (1).

$$TF(t) = \frac{\text{Number of times term } t \text{ appears in a document}}{\text{Total number of terms in the document}} \tag{1}$$

The IDF described the importance of terms, which possess equal significance throughout TF computation. Some terms, namely is, of, and that, often occur frequently but lack specific significance. Therefore, it is crucial to assign greater importance to infrequent terms while reducing the significance of common ones, achieved by using the formula in (2).

$$IDF(t) = \log \frac{\text{Total number of documents}}{\text{Number of documents with term t in it}} \tag{2}$$

The TF-IDF score was obtained by multiplying the TF and IDF, showing the frequency of a term while considering the relevance in the entire collection of documents. Subsequently, this score served as an input for the LSTM model.

2) Unigram

The unigram in text analytics is a probabilistic language model used to predict the next item in a given sequence. The model assumed terms are independent of each other, treating respective word occurrence as a separate event. The probability of a word in a document can be determined by calculating the relative frequency, as stated in (3).

$$P(word) = \frac{frequency\ of\ the\ word}{total\ number\ of\ words} \quad (3)$$

The feature vector for each document was constructed by calculating the probabilities of all distinct words in the text. This method offered a straightforward and efficient way to represent text data in the proposed model.

3) FastText

FastText was designed to capture the morphological structure of terms to augment the powers of the Word2Vec model. This method ensured each word was viewed as a mixture of n-gram characters, enabling the exchange or representations between terms with related subwords. Moreover, the method practically controlled the agglutinative traits of the Indonesian language, in which several word variants had the same root. FastText generated vector representations for every word, characterized by unique n-gram subwords and trained to decipher an entire document collection. During the training phase, the size of each word vector was altered in order to obtain terms found in the corpus, including comparable contexts in a spatial representation. On completion of the training process, a FastText vector was assigned to every word contained within a document. The vectors were combined by averaging or adding it together. This led to the generation of a feature vector of predefined size covering the entire composition, later used by the LSTM model.

D. Classifier Modelling using LSTM

LSTM is a variation of Recurrent Neural Network (RNN), specifically designed to efficiently solve the sequence prediction challenge [52], [53]. Due to the vanishing and ballooning gradient difficulties, traditional RNNs struggled with long-term dependencies even when good at remembering information from past inputs. LSTM—developed by Hochreiter & Schmidhuber [52] provided a way to either selectively retain or discard data obtained for long durations. This method was fundamentally based on maintaining the state of cells traversing the whole sequence ensuring the network decides to discard or preserves information.

The gates helped the model to develop the following capability [52]. Forget Gate (f_t) decides what information from the cell state should be discarded or retained, using a sigmoid function. Supposing *ft* outputs has a value approximate to 0, it shows forget, while 1 implies retain [52]. Input Gate (i_t) updated the cell state with new information, generating a vector of new candidate values and combining it with the output of the forget gate. Furthermore, the gate comprised two parts, namely sigmoid layer that decides the values to be updated and tanh layer, which generates a vector of new candidate values. Output Gate (o_t) was used to determine the next hidden state, considering the current input, previous hidden, and recent cell states in terms of making this decision.

The LSTM gates obtained a hidden state from prior steps, in addition to the current time step input. The values of the input, forget, and output gates were defined by three fully connected layers of sigmoid activation functions applied on the input data. It was ensured that the values of the three gate variables remained between 0 and 1 in respect to the sigmoid activation function. Moreover, the input nodes are important, usually assessed with hyperbolic tangent (tanh) activation function. The input gates controlled or regulated how internal state of the memory cell incorporated values of input nodes. The forget gates were responsible for deciding whether to preserve or discard the present memory value. The output gates determined whether the memory cell should affect the output at the current time step. The architecture of the LSTM gates enabled the cell state to control long-term dependencies by ensuring a balance between incorporating and maintaining current information. This helps the LSTM to be adept at tasks language modeling, speech recognition, and time series prediction when knowledge of context is essential. The LSTM model consisted of several layers, including a sequence of memory cells capable of long-term memory retention. However, the model depth—or number of layers—directly influenced the complexity and ability to learn from data including high degrees of abstraction. The LSTM model was developed by incorporating several layers— specifically 2, 3, and 4. The research investigated how network depth influences the capacity to manage and examine advanced linguistic data realized from social media.

In line with this perspective, hyperparameters played a crucial role in maximizing the performance of the LSTM model. Furthermore, this was influenced by the memory cell size, learning rate, dropout frequency, and the number of layers. The performance of the LSTM model was highly affected by the architecture and hyperparameter settings. The optimization of hyperparameter generally relied on empirical methods comprising multiple experiments to determine the optimal configuration for a specific task.

E. Evaluation

LSTM-based prediction models were defined by relevant performance indicators and strategies to assess the capacity in forecasting future occurrences or classifying data relying on historical trends. Additionally, prediction performance was assessed using Precision, Recall, and F1-score. Precision refers to a statistic measure of the exact expected positive events among all the forecasts. Recall computes the fraction of exact expected positive outcomes in respect to the total count of actual positive events. In cases when the effects of false positives and negatives are significant [54], [55], these tests provided an entire picture of the model performance.

