Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

A Classification of Disaster Responses Based on an Analysis of Data from Social Media

Muhabbat Hussain (Corresponding Author)
FAST National University of Emerging Science and Technology, Karachi, Pakistan. Email: muhabbathussain777@gmail.com

Hussain Akbar
School of Mechatronics Engineering and Automation, Shanghai University, Shanghai, 200444, China. Email: Hussainakbar@shu.edu.cn

Fateh Aman
Department of Computer science & information technology, Superior university main campus Sargodha (40100), Punjab, Pakistan.
Email: amanmalik967@gmail.com

Hafiz Muhammad Sarmad Abid
Department of Project and Operations Management, The Islamia University of Bahawalpur, Punjab, Pakistan. Email: Sarmad6868@gmail.com

Shakeel Ahmed
Bs Hons Final Year Library Information science and archive studies, University of Sindh, Jamshoro. Email: shakeel.ahmeduos@gmail.com

Masooma Soomro
Bs Hons Final Year Library Information science and archive studies, University of Sindh, Jamshoro. Email: masoomasoomro05@gmail.com

Abstract
During any natural disaster, a lot of information is created on social media; users produce a lot of information, such as Twitter, to post textual and multimedia content to report updates about injured or dead/missing people, needs and other information types. However, in the past, research in this field has not had much data available; for this purpose, we will require entirely different approaches, tools, and techniques to help inform decision-making under uncertain conditions. The fundamental target of our research in this field is to improve disaster relief efficiency and attention and extract useful information from social media data, like public attitude toward disaster response and the public demands for the targeted based on properties such as needs, damages, etc. In this study, public perception is assessed qualitatively by manually classifying, which contains lots of information like demand for target relief supplies, satisfaction with the disaster response, and public fairness. So, by using public tweets, are analyzed using different machine-learning models. And to better provide the decision maker with the appropriate model, the comparison of different machine learning models based on computational time and prediction accuracy is conducted.
Keywords: Natural Disaster, Social Media, Twitter, Machine Learning, Disaster Relief, Public Perception.

Introduction

Social media platforms, particularly Twitter, have become increasingly important for real-time information sharing during disasters and emergencies. User-generated content, including text, images, and videos, can significantly enhance situational awareness for emergency responders and other stakeholders, enabling more informed decision-making. However, the sheer volume of social media communication during a crisis often includes substantial noise and irrelevant information. It is crucial to develop methods for efficiently filtering and extracting actionable information from this ever-growing data stream. This necessitates categorizing messages into meaningful, high-level classifications to identify relevant and actionable data. Given the overwhelming data volume, manually analyzing each tweet in real-time is impractical for emergency managers and other stakeholders.

This study explores the use of various artificial intelligence techniques, such as information extraction and classification, to automate this process. During a natural disaster, numerous tweets are posted, but determining which tweets contain crucial information and user intentions is challenging. The rapid growth of social media users and the corresponding increase in disaster-related tweets present both a challenge and an opportunity. Many of these tweets can be invaluable to disaster management teams, enabling them to provide timely and targeted assistance to affected populations.

By analyzing social media posts, we can identify critical tweets, verify their relevance to the disaster, and extract key information such as location, expressed needs, and required items. This involves identifying location details (country, city, region) and searching for keywords indicating specific needs or requests, such as "need," "needed," "requiring," "urgent," "emergency," and "important." Furthermore, extracting information about necessary items like tents, food, water, medicine, and hygiene supplies can aid in coordinating relief efforts.

This analysis will enable us to provide disaster management teams with actionable information, including the precise location of the disaster, the specific needs of those affected, and the types and quantities of emergency supplies required. This information can significantly improve disaster response, particularly in regions like Pakistan, which currently lack a robust and efficient disaster management system. By leveraging social media data, we aim to bridge this gap and facilitate a more timely and effective response to emergency situations.

After analyzing the tweets, we can find what is required of disaster-affected people. How we can help them based on places and requirements in a short time.

After analyzing the tweets, we can find what is required of disaster-affected people. How we can help them based on places and requirements in a short time

Literature Review

Such countless works they are comparative with at present working, yet they have in an unexpected way. We have reviewed lots of literature, but some are mentioned here.

Real-time Disaster Mitigation: focuses on accelerating disaster mitigation efforts by mining and analyzing social media data. This includes examining public attitudes towards disaster response and identifying community needs for targeted relief supplies during various types of disasters. The study primarily focuses on high-impact events, analyzing properties such as duration and damage assessment based on a dataset of 41,993 tweets [1].

Political Sentiment Analysis: explores mining tweets to extract political sentiments and model them as a supervised learning problem. The study analyzes tweets related to the 2019 Indian General Elections, examining public sentiment towards major political parties. A predictive model, trained on sentiment data, is used to forecast election outcomes. Long Short-Term Memory networks are employed for representation modeling, and their performance is compared with traditional machine learning models [2].

Multimodal Disaster Response: presents a comprehensive analysis of textual and multimedia content from a large volume of tweets posted during three disaster events. Using various AI techniques from NLP (Natural Language Processing) and CV (Computer Vision), the study leverages diverse computational approaches to process disaster-related data. The research highlights the identification of various types of actionable information to inform emergency managers and responders and discusses the development of future automated disaster management systems [3].

Multimodal Dataset for Crisis Response: addresses the need for comprehensive datasets by presenting a large multimodal dataset collected from Twitter during various disaster events. The study provides three types of annotations useful for various crisis response and management efforts by different humanitarian organizations [4].

Big Data and Disaster Management: emphasizes the role of advancements in computing, including mobile devices, internet access, social media, and big data analytics, in transforming data exchange during disasters. This supports disaster managers with data-driven responses to disaster management challenges. Big data enables risk assessment through infrastructure data and sensor data, provides insights into affected populations through smartphone and social media data, and helps establish objective functions for local emergency response. Furthermore, big data offers continuous streams of on-site disaster information through data mining. By analyzing continuous disaster information, assessments create dynamic information loops on disaster events, assisting managers in developing real-time, accurate, and proactive rescue strategies. Big data can contribute to all phases of disaster management: prevention, preparedness, response, and recovery, ultimately improving a city's resilience to disasters [5].

A unique catastrophe response system known as HAC-ER is proposed. HAC-ER connects individuals, trained professionals, automated systems, and programmed responses in cooperative associations to enhance their individual and collective capabilities. The development of HAC-ER involved end-users, including professionals and volunteers, in participatory design workshops, practical training, and field tests of progressively advanced models of individual components and the complete system. This collaborative approach yielded both quantitative and qualitative results, as well as new research questions. HAC-ER demonstrates how human-agent collectives can address key challenges in disaster response. Specifically, it utilizes crowdsourcing combined with AI to gain comprehensive situational awareness from large volumes of reports posted by the public and trusted organizations. This information can then inform human-expert teams in organizing multi-UAV operations and task allocation plans for ground responders. Finally, HAC-ER incorporates a mechanism for tracking and

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

utilizing the attribution of data shared across the system to ensure accountability. All components of HAC-ER are openly supported, and their performance is evaluated against standard (non-HAC) baselines to demonstrate the value of the complete structure [6].

Studies using AIDR have shown its potential in processing social media messages during disasters. These messages offer real-time or low-latency situational awareness, enabling more effective crisis response. Different emergency response organizations need different types of information. For example, infrastructure damage reports should go to specific agencies, while water and food availability reports should be directed elsewhere. Disaster response strategies vary across different phases of disaster preparedness, response, and recovery, each with unique data requirements [7].

