Pattern Recognition in Disaster Response: Leveraging Machine Learning for Twitter Analysis

Md Impreeaj Hossain, Musaib Ibn Habib Mikdad, Ridwanul Haque, Farah Binta Haque, Ehsanur Rahman Rhythm, and Annajiat Alim Rasel
Department of Computer Science and Engineering (CSE)
School of Data and Sciences (SDS)
Brac University
KHA 224, Progati Sarani, Merul Badda, Dhaka - 1212, Bangladesh
{md.impreeaj.hossain, musaib.ibn.habib.mikdad, ridwanul.haque, farah.binta.haque, ehsanur.rahman.rhythm}@g.bracu.ac.bd, annajiat@gmail.com

***Abstract*—This study is groundbreaking because it provides a novel method to detect tweets about catastrophes by using modern machine learning techniques that include Long-Short-Term memory in addition to Natural Language Processing models (NLP). Each model is essential in the study of tweets from social media that react to disasters through Twitter Its primary function is to detect patterns within tweets and to identify the keywords or phrases that need to be spotted. They permit an efficient and thorough analysis of social media quickly and accurately - this is extremely useful during situations of crisis, when the decisions must be taken rapidly to protect the public. Integration between these models can lead to significant improvements to disaster response activities, while the framework showcases machines learning's power to gather valuable data from non well-structured, like social media. Frameworks for emergency response should provide broad-based options that meet demands of disaster management and introduce new concepts for the discussion group. The study shows that machine learning's capacity to increase the efficiency of disaster management techniques, and also open new research opportunities has significant impacts.**

***Index Terms*—Pattern Recognition, disaster management, social media analysis, LSTM, tweet classification, Keyword extraction, emergency response, real-time analysis.**

I. Introduction

Our study addresses the toughest challenge of applying machine learning techniques for Twitter data to provide aid in emergencies. Social media platforms provide a vital tool for disseminating important information during situations of emergency and their value for analysis has grown more important within the modern world. The primary focus is on developing sophisticated algorithms capable of sifting through tweets to identify essential information pertinent to disaster scenarios. At the core of this research are modern pattern recognition technologies, which form the basis of models designed to recognize specific patterns and words indicative of emergency situations. This enables a rapid and efficient response to unfolding crises. By utilizing Natural Language Processing (NLP) in tandem with machine learning, our study aims to provide insights into public sentiment during emergencies, significantly aiding in coordinated response efforts. Through our research and analysis we explore ways technology can improve the efficiency of disaster mitigation by providing emergency workers and decision-makers with accurate, real-time information required to respond during emergencies. Twitter data will help to in bridging the gap between aspects of social media networks such as Twitter and plans of action that emergency services have developed as well as aiding in the development of an disaster management strategy while increasing knowledge of the importance in such situations.

II. Literature review

The latest research studies focus on machine learning, as well as the use of natural process of language (NLP) methods for the analysis of social media posts during crises instances. Toraman et al. developed a specific software to analyze Turkish tweets that are posted during relief efforts. It assesses need for help through an interactive map while taking notes simultaneously [1]. Though this study was focused on Turkish tweets and did not take the seriousness that their posts carry into their importance, the impact of these tweets should be not understated that social media plays a vital part in responding to natural disasters, as well helping rescue operations during emergencies. Zhou et al. employed BERT to identify rescue requests on platforms like Twitter and Facebook during disasters, with their BERT-based model achieving a 91% accuracy rate [2]. This offers valuable insights for effective disaster response via social media.

Maharani along with al. performed their study using BERT to analyze flooding sentiment analyses of the tweets published on Twitter and demonstrated its capacity to improve awareness of natural catastrophes, while providing insight into the reaction of the public following these disasters. [3]. Meanwhile, Maulana and Maharani proposed a novel method for classifying disaster tweets by combining geospatial data with BERT MLP, achieving a high accuracy of 90.7% when tested on tweets related to the 2023 Beirut explosion [4]. Wang et al. presented a method for extracting consumer insights from tweets during public health crises using the BERT model, which was successful in achieving an accuracy rate of 87.3%

on a COVID-19 pandemic tweet dataset from four American towns called [5]. These techniques show guarantee as important resources for associations looking to really comprehend and answer emergencies more.