Precision was calculated using (4).

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

Recall was calculated using (5).

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

Where, TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives.

In order to calculate to the harmonic mean, the F1-score was combined with Recall and Precision. This mean is particularly useful for evaluating models aimed to balance recall with Precision. F1-score, the harmonic mean of Precision and Recall, can be calculated using (6).

$$F1 - score = 2x \frac{Precision\ x\ Recall}{Precision+Recall} \quad (6)$$

IV. RESULTS

This research aimed to determine the most effective combination of features and model configurations in analyzing sentiment expressed in social media data on forest fires. The method ensured that sentiment analysis model captured various linguistic variations and contextual information, thereby enabling improved performance. Two experimental scenarios were carried out using different combinations of feature sets, including LSTM layers. This was aimed to investigate the effects of feature set variations and LSTM depth. The first scenario focused on a combination of TF-IDF and Unigram features, with TF-IDF assessing the relative value of the term in a document compared to the entire corpus. This method discovered significant terms in dataset, by ensuring word frequency was in line with the inverse. Meanwhile, unigram features identified words with different attributes, representing the specific frequency in the text. The integration of TF-IDF with unigrams aimed to encapsulate both lexical significance and frequency, thereby offering a more thorough representation of textual data. This research examined LSTM models with differing depths of 2, 3, and 4 layers to evaluate the impact of model complexity on performance.

In the second scenario, the feature set was enhanced by aggregating TF-IDF and unigram features with FastText embeddings. This feature, designed by Facebook AI Research (FAIR), represents every word as a vector in a continuous space, which led to the capturing of semantic links. This enabled the model discover word frequency and relevance, contextual similarities and variations. FastText embeddings combined with TF-IDF and unigram features improved the feature set, allowing the model to incorporate semantic information, including statistical significance and word frequency. Similarly, in Scenario 1 the LSTM models with 2, 3, and 4 layers were investigated to assess how the integration of these characteristics, affected the general performance of sentiment analysis.

The incorporation of multiple layers in a network enabled the acquisition and generalization of different characteristics levels from the input data. In the LSTM models, layers 2, 3, and 4 were used to investigate how model depth affected the ability to capture and generalize different levels of characteristics. The baseline is a two-layer LSTM model, in which the first layer captures the basic characteristics of the input data, with the second enhancing these features in order to identify intricate patterns. This configuration generated a structure to evaluate the effectiveness of incorporating more complexity functions. The 3-layer LSTM model provided an extra level of abstraction by integrating features acquired from previous layers. This structure aided in increasing the model ability to generalize and understand complex patterns, thereby improving performance on classification tasks. However, the incorporation of additional layers into the 4-layer LSTM model, enhanced the ability to capture complex and linguistic nuances. The model tend to represent complex data by increasing the depth. This was proven to be important for sentiment analysis tasks constituting complex data, depending on the context of social media. The experiment aimed to achieve an optimal balance between model complexity and performance, by investigating the comparison of several variations in depth. In addition, the proposed method was compared with lexicon-based processes including Vader, and SentiWordnet [57].

The properties of each layer, and distinctive appropriateness for the assigned task, are taken into consideration when determining the activation function. The Rectified Linear Unit (ReLU) activation function for the hidden layer in the LSTM model was adopted. ReLU tend to avoid issues associated with the vanishing gradient, because the function captures more complicated patterns by representing non-linearity in the neural network. This enabled the reflection of unlike sigmoid and tanh functions, thereby enhancing and optimizing deep network training. The function also diminishes model complexity, improving computing efficiency. These advantages allowed the application of ReLU in the hidden layer of the proposed LSTM model to ensure efficient learning and enhanced performance. Meanwhile, softmax is the appropriate method for generating probability distribution of classes based on the research objective. This was realized by assigning a single sentiment class to every text input. The model architecture, which comprised the number of layers and activation functions had proven optimal for accurately classifying sentiment, in accordance with the main research objective.

Based on this perspective, 50 epochs of training were conducted for the proposed LSTM models. The first experiment produced consistent convergence values and optimal performance without overfitting. The model successfully achieved a balance between training time and performance by continuously monitoring training and validation loss during each epoch. Additionally, the hyperparameter settings are shown in Table 3.

TABLE 3
THE HYPERPARAMETER

Methods	Type
Batch size	128
Max_features	20,000
Max_len	200
Activation	Relu / Softmax
Learning rate	0.001
Hidden Unit	512
Layer	2 / 3 / 4
Dropout	Yes
Embedding_size	300

Table 4 shows dataset performance based on a ratio 80:20, and in comparison, with the LSTM model with 2-layers, it was reported that the proposed 4-layers model improved performance by approximately 10%. In the second scenario, this feature was enhanced by adopting FastText. The performance of the LSTM model incorporating FastText features, in contrast to Vader and SentiWordNet, are shown in Table 5. Fig. 2 focused on the most effective performance of the combined LSTM model, comprising 4-layers, TF-IDF, unigram, and FastText. The 4-layers model outperformed the others in terms of extracting and interpreting the nuanced context in social media texts concerning forest fire.