In the wake of investigating their previous work, we appreciate that in internet-based media, consistent response to disaster follows a comparable model; in other words, messages posted through electronic media during the early phases of a calamity talk about the readiness and notification ahead of time, while messages posted during the later stages report system mischief, gift, and misfortunes expected to report or available, etc. [9].

In another paper (Artificial Intelligence for Disaster Response) AIDR was used to channel the Twitter stream where the client can screen the grouping of the status like total dealt with the thing, the time sneaked past, etc. and a short time later finally, a consequence of messages organized into the classes is created, and after these are used to different crisis maps on the other sort of reports. For example, the buyer application is the current variation of Crisis Tracker, which is used AIDR to engage to play hooky of income, which change by sending circumstance to consolidate for instance spectator accounts, reports of damage establishment, or reports of mercilessness [7].

In the [11] paper, they apply social media intelligence for disaster response and management in smart cities. They aim to develop a cloud based on the big data framework; also, they were able to get different types of data from different sources. And apply machine learning techniques to collect and apply the process and get some results for the response for support of the different types of any emergency response and also focus on quickly. When they have collected the right data and further investigated the data sources and checked the data, it is suitable for their techniques where the data are working perfectly and well-extracted from the data representation. The main goal of their work is to display valuable information to the decision-maker for any emergency response that helped them, and finally, they want some contribution to research and facilitating comprehensive disaster management of their framework. And they will have helped us with any emergency response operation for smart cities.

In [12], the first increase in users on social media platforms, and they generate a huge amount of data in the form of the unstructured way. Like testing the form of messages, blogs, chats, and posts. The exchange of information on social media is one of the remarkably convenient mediums where it's easy to express opinions and ideas, and if a large number of users like it, then it gains popularity. Twitter is one of the popular social media platforms, and they have contained a large amount of data. So, in this paper, they will analyze people's predictions and opinions and predict the model, It's a supervised learning problem.

In [14], this research focuses on how to support disaster response. For this

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

purpose, they have proposed a mathematical programming approach in real time with the related information, and by using this real-time data, they provide how to optimize the post-disaster decision for any emergency response. So, this decision can support a tool chat that can provide effective and rapid solutions, and such a type of support is essential for disaster.

In [15], this paper focuses on the disaster situation of the affected family and, due to this disaster, the separation of the family and its default to tracking refugees. For this purpose, they have designed a system for tracking refugees, and this system is a bigger part of the system for more comprehensive disaster response. By using this system, they identify the personal ID and their location based on the record. After finding out the ID and location, they search for the family and related group for better treatment.

In [16], this paper analyzes the uploaded pictures after the disaster occurs and some issues that occur and the difficulty to find out the exit location on the map. There are two main reasons for such a type of difficulty: one is the picture in a different direction in terms of different heights of the image.

In [17], this paper focuses on big disasters, both natural disasters and man-made disasters, for example, the Wenchuan earthquake and Fukushima nuclear disaster, and such types of large disasters make it impossible to dispatch the relevant person to search for or immediately take any action. So, for this purpose, they have proposed the use of the architecture of the Internet of Things (IoT) (artificial intelligence + Internet of Things) to cooperate with the surface of the ground and underwater robots to apply to such types of disasters. For this issue in training, they use a deep learning model and different model verification to train in the deep learning model and transmit it into the Internet of Things and transmit it to the ground, and they apply it on robots and detect continuous object classification and verification and take a good decision for the emergency response.

In [19] this paper, they design and develop a crowdsourcing mobile application for any emergency disaster response on the basis of different types of requirements, user anonymity and friendliness, viewing of geographic information, and bi-directional and real-time updating. And this app is very useful for disseminating quality and timely geographic information during any disaster emergency.

To achieve this aim, they used natural language processing and also used machine learning and responses by taking the first post to the right person that may have helped us with issues posting. And the disaster volunteers of the management are trusted to show that according to achieving the results.

Now, in the below section, we have compared different types of papers, what they have done, what their problem statement is, what they have proposed as solutions, and also checked the gap or future work if they have mentioned it.

In [31] this paper they have worked on crisis-related data from Twitter by using contextual representations. As we know, in the past few years, working on neural-based representations like word embedding has frequently been used for natural language processing for different types of NLP tasks. As we know, that word embedding plays a significant or important role in different types of deep learning natural language processing (NLP) tasks.

Therefore, in this paper, they have found the best word embedding used for a particularly important and specific task. A dense classifier with contextual representations on embedding from the language model is proposed, and such a type of embedding model is used for crisis-related data on social media during any emergency disaster or any other natural disaster. They have to work in real-time from Twitter datasets; examples include the California Earthquake, Nepal Earthquake, and Typhoon Hagupit, and they have tried to perform and analyze with different parameters such as recall, precision, F1 score, and accuracy.

So, in this paper, they have proposed the dense classifier with ELMo embedding models to give better accuracy as compared to the deep learning classifier (Convolutional Neural Network and Multilayer Perceptron Convolutional Neural Network with Crisis word embedding) and other traditional classifiers (Support Vector Machine).

In [32], this paper discusses their work on informative tweet classification of the earthquake disaster situation in Indonesia. In this paper, they have explained that Twitter is one of the online media that produces a large number of important data on significant disaster-related data tweets. And these tweets contain information to take humanitarian assistance measures. In this paper, their main purpose is to train such a type of model where the model is based on the information sourced from Twitter and classifying different types of tweets and checking the tweets that are informative and which are non-informative by using an algorithm for classification.

They explain that they work according to the previous research algorithm that has been considered appropriate in dealing with problems: Support Vector Machine algorithms. Based on that research, they develop research models by adding features of the Smote Up sampling imbalance and the Gini Index and also adding the Naive Bayes algorithms and comparing them with the accuracy of the classification algorithm. The data that are used from a tweet are related to Indonesia and have been collected using the Rapid Miner application and also by the use of GataFramework in the text processing. After applying this proposed method to the Support Vector Machine algorithm, it produces 81.03% accuracy and is superior to the Naïve Bayes Algorithm, which produced an accuracy of 80.30%. Based on the result of the accuracy, both enter into a good classification algorithm, resulting in accuracy.

In [33], this paper they have also worked on real-time earthquake detection using Twitter tweets. For this purpose, first, they have explained that, as we know, nowadays social media networks have become a part of daily life due to the rapid development of new and advanced technology and also the usage of such types of networking sites.

And people come to know about the latest news alert related to natural calamities very quickly. But the authenticity of such a type of news needs to take care of developing a physical type of sensor system in the particular residential area for the earthquake to detect, which is very difficult, and also such a type of solution is very expensive. So, to resolve such types of issues, they used social media networks data, and by using this data, they help us and save lives.

Twitter is one of the most essential, and it also contains lots of useful information. Such types of data are mentioned by users. By using Twitter, tweets related to the earthquake can be used to detect the temporal occurrence as well as try to find the locations of a humanitarian organization. For this purpose, they have proposed the system will be deep learning techniques such as RNN/LSTM to check the validity of Twitter tweets and try to give real-time detection of the

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

given event. In this, they have trained the machine learning model to train the tweets related to the earthquake. In the past, and that have been labeled by using crowdsourcing, plays an important role as the classifier to predict the validity of the tweets. In this system, they also used the Twitter API to listen to a particular keyword like earthquake and take these tweets as input. And then these tweets are converted to the word embedding's using BERT before applying them to the model. And the proposed system can also detect earthquakes. When it happens at a level of tolerance and ensures earlier warning to the public and then a particular or any websites.

In the [34] paper, they are also working on focusing on a Twitter-based response system that uses the recurrent net of the different training classifiers on a disaster of the specific datasets from the Tweets. They use datasets from Twitter because nowadays Twitter has become one of the major sources of data for any research community working on the base social media domain.