Balakrishnan et al. also conducted extensive research on transformer-based algorithms for disaster tweet classification, evaluating a variety of models and determining the factors that influence them [6]. While catchphrase based approaches give straightforward yet mistaken grouping, AI approaches like BERT conquer constraints of watchword based techniques by precisely classifying catastrophe tweets, as exhibited in Pranay Kumar's concentrate on disdain discourse examination in calamity tweets utilizing BERT [7]. Furthermore, Saha et al. fostered a BERT and Container Organization mix model for tweet classification, helping fiasco alleviation associations in perceiving and examining calamity related tweets [8].

Significant progress has been made in the areas of public health and intimate partner violence (IPV). Al Garadi et al. used NLP to perceive IPV-related tweets, accomplishing a 88% exactness rate with a BERT-based model [9]. This exploration highlights the capability of NLP in helping IPV acknowledgment. Likewise, Le et al. fostered a BERT-based technique to characterize debacle tweets with high precision, displaying the flexibility of BERT in different catastrophe related errands [10].

Shan et al.'s work in real-time disaster damage estimation proposed a creative strategy utilizing cell phone-gathered virtual entertainment information, accomplishing high exactness in arranging catastrophe related message and evaluating fiasco damage [11]. This technique offers convenient and precise debacle harm appraisals, developing customary methodologies. Prasad et al. enhanced BERT for identifying and categorizing transportation disaster tweets, achieving over 95% accuracy in identifying and classifying these tweets into relevant categories [12].

Lastly, Dharma and Winarko's method for classifying natural disaster-related tweets using CNN and BERT embeddings [13] , and Huang et al.'s novel text clustering method for early detection of emergencies from social media, both achieved over 89% accuracy [14]. These approaches mark significant improvements in disaster tweet classification and emergency detection. Bello et al. presented a BERT-based framework for tweet sentiment analysis with 87.3% accuracy, contributing to natural language processing and sentiment analysis, albeit with limitations in dataset reliance and generalizability [15].

III. Dataset

The enhanced dataset offers a detailed snapshot of Twitter dynamics, now featuring a balanced compilation of 25,000 tweets. Each tweet is meticulously categorized as either 'disaster-related' or 'non-disaster related'. Originating from the renowned data-sharing platform Kaggle, this repository includes tweets from a variety of global locations, including the United States, providing a comprehensive view of the linguistic patterns associated with natural disasters. Flood-related terms feature prominently, signalling prevalent themes.

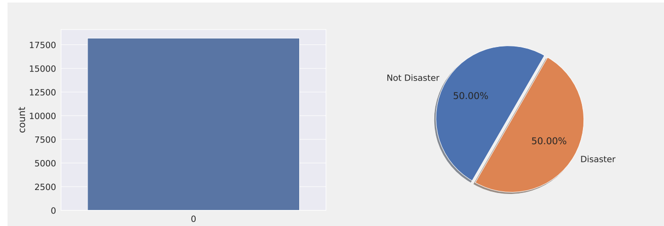

Fig. 1. Distribution of tweets

This expansion expands and diversifies our dataset's geographic and thematic coverage. Every tweet was classified using crowd sourcing classification to ensure accurate analysis and relevance for analysis purposes. Importantly, this symmetry makes the gathering more suitable to cutting-edge NLP models like LSTM (Long Short-Term Memory). Accordingly, this dataset serves as an excellent starting point for developing and testing disaster response models on social media data , underscoring its value in disaster assistance programs as well as research programs. Scholars and technologists will find this resource especially invaluable as its extensive yet compact material may validate future and past accomplishments in social media analysis as well as enhance disaster response models.