The obtained results showed that the LSTM model with FastText exhibited better Precision, Recall, and F1-score. The results showed specific trends in the performance of improved models independent of the settings applied. The performance depended on the number of layers. For example, a 2-layer LSTM model showed a modest precision and recall values of 0.619 and 0.614, respectively. This implied that the LSTM model can balance the trade-off between Precision and Recall. The addition of another layer to the 3-layered model produced a slight improvement in precision (0.631), recall (0.620), and F1score (0.625). This showed that the incorporation of more layers enabled the optimally analysis of complex data patterns, resulting in better prediction outcomes. LSTM model with 4-layers achieved the best performance with a precision, recall, and F1-score of 0.659, 0.641, and 0.649, respectively. Generally, the results of the experiment showed that it outperformed the lexicon-based model.

TABLE 4
THE PERFORMANCE (SCENARIO 1)

Methods	Precision	Recall	F1-score
LSTM (TF-IDF, unigram, 2 layers)	0.575	0.488	0.528
LSTM (TF-IDF, unigram, 3 layers)	0.586	0.474	0.524
LSTM (TF-IDF, unigram, 4 layers)	0.604	0.591	0.597

TABLE 5
THE PERFORMANCE (SCENARIO 2)

Methods	Precision	Recall	F1-score
LSTM (FastText, TF-IDF, unigram, 2 layers)	0.619	0.614	0.616
LSTM (FastText, TF-IDF, unigram, 3 layers)	0.631	0.620	0.625
LSTM (FastText, TF-IDF, unigram, 4 layers)	0.659	0.641	0.649
Vader	0.645	0.627	0.635
SentiWordNet	0.628	0.630	0.628

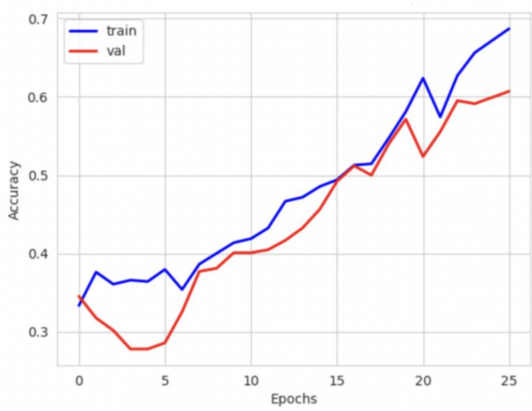

Fig. 2 LSTM sentiment with FastText, TF-IDF with 4 layers

In order to discover the statistical significance of the correlation between the number of LSTM layers and the model performance, an ANOVA (Analysis of Variance) analysis was conducted on Precision, Recall, and F1-scores using Python with SciPy and statsmodels, as shown in Table 6.

The ANOVA showed that the Precision, Recall, and F1-score of the LSTM model were significantly influenced by the number of layers. These results were proven by the extremely high F-statistic and near-zero p-values observed for all three metrics. The results presented strong statistical evidence proving that deeper LSTM models performed exceptionally in terms of Precision, Recall, and F1-score for sentiment analysis of social media text concerning forest fire.

TABLE 6
THE ANOVA FOR PRECISION, RECALL, AND F1-SCORE

Metric	Source	df	Sum Sq	Mean Sq	F-value	P-value
Precision	C(Layers)	2	2.528000e-03	1.264000e-03	4.101915e+28	3.912046e-85
	Residual	6	1.848893e-31	3.081488e-32		
Recall	C(Layers)	2	1.206000e-03	6.030000e-04	4.892117e+28	2.306074e-85
	Residual	6	7.395571e-32	1.232595e-32		
F1-score	C(Layers)	2	1.746000e-03	8.730000e-04	2.833047e+28	1.187415e-84
	Residual	6	1.848893e-31	3.081488e-32		

V. DISCUSSION

LSTM models were specifically designed for the sequential nature of data, due to the unique architecture including memory cells and gates controlling the flow of information. During the examination of social media text, these models succeeded in capturing long-term dependencies in data sequence. Social media posts occasionally featured a sequence of threads or messages whose context and sentiment evolved with time. Conventional models based on the bag-of-words method or other neural networks such as feedforward designs, often failed to capture the temporal dynamics because it handled each input independently, regardless of the order or sequence [52]. The capacity of LSTMs to selectively retain and discard data practically led to enhanced ability. The gating mechanism combined input, output, and forget gates to provide this capacity. Additionally, this feature enabled LSTMs to discard data irrelevant to future states, as well as retain relevant information over long sequences. For social media sentiment analysis, LSTMs were used to retain important emotional cues from past interactions, including understanding complex user interactions lasting for hours or days. Sentiment expressed initially influenced the background of next postings. Furthermore, LSTM effectively addressed the vanishing gradient problem. The ability to model long-term dependencies was limited by the gradients propagated through time and layers, becoming either extremely small (vanishing) or large (exploding). This issue was resolved through LSTM gate structure— regulating the gradient flow. LSTM performed better in terms of learning efficiency, achieving higher accuracy and stability, when processing long or complex data sequences discovered on social media streams [52]. The results showed that the models were properly-suited for tasks requiring an understanding of data temporal structure. Sentiment analysis on social media, refers to the expressions of feelings and opinions interwoven through temporal interactions, which requires the ability to accurately process and evaluate these sequences.