As we know, micro blogging sites receive millions of tweets in a single day on their platform. They explain that in the earlier study of the papers, they explain that during any disaster, the frequency of social media tweets grows exponentially. By using these tweets after analyses, they have obtained some actionable and important information related to the specific disaster or any specific event.

So many of the tweets are used by the design of a semi-automated AI (artificial intelligence)-based system, and they extracted the actionable information. By the use of this actionable information, one can take effective disaster response. So, in this paper, the main goal is to classify such types of tweets on the recurrent nets for training a classifier on the specific tweets. Their proposed system is enabling the timely dissemination of information to the various stakeholders so that action is on time for the disaster response.

And also, proactive measures are taken to reduce the many consequences of disasters. Experimental results show that the recurrent nets outperform the traditional of the different machine learning algorithms with accuracy in classifying disaster-specific tweets, and these tweets are taken from social media like Twitter.

In [35] this paper explains the previous work on earthquake impact assessments normally done by non-governmental organizations (NGOs) sponsoring and collecting the data. So, in that approach, more time is consumed, and as a result, they become more expensive. So, they explain that recently social media has become more valuable and more important, and easily collected data from different tools and techniques. A large amount of first-hand data after the disaster helps us easily see great potential for decision-making. Nevertheless, extracting valuable and meaningful information from social media is an ongoing area of research. So, in this paper they test the accuracy of the pre-trained sentiment analysis (SA) type model developed by the no of the code. Machine learning type platform and Monkey Learn using the text of the disaster-related to the emergency response and one early recovery phase of three types of the major earthquake and that type of the struck at Albania on the 26th of November 2019, In this disaster, 51 deaths, 3000 injuries, and a large number of extensive damages.

From this disaster, they have contained 695 tweets with hash tags. #Albania #Albanian Earthquake, #Albanian earthquake between 26th November 2019 and

3rd February 2020, In this paper, they used these data to test the accuracy of the SA pre-trained classification model, and they are developed by Monkey Learn to identify polarity in the text data. And this test explores the feasibility of automating the classification process to extract valuable and meaningful information from the text data from social media in real-time in the future. In which they test the no-code machine-type learning platforms and their performance by using the confusion matrix.
Finally, they obtained an overall accuracy (ACC) of 63%, and also the misclassification of the rate is 37%. They conclude that the overall accuracy (ACC) of the unsupervised classification is sufficient for a preliminary assessment, but further research is needed to determine if the accuracy is improved by customizing the training model of the machine learning platform.

Research Questions

- **Question 1.**

How to detect any disaster and manage a quick response?

- **Question 2.**

How good is the detection of disaster response in real-time generalization ability?

Research Objectives

- **Objective 1.**

To detect natural disasters by using different machine learning techniques

- **Objective 1.**

To identify the needs of the people affected by the disaster area

Research Gap

However, in the past, research in this field has not had much data available. For this purpose, we will require entirely different approaches, tools, and techniques to help inform decision-making under uncertain conditions. The mismatch between high disaster response and disaster resilience becomes a critical problem for emergency management.
"*However, there is still a need to quickly respond to any disaster by using social media (e.g., Twitter and Facebook (Meta)) tweets to detect early response detect the need on affected by the affected area.*"

Problem Statement

Whenever a disaster occurs, we have been too late informed, and our disaster response agencies' response is delayed. Due to the delayed response of our disaster response agencies as a result, the losses are increased. Our target is to detect disasters early and early respond to the disaster agencies for what is needed.

Methodology

Block diagram for proposed research methodology

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

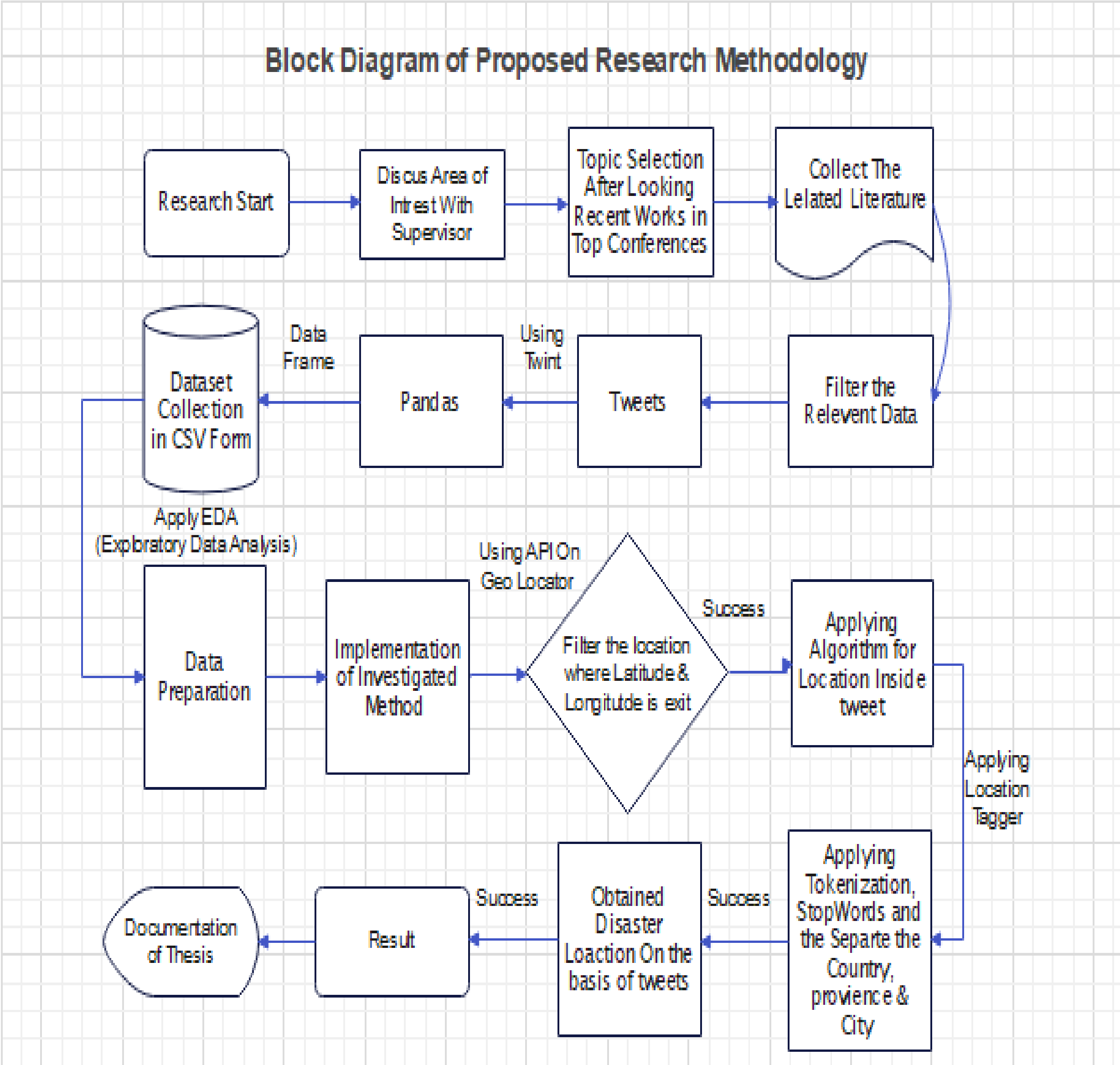


Figure 1 : Block Diagram for Research Methodology

Data Collection

- **Source of Data:** The primary data for this research comes from social media platforms, specifically Twitter and Facebook. Tweets related to earthquakes from the period of 2019 to 2022 are collected using the Twitter API and Facebook Graph API.
- **Data Sources**: Earthquake records from the United States Geological Survey (USGS) are used as ground truth data to verify the relevance and accuracy of the social media data.
- **Data Filtering**: The data is filtered to include only tweets containing specific keywords related to earthquakes, such as "earthquake," "quake," "disaster," and others. The filtering process ensures that irrelevant tweets are excluded from the dataset.