TABLE I
Data Overview

id	keyword	location	text	target
0	ablaze	NaN	Communal violence in Bhainsa, Telangana. "Ston...	1
1	ablaze	NaN	Telangana: Section 144 has been imposed in Bha...	1
2	ablaze	New York City	Arsonist sets cars ablaze at dealership https...	1
3	ablaze	Morgantown, WV	Arsonist sets cars ablaze at dealership https...	1
5	ablaze	OC	If this child was Chinese, this tweet would ha...	0

IV. Methodology

Data collection is the initial step in the process. Twitter's API includes a complete toolkit that can help you collect tweets. Tweets with particular keywords or hashtags linked to natural disasters (e.g. floods, earthquakes) or data on location in the areas that are affected by natural disasters are gathered. Once the data is gathered, information needs to be thoroughly cleaned and reprocessing, including the elimination for URLs, punctuation marks and tweets that contain other than English tags. In addition, tokenization can be used for the purpose of separating tweets into phrases, or tokens.

Labeling the process is essential. To make sure that the effectiveness of supervised learning can be assured, tweets need to be categorized manually as being connected or not to natural catastrophes. It's not a simple process, it's essential for the efficacy of training in modeling.The entire dataset is split into two sets: the training set, 82%, as well as an experimental collection of 18%. Training sets help in modeling of models. The test set may be utilized to evaluate their performance.

The LSTM model, distinguished by its capacity to process sequences of data over long intervals and store this information efficiently is at the core of this study. In addition to unigram, bigram and trigram models to provide a broader contextual picture of text The LSTM framework has a knack for discovering complex linguistic patterns.

In the course of refinement The LSTM's performance will be assessed by measuring epochs and losses, as well as accuracy measures giving a complete knowledge of its ability to recognize content that is related to disasters. When fine-tuning this model, it can be used to determine the real-time status of tweets. It provides instant information for disaster management. Incorporating real-time tweet data allows for continuous calibrating of the model, which ensures its efficiency with respect to changing use of language in situations involving disasters.

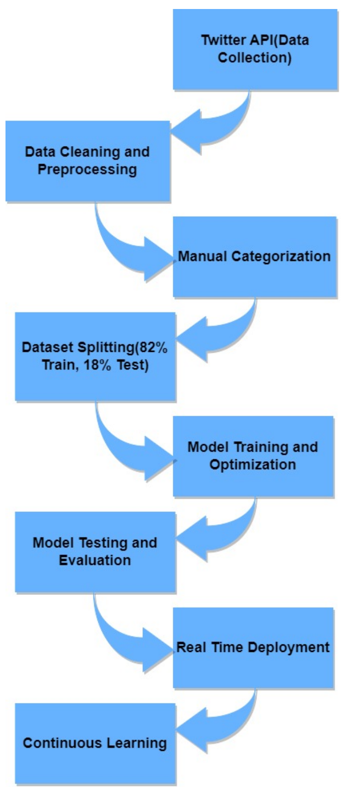

Fig. 2. Model Diagram

V. Result and Analysis

In order to analyze text, the auxiliary methods are utilized to clear and process the information. They're crucial in eliminating URLs as well as and punctuation marks from text. The text is changed to lowercase in the initial process to use NLP. Text tokenization is a crucial process of machine learning. It breaks text down into smaller parts for better processing.

The pie chart can be designed to visualise the data distribution when a disaster occurs. The figure 2 illustrates the fact that a majority of tweets concern the catastrophe and the other 50% are on different topics. Twitter has facilitated communication and interaction between people that were affected by the tragedy. Tweets about the catastrophe comprise the largest proportion of Twitter tweets. Further insights can be gleaned from the histograms in Figure 3 that shows the the distribution of characters in tweets that are not related to disasters. The average length of tweets is 130 characters. With more than 100 tweets in this category. Certain tweets go over the 220 character limit. The CDF illustrates that 75% of tweets contain only 150 words or less. In order to improve the accuracy of smaller numbers of tweets, an approximate cumulative distribution function (ECDF) is utilized and relies on the actual evidence instead of theoretical assumptions.

Tweets that are not related to disasters tend to be brief messages or reflections. However, the tweets for disasters tend to be longer and give information or support. Figure 4 illustrates the distribution of characters of tweets relating to disasters. The average tweet is approximately 130 characters in length most of them falling between this and 140 characters. Further, analysis of CDF indicates that about 75% of all Tweets related to disasters have $\leq$ 150 characters. The ECDF is based on data from empirical studies is in agreement with these data.