TF-IDF considered the occurrence of the most significant words in a given text collection. In addition, FastText processed terms that were not encountered during training by decomposing the invisible words into fragments known as n-grams, as well as merging the respective vectors. The process allowed FastText to provide a meaningful representation even for out-of-vocabulary (OOV) words frequently encountered in social media language. The method is important in the context of social media, where new words and slang are used rapidly, specifically during natural disasters namely forest fires. Therefore, it was inferred that the use of a classification model incorporating the features facilitated improved performance. This effectively captured the distinct characteristics of the input text based on respective operational modes.

A positive relationship existed between the layer depth and distinct performance as showed in Table 6. The result showed that deeper networks effectively evaluated the diverse and intricate linguistic characteristics of social media text in respect to sentiment analysis. The increase in the number of LSTM layers showed an enhanced overall performance which, resulted in the marginal enhancement of precision, and a decline in recall. The model ability to accurately detect positive cases was improved at the expense of overlooking cases that should had been detected. The F1-score improvement was negligible, showing a minimal overall enhancement in model performance. The LSTM model with 4-layers showed a significant improvement in both precision and recall when compared to the two preceding models. A higher level of precision showed a superior capability to accurately identify true cases, while increased recall led to greater capture of a larger proportion of relevant cases. The F1-score exhibited substantial superiority over the other two models, implying an improved performance in terms of precision and recall.

The performance metrics of the 4-layer LSTM model, with Precision, Recall, and F1-scores less than 70%, signify the challenges and constraints of sentiment analysis associated with social media data regarding forest fires. Despite efforts to address informal language, slang, abbreviations, and emoticons, these terms complicated sentiment analysis task. The complex characteristics of dataset and subtle differences in language used were significant factors of the low-performance metrics.

The 4-layers LSTM model obtained performance metrics (Precision, Recall, and F1-score) comparable to Vader and SentiWordNet. The model achieved higher precision and F1-score compared to rule-based ones, such as Vader and SentiWordNet, despite the competitive nature, particularly in situations where computational resources or extensive training data was unavailable. This is in line with the results by Khan et al. [58], that rule-based models succeeded in performing sentiment analysis on social media. Additionally, Vader used specifically designed structure for microblog data to accurately interpret contemporary slang and emoticons. As a result, it achieves a high level of precision and recall without requiring extensive training datasets. Similarly, SentiWordNet adopted a lexicon-based method to maintain a well-balanced level of accuracy and completeness, showing the effectiveness in various scenarios.

Madichetty et al. [59] stated that while CNNs are effective for short texts, CNN overlook long-term dependencies and LSTM models, important for understanding sentiment in social media posts. Meanwhile, Koto et al. [29] proved that IndoBERT improved sentiment analysis using the Indonesian language, although this required significant computational resources. The proposed LSTM model was more suitable for real-time applications with limited resources. Barik and Misra [60] stated that enhancing the VADER lexicon classifier improved sentiment classification accuracy. This is in line with the results that VADER remained competitive, specifically in resource-constrained settings. Izsak et al. [61] investigated efficient BERT training methods with limited resources, focusing on the adaptability of transformer models while requiring more resources than the proposed LSTM model. The results showed that advanced models including BERT and IndoBERT offered superior performance, characterized by higher computational costs and data requirements. In resource-constrained environments, rule-based models such as VADER and SentiWordNet were regarded as competitive alternatives. The 4-layers LSTM model provided a balance between performance and efficiency, resulting in the suitable for real-time sentiment analysis.

This comparison focused on the benefits of using sophisticated and adaptive neural architectures such as LSTMs over stationary rule-based models for challenging language understanding tasks. Despite the simplicity and less computational features, lexicon-based models are not as functional as LSTMs [62], which had proven to be more robust and flexible in analyzing wildfire sentiment on social media, where language and emotions change fast and are mostly idiomatic. The exceptional accuracy helped to identify relevant emotions, facilitating effective analysis. The results obtained suggested that increasing LSTM layer counts improved model performance, and this is in line with the present research. According to Zhang et al. [63], neural networks produced better results in text classification since it can effectively show complex patterns in large datasets. In the field of sentiment analysis, marked by informal language and rapid changes, knowledge was advanced by precisely measuring the impact on Precision, Recall, and F1-score. The research by Wankhade et al. [64] is consistent with the proposed model demand for improved performance criteria. The results outlined the importance of maintaining a balance between precision and recall ensuring the practical efficacy of sentiment analysis models in real-world scenarios. This research extended the previous investigation carried out by Mohamed Chiny et al. [65] using a sophisticated architecture of LSTM networks enhanced with FastText and multiple features. The enhancements were specifically designed for analyzing social media data regarding forest fires. The proposed models possessed distinct features enabling it to perceive the general sentiment, including identifying the nuanced changes essential in rapidly evolving situations.