Data Preprocessing

- **Data Cleaning:** Raw tweets are preprocessed by removing stop words, special characters, URLs, and other irrelevant text. The text data is tokenized into individual words and then lemmatized to reduce words to their base form.
- **Merging Earthquake Data:** The social media data is merged with the earthquake records from USGS based on location and timestamp to create a comprehensive dataset for analysis.
- **Handling Missing Data:** Missing values are handled through imputation techniques or by removing records that lack crucial information, ensuring the dataset is complete for analysis.

Feature Extraction

- **Text Features:** Key features for text classification are extracted from the preprocessed data using techniques such as:
  ☐ **Bag of Words (BoW)**: A common method for representing text data by counting the frequency of each word in the document.
  ☐ **TF-IDF (Term Frequency-Inverse Document Frequency)**: A statistical measure used to evaluate the importance of a word in the context of the document corpus.
- **Sentiment Analysis:** Sentiment scores are assigned to each tweet based on whether it expresses positive, negative, or neutral sentiment towards the disaster.

Machine Learning Models

- **Model Selection:** Eight different machine learning algorithms are used to classify tweets related to earthquakes. These models include:
  1. **Neural Network**: A deep learning approach used to classify text data based on learned patterns.
  2. **Nearest Neighbors (KNN)**: A simple classification algorithm based on the proximity of data points in a feature space.
  3. **Linear Regression**: Used as a baseline for prediction based on linear relationships.
  4. **Naïve Bayes**: A probabilistic model used for text classification.
  5. **Decision Tree**: A tree-based model used for classification tasks.
  6. **Random Forest**: An ensemble method using multiple decision trees for improved prediction.
  7. **AdaBoost**: An adaptive boosting algorithm that combines weak learners to improve prediction accuracy.
  8. **Support Vector Machine (SVM)**: A supervised learning model that finds the optimal boundary between classes.

Model Evaluation

- **Metrics:** The performance of each model is evaluated using the following metrics:
  ☐ **Accuracy**: The percentage of correct predictions.
  ☐ **Precision**: The percentage of relevant results among all retrieved results.

☐ **Recall**: The percentage of relevant results among all possible relevant results.
☐ **F1-Score**: The harmonic mean of precision and recall, providing a balance between the two metrics.
☐ **Computational Time**: The time taken by each model to process the data and produce results.

- **Cross-Validation**: K-fold cross-validation is used to assess the generalization performance of the models and reduce the risk of over fitting.

Disaster Detection and Response Management

- **Real-Time Detection**: The models are tested for their ability to classify tweets in real-time during earthquake events. The response time and accuracy of disaster detection are key factors in evaluating the effectiveness of the system.
- **Response Management:** The ability to classify tweets accurately is crucial for disaster response efforts. The models' output is used to generate insights for disaster management teams to respond more effectively based on the social media data.

Comparison of Models

- The models are compared based on their performance in terms of prediction accuracy, computational time, and their ability to generalize across different disaster events. The effectiveness of the models in real-time disaster detection and response management is also assessed.

Results and Discussions

After applying EDA (Exploratory Data Analysis), now the data can work on it. We have collected our data from tweets. After analysis of our data, the main purpose of the use of the data is to find out the location and then apply machine learning to it. we have done some work before on location; first, we have plotted the msno bar on our data. Then this type of graph is shown.

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

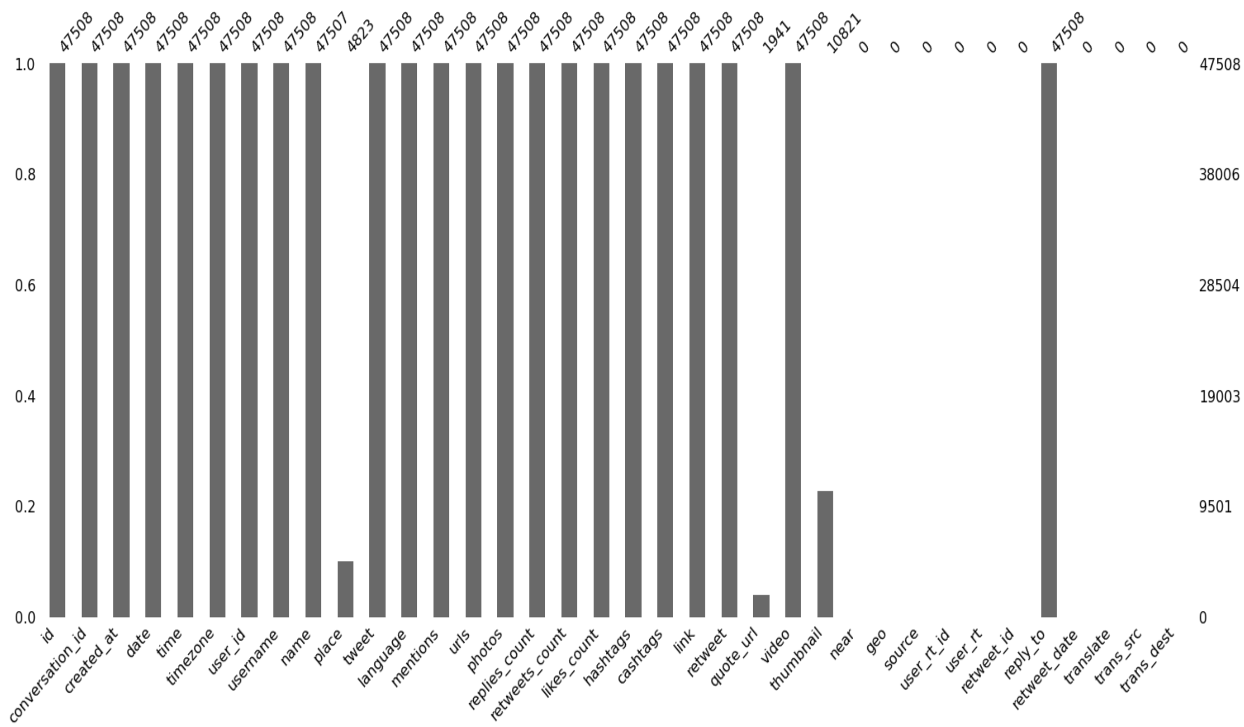


Figure 2: Msno Bar Chart of the Data

Our target is to find the out location from the tweets column and also find based on location and also on the basis of location and tweets we find out the response that is the need during any natural disaster and what the location on early and able to identify early on the place things is required.