Figure 5 shows a distinct contrast in the number of words between the tweets that are related to disaster and those not. The tweets related to disasters are usually smaller (10-20 words) in comparison to non-disaster related tweets and are generally larger (20-30 word). Though short tweets may be present in both types but they're not as common when compared with longer tweets. Twitters that are about emergencies tend to have short updates while non-disaster tweets offer more in depth data. Emotions drive shorter words in disaster tweets. The word count median for tweets about disaster is 15 and for non-disaster tweets the word count is 25. The word count in tweets that are not disaster-related vary in length, some being as little as a few words.

On the other hand, non-disaster Tweets generally convey more informative and descriptive information without any the occasional outliers. Figure 6 provides a visual contrast between the disaster and non-disaster tweet types. Non-disaster tweets average a word-length of 10.03 and the tweets related to disaster contain an average word count of 7.98. Tweets that are not related to disaster typically have larger words in comparison to related tweets to disasters. The word count in non-disaster tweets is around 10 for disaster tweets, whereas in non-disaster tweets it's about seven. The non-disaster tweets usually have the minimum number of words. Disaster-related tweets typically contain at minimum eight words. The IQR for tweets that are not related to disasters is 3. For the tweets relating to disasters, it's 2. This suggests a greater variance of word lengths for non-disaster related content. Certain non-disaster content is notable due to their usage of lengthy phrases (15-20 words) that often convey specific and useful information.

Twitter is an active social network that is characterized by instantaneous interactions as well as a vast spectrum of human emotions. They range from needy messages in times of crisis to serene narratives of the daily life. The structures that are linguistic in the content of the platform are illustrated in Figure 7highlighting their variety. Numerous unigrams, such as “fire,” “earthquake,” and “flood” emphasise the urgency of the unfolding crisis. Expressions such as “love,” “friend,” and “watch” reflect the joy of making connections as well as nurturing relationships and having fun. The variety of language of Twitter displays a variety of emotions posted online.

For a deeper study of Figure 8 the large bigrams highlight this difference. The events such as “earthquake impacts” and “flood warnings” bring images of chaos and disruption which highlight the urgent demand for a collective response. However, contrasts like “having fun” or “watching film” give insight into everyday life and capture occasions of joy and bonding.The story’s conclusion is depicted in Figure 9, which reveals the most frequently used trigrams. The phrases like “Tsunami warnings issued” or “Firefighters battling fire” convey the gravity of a disaster and the response of communities. Trigrams like “Looking towards” express enthusiasm for anticipation and excitement, as well as the immense complexity of everyday interactions. Figure 10 gives a thorough depiction of the frequency of named entities within the data. The study highlights the most popular subjects themes, important points of discussion in media. This analysis provides insight into the main discussion areas as well as influential people.

To classify tweets related to disasters using the model LSTM. The process of training consisted of 10 epochs with the use of 455 batches for each epoch. The method of batching was developed to optimize efficiency and speed of processing which resulted in a fast learning process. In the first epoch, the model performed well, and a loss in training of 0.34. This early performance confirms the validity of predictions made in the beginning and establishes a positive path to future epochs.

It is evident that the LSTM model’s outputs across 10 epochs show a learning pattern. In the beginning, the loss is reduced by 0.3446 to 0.0142 while the accuracy increases by 85.17% to 99.62%, which suggests that the model is able to learn. The loss in validation rises by 0.2439 to 0.5330 and suggests that the model is overfitting since the model is too tightly adapted to the data it is trained on. The validation accuracy proximates to its peak level at 91.95% in the third epoch, but by the tenth, it gets down to 89.39%. This might allude that as training goes on, the model’s potentiality to adjust to new data becomes less and less. This implies that in order to maintain good performance in the presence of unseen data, regularization or a sudden halt to the LSTM model’s training is required.