The dynamic characteristics of social media, as determined by user interaction and data reliability, presented new challenges and opportunities for sentiment analysis. The diverse platforms provided an abundance of real-time, unorganised language, including rapid responses to events such as forest fires, it also had the potential to distort societies due to the biased demographics of users. The misrepresentation can reduce the efficacy of the proposed models, which heavily depended on the quantity and quality of linguistic training data, specifically Indonesian. The inaccuracies encountered during the process of normalising text for slang or idiomatic expressions in dataset had an impact on the accuracy of sentiment analysis. Meanwhile, the present research examined linguistic features in Indonesian, reporting that the application of the results to other languages or cultural contexts would require substantial modifications and adaptations. This was because the general relevance of the models were limited. Further research should be carried out to modify the proposed model for global deployment. The results showed that the various attributes of social media data played a crucial role in enhancing the effectiveness of the models. LSTM and advanced text processing techniques including FastText were used to effectively detect forest fires by capturing intricate details. Additionally, LSTM models were efficient at contextualising and prediction using complex, informal, and multifaceted language frequently encountered in social media data. These capabilities showed the effectiveness of the proposed model in providing early indicators and community responses, as well as insights usually lacked in more strict datasets.

VI. CONCLUSIONS

In conclusion, the main research objective focused on the adoption of relatively dynamic social media data to develop forest fire-related sentiment analysis using the LSTM model. The results showed that the model has superior performance compared to rule-based sentiment analysis including Vader and SentiWordNet when equipped with multiple layers. Additionally, it was reported that the incorporation of layers gradually improved the ability of the model to analyze complex and subtle linguistic data found in social media interactions.

The 4-layers LSTM model showed the most outstanding performance measures, thereby validating the ability to extract and analyze complex context on social media texts. Apart from exhibiting the ability of deep learning to offer more complex sentiment analysis compared to rule-based methods, this model attained outstanding degrees of accuracy and recall. Rule-based techniques, such as Vader and SentiWordNet, lacked the depth required for complex analysis, despite the strength and relatively balanced accuracy and recall. The current research improved the accuracy of sentiment analysis in the context of forest fire data by designing the model to suit the various linguistic components of Indonesian, using advanced feature extraction methods. These changes addressed how local idioms, slang, and culturally specific expressions were often ignored by traditional sentiment analysis systems.

Future research was anticipated to develop a hybrid model combining the effective computational capacity of rule-based systems with the strong analytical ability of LSTMs. This was aimed to obtain a combined impact of the speed and efficiency of lexicon-based analysis with the contextual sensitivity of deep learning to produce a scalable, accurate, efficient sentiment analysis tool. The hybrid model could completely change the knowledge and reaction of public opinion on important problems including wildfires. This was supposed to enhance community participation and emergency response plan methods. Therefore, the current research advanced the field of sentiment analysis by proving the practicality of state-of- arts LSTM models incorporated with real-world social media data. It also established the basis for future developments in integrating machine learning methods to properly address community requirements. The adoption of deep learning and rule-based systems helped in the production of instruments for evaluating public opinions in relation to disaster management.

Author Contributions: *Warih Maharani*: Conceptualization, Formal Analysis, Methodology, Investigation, Writing - Original Draft, Writing - Review & Editing. *Hanita Daud:* Data Curation, Funding Acquisition, Project Administration, and Recources. *Noryanti Muhammad*: Software, Supervision, Writing - Review & Editing. *Evizal Abdul Kadir:* Validation, Visualization, Writing - Review & Editing.

All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Telkom University under Joint Research Project (JRP) Grand (2023).

Conflicts of Interest: The authors declare no conflict of interest

Data Availability: The data will be made available for sharing upon request.

Institutional Review Board Statement: Not applicable.

Informed Consent: Informed Consent was obtained, and a detailed explanation was presented in the Methods section.

Animal Subjects: There were no animal subjects.

ORCID:
Warih Maharani: https://orcid.org/0000-0003-1574-921X
Hanita Daud: -
Noryanti Muhammad: https://orcid.org/0000-0002-6112-3720
Evizal Abdul Kadir: -

REFERENCES

[1] A. Khorram-Manesh, K. Goniewicz, A. Hertelendy, and M. Dulebenets, *Handbook of disaster and emergency management*. Kompendiet, 2021.

[2] M. P. Thompson et al., "Potential operational delineations: New horizons for proactive, risk-informed strategic land and fire management," Fire Ecol., vol. 18, no. 1, 2022, doi: 10.1186/s42408-022-00139-2.