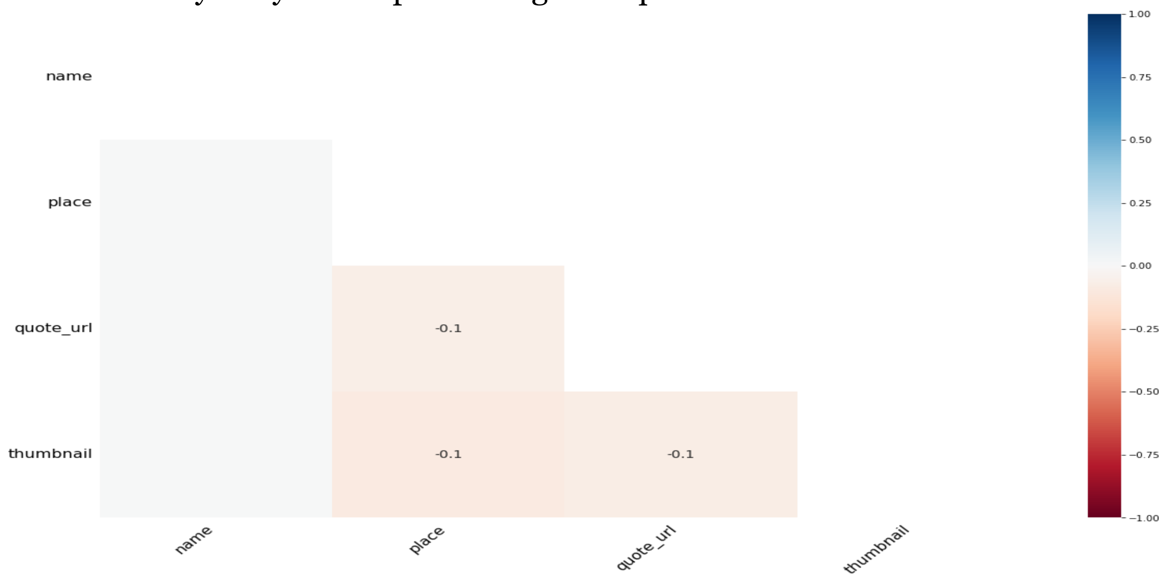


Figure 3 : Are Null Column are Correlated

In our dataset, there is a place column. When we are going to find out the location from the dataset, it is difficult to find the location. We have more than 1000 tweets. We have found only two tweets where place columns are not empty, and other than that, all data are empty.

So, we need such type of data where if some location is existing. So, again we

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

need new data from tweets to find out some locations exist. So, after some struggling again, we have been able to find out about another new dataset where more than a hundred locations are existing,
Now we apply EDA and filter the data. About Datasets: After describing our new dataset, there are a total of 35 columns, and the data types are in different types, like bool (1), float64 (10), int64 (8), and object (17). When we are checked the particular place column where the place column is not null, There is not any proper location name exit, but their place data is in XML format we have attached the picture of the place column, Below have plotted the dendrogram to separate the null values.

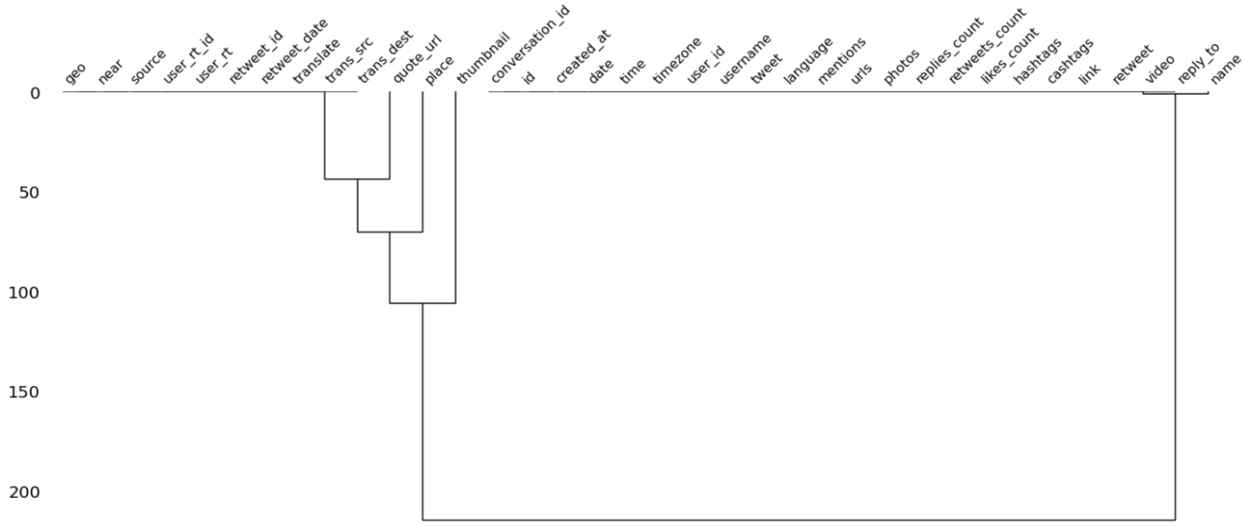


Figure 4 : Are there any common patterns of nullity between columns?

We have our data set according to our tweets based on. When we checked, the tweets were in different languages. So, first of all, we have removed all rows from the data set where the tweets are not in the English language. So, only the English tweets are remaining to exist.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

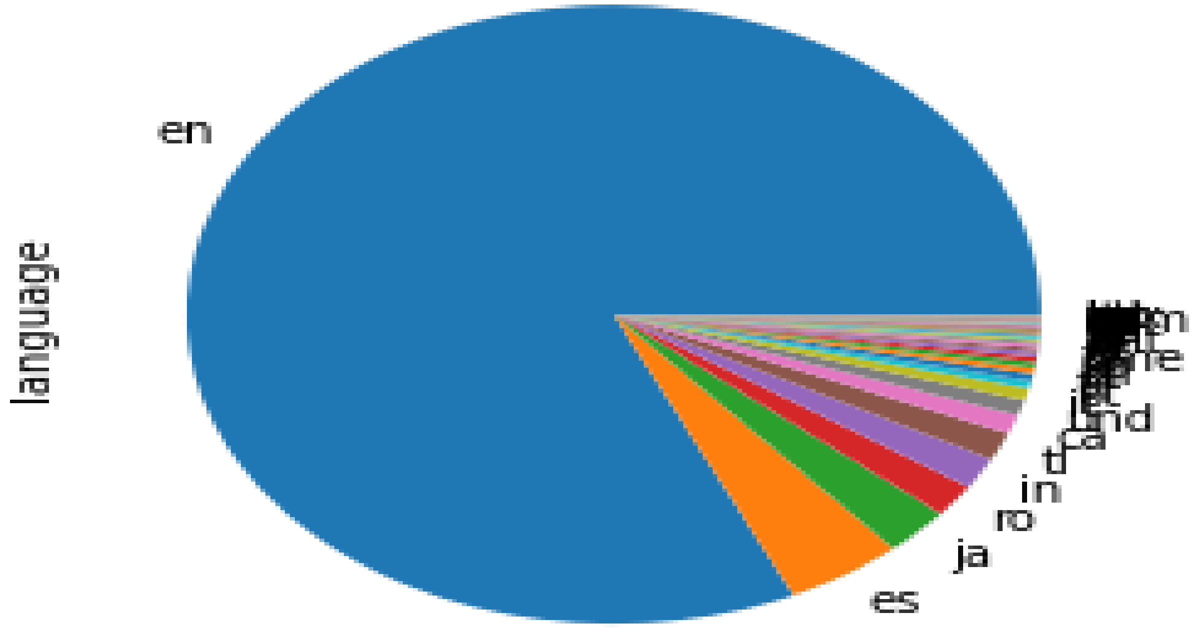


Figure 5 : Many languages are used in tweets?