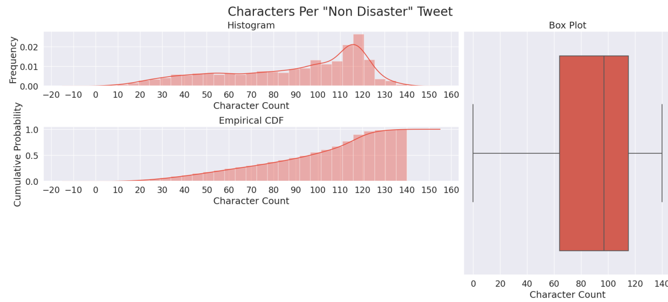

Fig. 3. Characters for each Non Disaster tweet

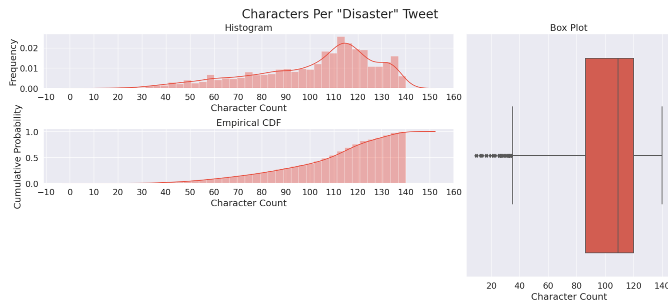

Fig. 4. Characters for each Disaster tweet

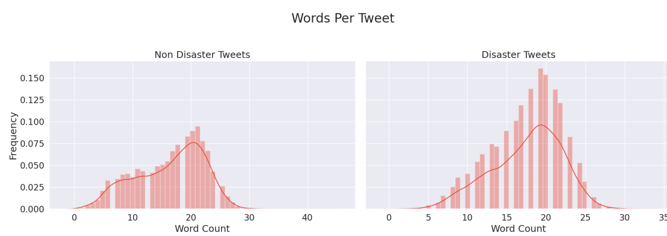

Fig. 5. Words for each Tweet

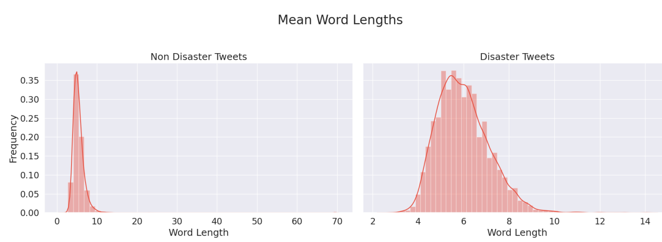

Fig. 6. Word lengths Mean

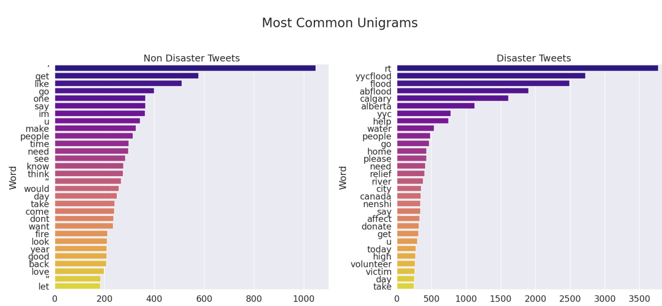

Fig. 7. Most Common Unigrams

TABLE II
MODEL TRAINING AND VALIDATION METRICS OVER EPOCHS

Epoch	Loss	Accuracy	Validation Loss	Validation Accuracy
1	0.3446	85.17%	0.2439	91.76%
2	0.1364	95.61%	0.2589	91.70%
3	0.0713	98.08%	0.2315	91.95%
4	0.0488	98.70%	0.3071	91.32%
5	0.0313	99.18%	0.4394	90.66%
6	0.0181	99.55%	0.3301	89.94%
7	0.0163	99.56%	0.4318	87.14%
8	0.0178	99.51%	0.4261	89.78%
9	0.0180	99.53%	0.4283	89.78%
10	0.0142	99.62%	0.5330	89.39%