[3] F.-N. Robinne and F. Secretariat, "Impacts of disasters on forests, in particular forest fires," *UNFFS Background paper*, 2021.

[4] J. A. Wang, J. T. Randerson, M. L. Goulden, C. A. Knight, and J. J. Battles, "Losses of Tree Cover in California Driven by Increasing Fire Disturbance and Climate Stress," *AGU Adv*., vol. 3, no. 4,2022, Art.no. e2021AV000654, doi: 10.1029/2021AV000654.

[5] J. MacCarthy, J. Richter, S. Tyukavina, M. Weisse, and N. Harris, "The latest data confirms: Forest fires are getting worse". Accessed: August 29, 2023. [Online]. Available: https://www. wri.org/insights/global-trends-forest-fires

[6] S. Thapa *et al.*, "Forest fire detection and monitoring," *Earth Observation Science and Applications for Risk Reduction and Enhanced Resilience in Hindu Kush Himalaya Region: A Decade of Experience from SERVIR*, pp. 147–167, 2021.

[7] A. Nurkholis, D. Alita, A. Sucipto, M. Chanafy, Z. Amalia, and others, "Hotspot classification for forest fire prediction using c5. 0 algorithm," in *2021 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, 2021, pp. 12–16.

[8] R.S. Priya, K. Vani, "Deep learning based forest fire classification and detection in satellite images," in *2019 11th International Conference on Advanced Computing (ICoAC)*, 2019, pp. 61–65, doi: 10.1109/ICoAC48765.2019.246817.

[9] Y. Kang, E. Jang, J. Im, and C. Kwon, "A deep learning model using geostationary satellite data for forest fire detection with reduced detection latency," *GIsci Remote Sens*, vol. 59, no. 1, pp. 2019–2035, 2022, doi: 10.1080/15481603.2022.2143872.

[10] B. T. Pham *et al.*, "Performance evaluation of machine learning methods for forest fire modeling and prediction," *Symmetry,* vol. 12, no. 6, p. 1022, 2020, doi: 10.3390/sym12061022.

[11] T. Preeti, S. Kanakaraddi, A. Beelagi, S. Malagi, and A. Sudi, "Forest fire prediction using machine learning techniques," in *2021 International Conference on Intelligent Technologies (CONIT)*, 2021, pp. 1–6, doi: 10.1109/CONIT51480.2021.9498448.

[12] X. Wu, G. Zhang, Z. Yang, S. Tan, Y. Yang, and Z. Pang, "Machine learning for predicting forest fire occurrence in Changsha: An innovative investigation into the introduction of a forest fuel factor," *Remote Sens* vol. 15, no. 17, p. 4208, 2023, doi: 10.3390/rs15174208.

[13] S. Yang, M. Lupascu, and K. S. Meel, "Predicting forest fire using remote sensing data and machine learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, pp. 14983–14990, doi: 10.1609/aaai.v35i17.17758.

[14] A. Singla and R. Agrawal, "Social media and disaster management: Investigating challenges and enablers," *Global Knowledge, Memory and Communication*, vol. 73, no. 1/2, pp. 100–122, 2024, doi: 10.1108/GKMC-12-2021-0206.

[15] H. Li, D. Caragea, C. Caragea, and N. Herndon, "Disaster response aided by tweet classification with a domain adaptation approach," *Journal of Contingencies and Crisis Management*, vol. 26, no. 1, pp. 16–27, 2018, doi: 10.1111/1468-5973.12194.

[16] P. M. Landwehr and K. M. Carley, "Social media in disaster relief," in *Data mining and knowledge discovery for big data*, Springer, 2014, pp. 225–257, doi: 10.1007/978-3-642-40837-3_7.

[17] R. I. Ogie, S. James, A. Moore, T. Dilworth, M. Amirghasemi, and J. Whittaker, "Social media use in disaster recovery: A systematic literature review," *International Journal of Disaster Risk Reduction*, vol. 70, p. 102783, 2022, doi: 10.1016/j.ijdrr.2022.102783.

[18] J. Lever, R. Arcucci, and J. Cai, "Social data assimilation of human sensor networks for wildfires," in *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments*, 2022, pp. 455–462, doi: 10.1145/3529190.3534735.

[19] J. Lever and R. Arcucci, "Towards social machine learning for natural disasters," in *International Conference on Computational Science*, 2022, pp. 756–769, doi: 10.1007/978-3-031-08757-8_62.

[20] J. Lever, S. Cheng, and R. Arcucci, "Human-sensors & physics aware machine learning for wildfire detection and nowcasting," in *International Conference on Computational Science*, 2023, pp. 422–429, doi: 10.1007/978-3-031-36027-5_3.

[21] J. Lever and R. Arcucci, "Sentimental wildfire: A social-physics machine learning model for wildfire nowcasting," *J Comput Soc Sci*, vol. 5, no. 2, pp. 1427–1465, 2022, doi: 10.1007/s42001-022-00174-8.