So, you can see that the place column has contained the type in points and coordinates and also in some numbers for form.
After coordinates, the number is not a number but is latitude and longitude values.
So, we need to find out the location by using these latitude and longitude values and identify the location. When we started to find out the location using the given latitude and longitude, it was difficult to directly use the latitude and longitude. So, for this purpose, we need to apply the API. Before applying the API, we also used some important libraries to identify the place of the location column by using %matplotlib inline! pip install -U missingno geopy we have loaded our data set only a single time; they have successfully installed it on our local Anaconda on Jupyter notebooks. After installation, we applied our algorithm to it. So, after lots of errors, finally, we have been able to successfully find out the location on the basis of latitude and longitude. After one day, when we ran the same code, they took some time and did not find any location. we had already found the location one day before, but that time they showed some critical error.
The error is some server issue based on our tweets. It may be some security issue from the Twitter server side because their provided data maximum values of the location column are empty.
But the main issue is that when we hit the first time from our local server from Jupyter notebooks. Then that's time initially; our message successfully hits the location server and gets a particular location from the server according to our place column values of their key points and coordinates of their latitude and

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

longitude.
When we hit it the second time, they show an error on the JSON file and may have some API issues also; it's taken 15 to 20 min to get the location. After this issue occurred, we have been unable to resolve this issue for a few days to find the location.
So, as we know, Google Colab is one of the fastest and most advanced tools for the scientists. And then we think that maybe through Google Colab, location issues may be handled.
So, we have applied the same logic to Google Colab and then I'll successfully find out the location; also, it takes less time as compared to our local Anaconda on Jupyter notebooks.
So, Google Colab can successfully handle our API location issues and handle our request again and again. Below we have explained our logic applied for the location to create an algorithm for the location.
So, as we know, through the API, we can find out the location. So, first of all, we need to initialize the Nomination API. In the Nomination API, we get the geolocation by using the place column that exists in the useful data information, where points and coordinates and their latitude and longitude are. So, it's possible through geolocation that the user agent is'geoapiExercises.
Before applying the API, one thing is also needed to apply it. And also, one thing is that the place column is in XML format instead of the JSON format. As we know, XML format data is heavier as compared to JSON data. And JSON data is light data, and data is more efficient too as compared to all other formats of data. So, before applying the API, we need to convert it into JSON format, which is easy for us during the application of the API geolocator. So, when we applied geo-location, we were still unable to find the locations. And for this purpose, we have created an algorithm and found out the location. The algorithm for location according to our data format is also based on our requirement on our target.
After applying the algorithms, we have faced lots of issues when applying them separately on single tweets where latitude and longitude are existing, so that's why it's easy to extract the location, When the same logic applies overall to our dataset, then that error occurs, and we am unable to find out any location.
To resolve this problem, when we applied Lambda through the geolocator, then they could resolve this issue. The Python lambda is to execute expression, and a lambda is a functional use in Python. In Python, it is a small anonymous type function. Where that easily takes a lot of arguments and uses this, they find out the result and show the expression.
So, finally, we have successfully found the location. We have added a location column to our dataset to find out the result saved on our particular location. According to their place column, latitude and longitude, we have attached a fig below where you can see the image according to latitude and longitude, and their location is found.

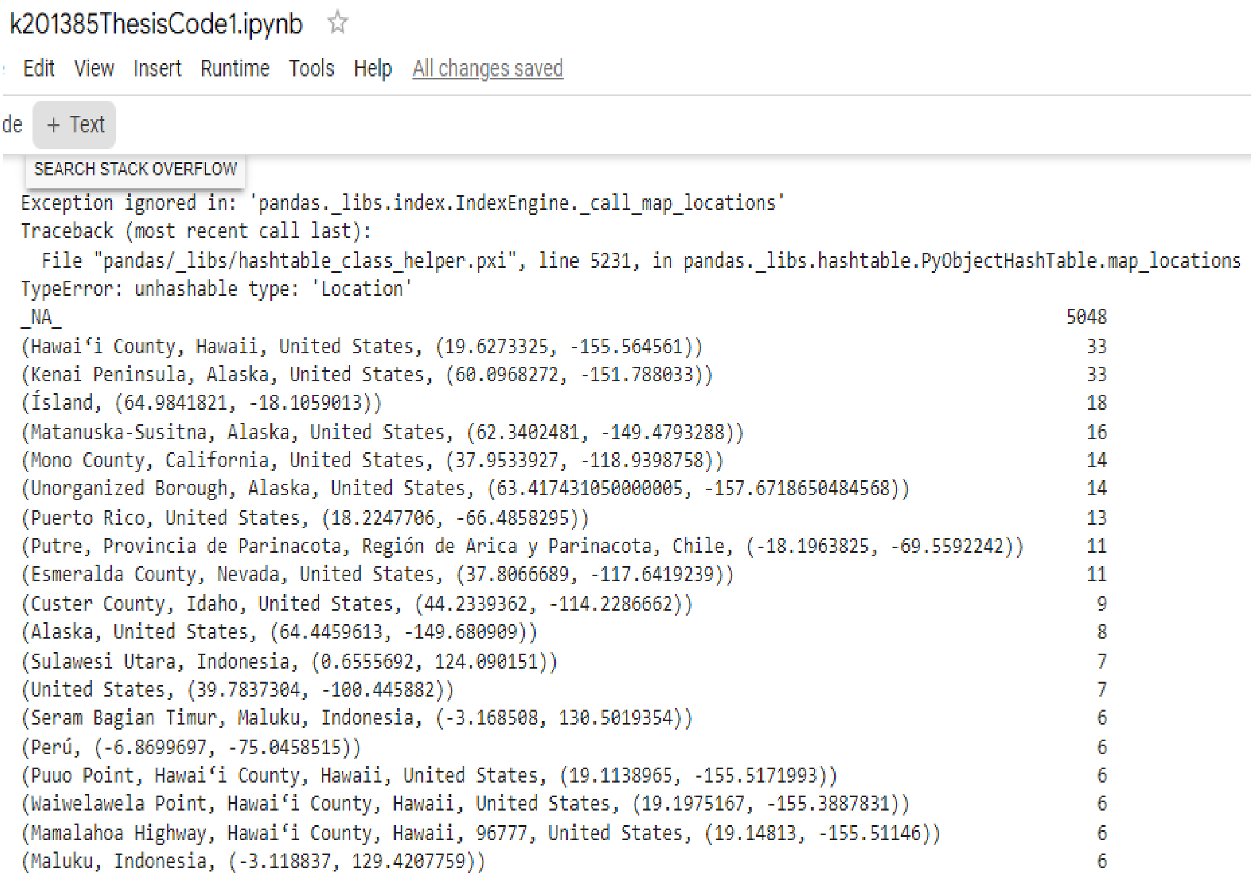
Result

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)



```
k201385ThesisCode1.ipynb
Edit View Insert Runtime Tools Help All changes saved
de + Text
SEARCH STACK OVERFLOW
Exception ignored in: 'pandas._libs.index.IndexEngine._call_map_locations'
Traceback (most recent call last):
  File "pandas/_libs/hashtable_class_helper.pxi", line 5231, in pandas._libs.hashtable.PyObjectHashTable.map_locations
TypeError: unhashable type: 'Location'
_NA_                                                                                        5048
(Hawaiʻi County, Hawaii, United States, (19.6273325, -155.564561))                            33
(Kenai Peninsula, Alaska, United States, (60.0968272, -151.788033))                           33
(Ísland, (64.9841821, -18.1059013))                                                           18
(Matanuska-Susitna, Alaska, United States, (62.3402481, -149.4793288))                        16
(Mono County, California, United States, (37.9533927, -118.9398758))                          14
(Unorganized Borough, Alaska, United States, (63.417431050000005, -157.6718650484568))        14
(Puerto Rico, United States, (18.2247706, -66.4858295))                                       13
(Putre, Provincia de Parinacota, Región de Arica y Parinacota, Chile, (-18.1963825, -69.5592242))  11
(Esmeralda County, Nevada, United States, (37.8066689, -117.6419239))                         11
(Custer County, Idaho, United States, (44.2339362, -114.2286662))                              9
(Alaska, United States, (64.4459613, -149.680909))                                             8
(Sulawesi Utara, Indonesia, (0.6555692, 124.090151))                                           7
(United States, (39.7837304, -100.445882))                                                     7
(Seram Bagian Timur, Maluku, Indonesia, (-3.168508, 130.5019354))                              6
(Perú, (-6.8699697, -75.0458515))                                                              6
(Puuo Point, Hawaiʻi County, Hawaii, United States, (19.1138965, -155.5171993))                6
(Waiwelawela Point, Hawaiʻi County, Hawaii, United States, (19.1975167, -155.3887831))         6
(Mamalahoa Highway, Hawaiʻi County, Hawaii, 96777, United States, (19.14813, -155.51146))      6
(Maluku, Indonesia, (-3.118837, 129.4207759))                                                  6
```