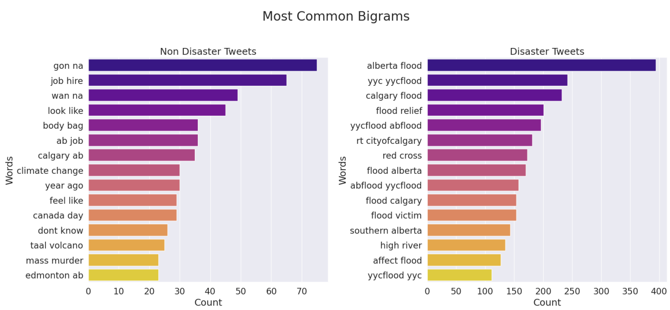

Fig. 8. Most Common Bigrams

Analyzing and evaluating the results and then analyzing these results is a crucial aspect of planning for disasters and the analysis of social media. This article focuses on the effectiveness and drawbacks in using machine learning methods in real-time disaster management apps and patterns recognition programs to analyze social media. This piece is intended to finish this study in that it provides an explanation about the conclusions

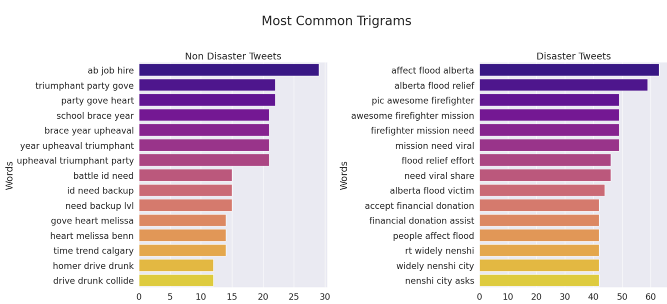

Fig. 9. Most Common Trigrams

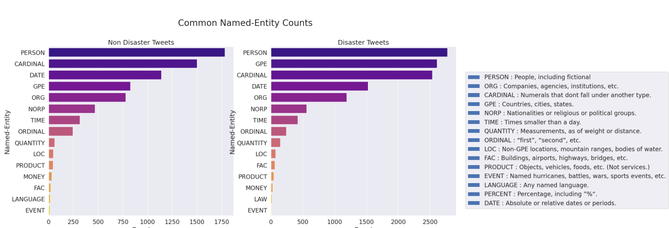

Fig. 10. Common Named-Entity Count

from this research and assisting in the advancement of this field.

The research conducted by our team has revealed that one of the major problems with the LSTM models is its inability to accurately fit and also fails to verify and validate that the training needs are met. The initial results revealed an enthralling learning curve with an accuracy range between 85.17% to 99.62% in a variety of test conditions. Loss reductions dramatically decreased from 0.3446 and dropped to 0.0.0142 once the training was finished. The preliminary findings suggest the possibility that data derived from sources other than traditional could be successfully utilized in the course of learning because it was able to adjust according to the demands of trainees. It is in contrast to models that have been fitted too much which do not correspond to the information they are using in training, and they change as the new data, which isn't validated enough comes in. This is often shown by an increase in validation loss. Further evidence of the model's decreased aptitude to adjust to new data comes from the fact that its accuracy on validation data peaks at 91.95% in the third epoch and then falls to 89.39% by the tenth epoch.

To address the shortcomings, our research recommends the inclusion of regularization, or pausing features in the learning portion that is part of an LSTM modeling. Regularization is a strategy to reduce the negative effects of overfitting in part, by penalizing models that comprise several layers. In this way and allowing regularization, it will aid the model to learn universal patterns, and will be more efficient with different types of data. Based on the validation results along with models' performance, the results of testing suggest that the earlier stopping mechanism might prevent the model from reaching its peak and may be terminated in a short time. This is essential in ensuring the model's quality and accuracy as well as adaptability to dealing with unknowable information as well as a possible improvements to be made regarding the training of models and the validation process in improved machine learning algorithms designed for situations involving disasters.

VI. CONCLUSION

In the course of its study of tweets that are related to disasters the research has proposed an ingenious hybrid system that combines the LSTM as well as NLP methods to classify tweets effectively and to categorize tweets about disasters more accurately. This system ensures effective recognition of keywords and its primary goal is in identifying patterns that are present within tweets. Categorization aids emergency decision-makers to make informed decisions faster, while also speeding up the process of analysis substantially. The collaboration of these models is an important step forward in emergency management. The model demonstrates the ways machine learning can be utilized to get valuable insights from social media in situations where millions of tweets can flood into the system via sites such as Twitter. It can at first be difficult to discern the flaws of a system like its propensity to be over-fitted during long time periods of training, or the inadequate flexibility when dealing using new data sources. The use of regularisation techniques or early stopping mechanisms can help and provide disaster professionals better methods.