[22] C. Boulton, H. Shotton, and H. Williams, "Using social media to detect and locate wildfires," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2016, pp. 178–186, doi: 10.1609/icwsm.v10i2.14850.

[23] G. Zamarreño-Aramendia, F. J. Cristòfol, J. de-San-Eugenio-Vela, and X. Ginesta, "Social-media analysis for disaster prevention: Forest fire in Artenara and Valleseco, Canary Islands," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 6, no. 4, p. 169, 2020, doi: 10.3390/joitmc6040169.

[24] A. A. Fitriany, P. J. Flatau, K. Khoirunurrofik, and N. F. Riama, "Assessment on the use of meteorological and social media information for forest fire detection and prediction in Riau, Indonesia," *Sustainability*, vol. 13, no. 20, p. 11188, 2021, doi: 10.3390/su132011188.
[25] A. Rachman, B. H. Saharjo, and E. I. K. Putri, "Forest and land fire prevention strategies in the forest management Unit Kubu Raya, South Ketapang, and North Ketapang in West Kalimantan Province," 2020.
[26] BNPB, "Berbagi Informasi Real-time Situasi Darurat Melalui PetaBencana.id.". Accessed: August 29, 2023. [Online]. Available: https://www.bnpb.go.id/berita/berbagi-informasi-realtime-situasi-darurat-melalui-petabencana-id-
[27] S. Madichetty, "Classifying informative and non-informative tweets from the twitter by adapting image features during disaster," *Multimed Tools Appl*, vol. 79, no. 39, pp. 28901–28923, 2020, doi: 10.1007/s11042-020-09343-1.
[28] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
[29] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," *arXiv preprint arXiv:2011.00677*, 2020.
[30] S. M. Davey and A. Sarre, "the 2019/20 Black Summer bushfires," 2020, *Taylor & Francis,* doi: 10.1080/00049158.2020.1769899.
[31] P. Brando *et al.*, "Amazon wildfires: Scenes from a foreseeable disaster," *Flora*, vol. 268, p. 151609, 2020, doi: 10.1016/j.flora.2020.151609.
[32] G. Zittis *et al.*, "Climate change and weather extremes in the Eastern Mediterranean and Middle East," *Reviews of Geophysics*, vol. 60, no. 3, p. e2021RG000762, 2022, doi: 10.1029/2021RG000762.
[33] Forest Watch Indonesia, "Tentang hutan indonesia." Accessed: August 29, 2023. [Online]. Available: https://fwi.or.id/
[34] A. V Mavrodieva and R. Shaw, "Social media in disaster management," *Media and Disaster Risk Reduction: Advances, Challenges and Potentials*, pp. 55–73, 2021, doi: 10.1007/978-981-16-0285-6_4
[35] L. Yan and A. J. Pedraza-Martinez, "Social media for disaster management: Operational value of the social conversation," *Prod Oper Manag*, vol. 28, no. 10, pp. 2514–2532, 2019, doi: 10.1111/poms.13064.
[36] Z. Wang and X. Ye, "Social media analytics for natural disaster management," *International Journal of Geographical Information Science*, vol. 32, no. 1, pp. 49–72, 2018, doi: 10.1080/13658816.2017.1367003.
[37] W. Budiharto and M. Meiliana, "Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis," *J Big Data*, vol. 5, no. 1, pp. 1–10, 2018, doi: 10.1186/s40537-018-0164-1.
[38] T. Mustaqim, K. Umam, and M. A. Muslim, "Twitter text mining for sentiment analysis on government's response to forest fires with vader lexicon polarity detection and k-nearest neighbor algorithm," in *Journal of Physics: Conference Series*, 2020, p. 32024, doi: 10.1088/1742-6596/1567/3/032024.
[39] I. Irawanto, C. Widodo, A. Hasanah, P. A. D. Kusumah, K. Kusrini, and K. Kusnawi, "Sentiment analysis and classification of Forest Fires in Indonesia," *ILKOM Jurnal Ilmiah*, vol. 15, no. 1, pp. 175–185, 2023, doi: 10.33096/ilkom.v15i1.1337.175-185.
[40] B. S. Negara, R. Kurniawan, M. Z. A. Nazri, S. Abdullah, R. W. Saputra, and A. Ismanto, "Riau forest fire prediction using supervised machine learning," in *Journal of Physics: Conference Series*, 2020, p. 12002, doi: 10.1088/1742-6596/1566/1/012002.
[41] L. W. Evelina, M. Rizkiansyah, I. Prawira, R. E. Irawan, and others, "Comparative Forest Preservation Campaign Government and NGO on Indonesia Instagram," in *2020 International Conference on Information Management and Technology (ICIMTech)*, 2020, pp. 370–375, doi: 10.1109/ICIMTech50083.2020.9211190.
[42] S. Hasyim, R. Abdullah, and H. Ibrahim, "Forest damage and preservation through forest resources management in Indonesia," *GeoJournal*, vol. 86, pp. 2183–2189, 2021, doi: 10.1007/s10708-020-10177-5.
[43] M. Muslikhin, L. W. Evelina, M. Rizkiansyah, I. Prawira, and R. E. Irawan, "Environmental communication network on Indonesian forest fires on Twitter in pandemic COVID-19 era," in *IOP Conference Series: Earth and Environmental Science*, 2022, p. 12026, doi: 10.1088/1755-1315/951/1/012026.
[44] T. Nikonovas, A. Spessa, S. H. Doerr, G. D. Clay, and S. Mezbahuddin, "ProbFire: a probabilistic fire early warning system for Indonesia," *Natural Hazards and Earth System Sciences*, vol. 22, no. 2, pp. 303–322, 2022, doi: 10.5194/nhess-22-303-2022.
[45] M. Kibanov, G. Stumme, I. Amin, and J. G. Lee, "Mining social media to inform peatland fire and haze disaster management," *Soc Netw Anal Min*, vol. 7, pp. 1–19, 2017, doi: 10.1007/s13278-017-0446-1.
[46] V. Slavkovikj, S. Verstockt, S. Van Hoecke, and R. de Walle, "Review of wildfire detection using social media," *Fire Saf J*, vol. 68, pp. 109–118, 2014, doi: 10.1016/j.firesaf.2014.05.021.
[47] S. B. Kotsiantis, D. Kanellopoulos, and P. E. Pintelas, "Data preprocessing for supervised leaning," *Int J Comp Sci*, vol. 1, no. 2, pp. 111–117, 2006.
[48] L. Hickman, S. Thapa, L. Tay, M. Cao, and P. Srinivasan, "Text preprocessing for text mining in organizational research: Review and recommendations," *Organ Res Methods*, vol. 25, no. 1, pp. 114–146, 2022, doi: 10.1177/1094428120971683.
[49] A. Aizawa, "An information-theoretic perspective of tf-idf measures," *Inf Process Manag*, vol. 39, no. 1, pp. 45–65, 2003, doi: 10.1016/S0306-4573(02)00021-3.
[50] E. A. Rissola, D. E. Losada, and F. Crestani, "A survey of computational methods for online mental state assessment on social media," *ACM Trans Comput Healthc*, vol. 2, no. 2, pp. 1–31, 2021, doi: 10.1145/3437259.
[51] A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, and T. Mikolov, "Fasttext. zip: Compressing text classification models," *arXiv preprint arXiv:1612.03651*, 2016.
[52] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Comput*, vol. 12, no. 10, pp. 2451–2471, 2000, doi: 10.1162/089976600300015015.
[53] H. Abbasimehr, M. Shabani, and M. Yousefi, "An optimized model using LSTM network for demand forecasting," *Comput Ind Eng*, vol. 143, p. 106435, 2020, doi: 10.1016/j.cie.2020.106435.
[54] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.
[55] B. CM, "Pattern Recognition and Machine Learning,‖ Springer," 2006.
[56] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of the international AAAI conference on web and social media*, 2014, pp. 216–225, doi: 10.1609/icwsm.v8i1.14550.
[57] A. Esuli and F. Sebastiani, "Sentiwordnet: A publicly available lexical resource for opinion mining," in *Proceedings of LREC*, 2006, pp. 417–422.
[58] L. Khan, A. Amjad, K. M. Afaq, and H.-T. Chang, "Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media," *Applied Sciences*, vol. 12, no. 5, p. 2694, 2022, doi: 10.3390/app12052694.