Figure 6: Location According to figure 3, latitude and longitude

In the above figure, you can see their complete location is met. In the above result, there are five different pieces of information that exist. In figure line 2, you can see that the first words are Kenai Peninsula, the second word is Alaska, and the third word is the United States, and the remaining numbers are latitude and longitude. Now, the first working is City Name, and the second word is States or Provinces Name and the third word is Country Name. So, the city name, province, and country name are existing, so by using these locations and comparing them with tweets, we can check the tweets information for our final goal. When we had found out the location and also cleaned the data, then we plotted the heat map for missing values of the dataset.

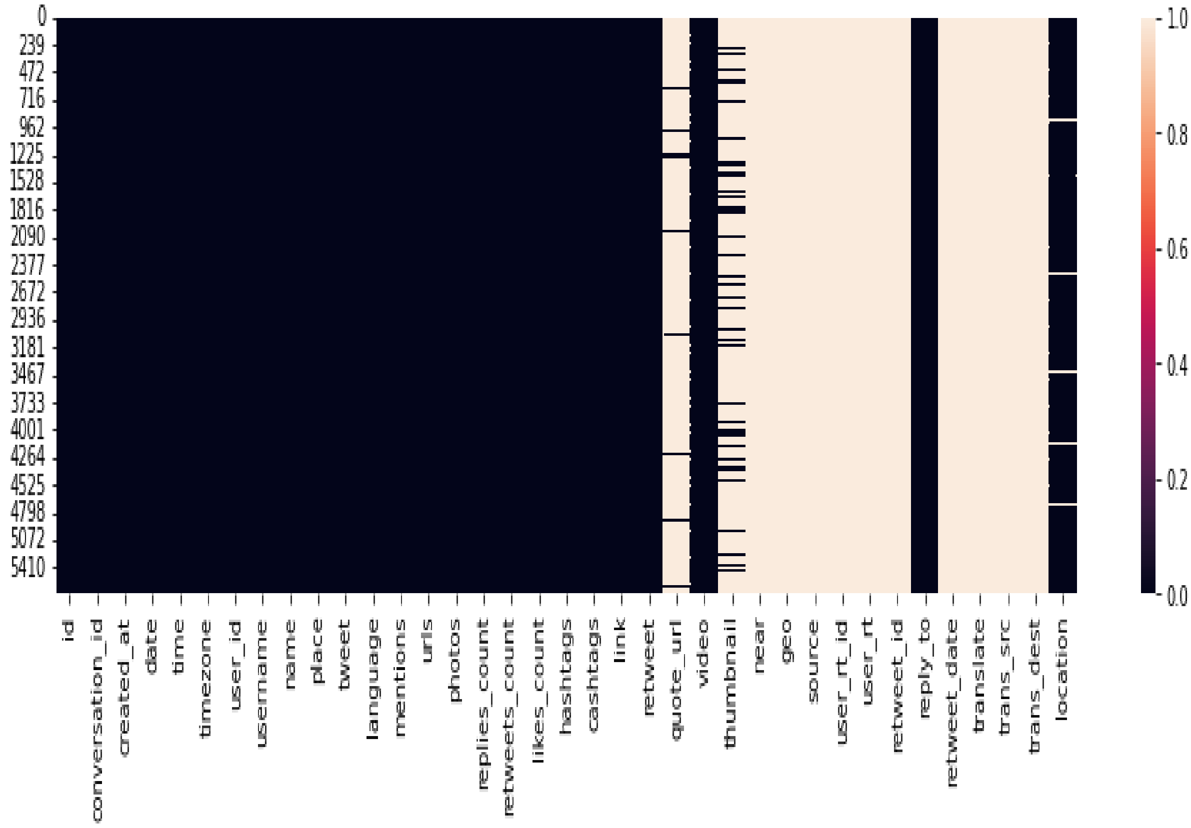


Figure 7: Heat Map Missing Values

In the above heat map plot, we can see that the end of the column is the location. We have added a new location column where their value is exiting. And also, we can see that there is some column that is empty, so we need this empty column, which is not required for our goal.
After finding out the location, we need to do some filtrations on our dataset. As we know that, the main target of the data parts is the tweets. When we check the tweets, there are different types of URLs that exist.
So, we need to remove all different types of URLs from tweets. It's not necessary to link use for our further work on our data. So, to remove links from the tweets. We need to create a function, so that's why we have used remove URLs (text); then we set the pattern and then compile the text of the data. After creating the function, we then applied the remove (URLs) on our tweets data and they have removed all types of URLs from our tweets.
After removing all types of URLs from the tweets, the next thing is that some HTML also exists. So, for this purpose, we need to remove all different types of the HTML from tweets. It's not necessary to link use for our further work on our data. So, we need to remove HTML from the tweets. We need to create a function. Before creating the function in Python, there is a library for HTML that helps us to remove HTML from the whole data.
So, after searching, we have found the library, i.e., from bs4 import Beautiful Soup, and after importing, the library defines the HTML text in Lxml form. And finally, by using the apply function on HTML text, that's why we have used remove URL (text), and then we set the pattern and then compile the text of the

data. After creating the function, we then applied the remove (URLs) on our tweets data and they have removed all types of URLs from our tweets.

Location Address by Using latitude and Longitude

Here we have attached our logic that applies to the thesis and how to get the location country name, province name, and city name if latitude and longitude have existed on tweets.

Remove Punctuations

To remove punctuation from tweets, first we have to import string punctuation. After importing the string, we defined the text and check method and then set the variable of x for punctuation. When we searched how many punctuations string the following punctuations i.e. '!"#$%&\'()+,-./:;< = >?@[\\]^_'{|}.

Now, we are going to apply the above string to all our tweets from the dataset. First of all, we have to define removing punctuation and then we have declared the text as null and the text is not null, then check the char in the text and apply the for loop, and if the char is not in string punctuations, also use resub for 0-9 values inside the text.

We have done the above conditions, and then we have all these conditions on our dataset by using the apply lambda and removing punctuations from tweets through referencing the declared variable after applying the above logic to our tweets, then the tweet column, then another column, and saving other data on the newly created.

Words Tokenization

After removing punctuations, the next step is to tokenize our dataset's Tweet_punct column tweets of the words. To remove the tokenization, first of all, we have declared the method for tokenization, and inside the tokenization, we have given the text.

After declaring the method, we split the words using the re.split function, and when they split the words, they return the words into the declared string value X.

When we have done the above conditions, we apply all the above conditions to our Tweet_punct column and also convert all words into lower words.

When we applied our above logic to the Tweet_punct column, then that's the time we created another new column named Tweet tokenized and saved the result in this column. Also, we have added logic that has been used for tokenization and how to tokenize the words from tweets and get the results properly.