FUTURE WORK

Machine learning strategies can assist in emergency situations by employing a variety of efficient models, for example the GP-3 text-creation technology is much more advanced and provides more complete information analysis and data mining of social media. RoBERTa enhances comprehension of language, while the approach to autoregressive learning used in the XLNet model detects two-way contexts with precision for intricate analysis of text; and T5 is able to effortlessly manage tasks as well as data inputs through its text-totext technique, which makes these models the most advanced technology in the field of natural language processing which could improve effectiveness and precision for emergency response.

REFERENCES

[1] C. Toraman, I. E. Kucukkaya, O. Ozcelik, and U. Sahin, "Tweets under the rubble: Detection of messages calling for help in earthquake disaster," *International Journal of Disaster Prevention and Research*, vol. 11, no. 3, pp. 285–297, 2023.

[2] B. Zhou, L. Zou, A. Mostafavi, B. Lin, M. Yang, N. G. Gharaibeh, H. Cai, M. J. Abedin, and D. Mandal, "Victimfinder: Harvesting rescue requests in disaster response from social media with bert," *Computers, Environment and Urban Systems*, vol. 95, no. 1, p. 101824, 2022.

[3] W. Maharani, "Sentiment analysis during jakarta flood for emergency responses and situational awareness in disaster management using bert," *IEEE Conference Publication*, June 2022.

[4] I. Maulana and W. Maharani, "Disaster tweet classification based on geospatial data using the bert-mlp method," *IEEE Conference Publication*, August 2022.

[5] Y. Wang, E. Willis, V. K. Yeruva, D. Ho, and Y. Lee, "A case study of using natural language processing to extract consumer insights from tweets in american cities for public health crises," *BMC Public Health*, vol. 23, no. 4, 2023.

[6] V. Balakrishnan, Z. Shi, C. L. Law, R. Lim, L. P. Teh, Y. Fan, and J. Periasamy, "A comprehensive analysis of transformer-deep neural network models in twitter disaster detection," *Mathematics*, vol. 10, no. 24, p. 4664, 2022.

[7] M. S. B. V. Pranay Kumar, "View of hate speech and reality check analysis of disaster tweets using bert deep learning model," 2023.

[8] S. S. Tulika Saha, Srivatsa Ramesh Jayashree, "Bert-caps: A transformer-based capsule network for tweet act classification," *IEEE Journals & Magazine*, October 2022.

[9] M. A. Al-Garadi, S.-J. Kim, Y. Guo, E. Warren, Y. Yang, S. Lakamana, and A. Sarker, "Natural language model for automatic identification of intimate partner violence reports from twitter," *Array*, vol. 15, p. 100217, 2022.

[10] A. D. Le, "Disaster tweets classification using bert-based language model," *ResearchGate*, 2022.

[11] S. Shan, F. Zhao, Y. Wei, and M. Liu, "Disaster management 2.0: A real-time disaster damage assessment model based on mobile social media data—a case study of weibo (chinese twitter)," *Safety Science*, vol. 115, pp. 393–413, 2022.

[12] R. Prasad, A. U. Udeme, S. Misra, and H. Bisallah, "Identification and classification of transportation disaster tweets using improved bidirectional encoder representations from transformers," *ScienceDirect*, vol. 3, no. 1, p. 100154, 2023.

[13] L. S. A. Dharma and E. Winarko, "Classifying natural disaster tweet using a convolutional neural network and bert embedding," *IEEE Access*, vol. 10, no. 1, pp. 755–764, 2022.

[14] L. Huang, P. Shi, H. Zhu, and T. Chen, "Early detection of emergency events from social media: a new text clustering approach," *Natural Hazards*, vol. 111, no. 1, pp. 851–875, 2022.

[15] A. Bello, S. Ng, and M. Leung, "A bert framework to sentiment analysis of tweets," *Sensors*, vol. 23, no. 1, p. 506, 2023.