[59] S. Madichetty and M. Sridevi, “A neural-based approach for detecting the situational information from Twitter during disaster,” *IEEE Trans Comput Soc Syst*, vol. 8, no. 4, pp. 870–880, 2021, doi: 10.1109/TCSS.2021.3064299.
[60] K. Barik and S. Misra, “Analysis of customer reviews with an improved VADER lexicon classifier,” *J Big Data*, vol. 11, no. 1, p. 10, 2024, doi: 10.1186/s40537-023-00861-x.
[61] P. Izsak, M. Berchansky, and O. Levy, “How to train BERT with an academic budget,” *arXiv preprint arXiv:2104.07705*, 2021.
[62] K. L. Tan, C. P. Lee, K. S. M. Anbananthen, and K. M. Lim, “RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network,” *IEEE Access*, vol. 10, pp. 21517–21525, 2022, doi: 10.1109/ACCESS.2022.3152828.
[63] L. Zhang, S. Wang, and B. Liu, “Deep learning for sentiment analysis: A survey,” *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 8, no. 4, p. e1253, 2018, doi: 10.1002/widm.1253.
[64] M. Wankhade, A. C. S. Rao, and C. Kulkarni, “A survey on sentiment analysis methods, applications, and challenges,” *Artif Intell Rev*, vol. 55, no. 7, pp. 5731–5780, 2022, doi: 10.1007/s10462-022-10144-1.
[65] M. Chiny, M. Chihab, O. Bencharef, and Y. Chihab, “LSTM, VADER and TF-IDF based hybrid sentiment analysis model,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, 2021, doi: 10.14569/IJACSA.2021.0120730.