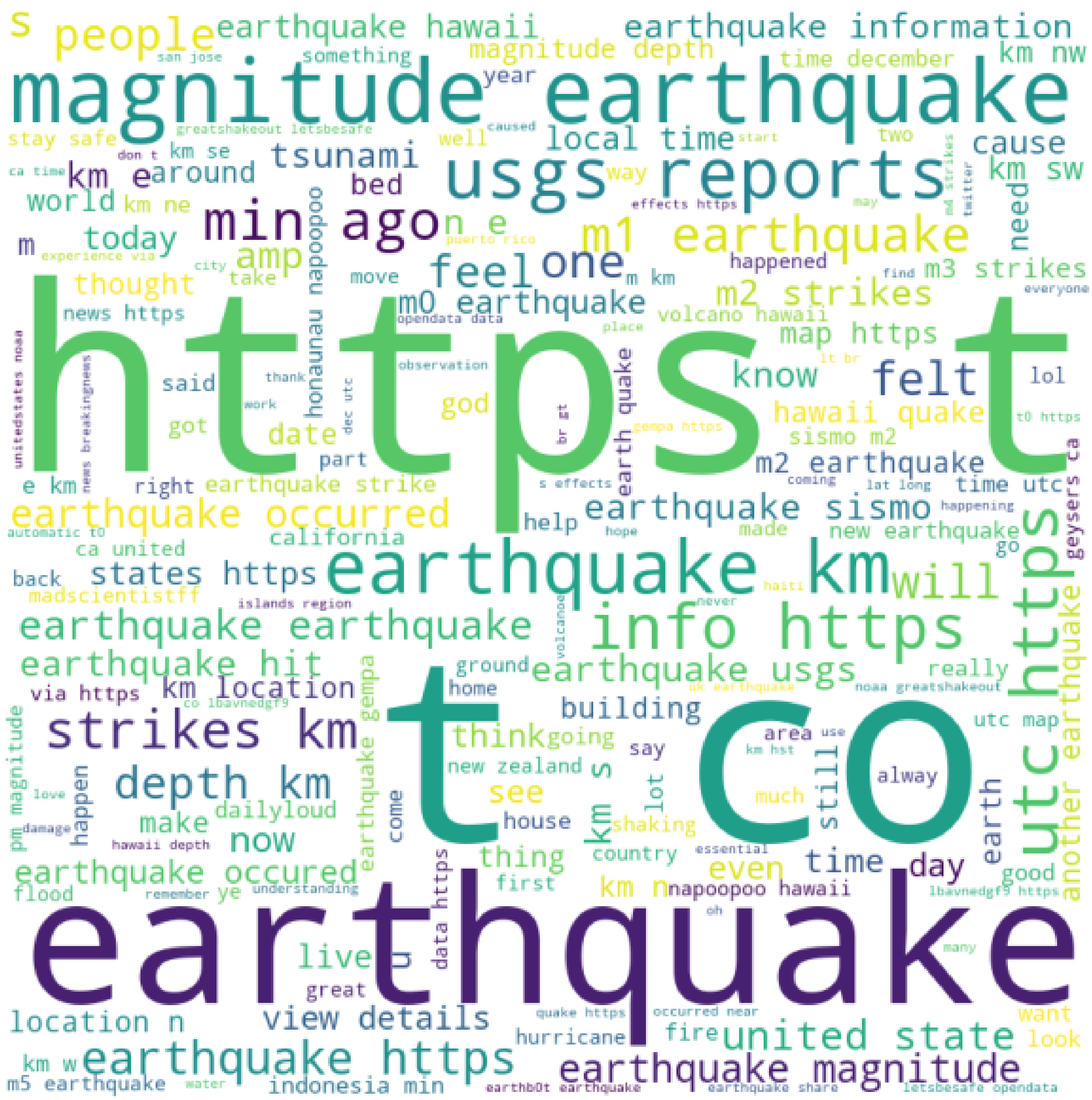


Figure 8: Word Cloud of Tweets 1

After removing the stop words, the picture is better than in figure 7.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

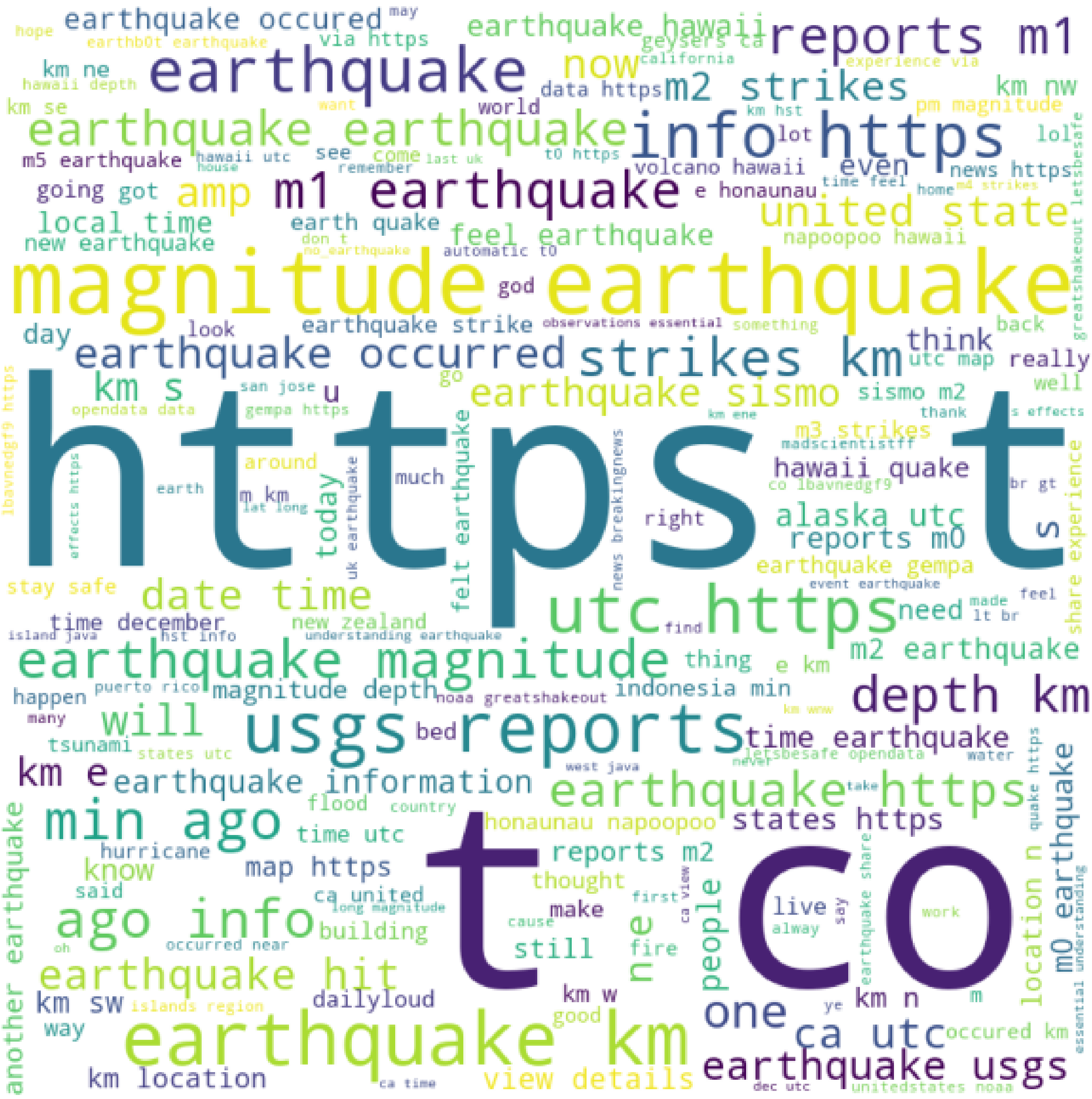


Figure 9 : Removing Stop Words

The above figure is further improved when we have removed the punctuation.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

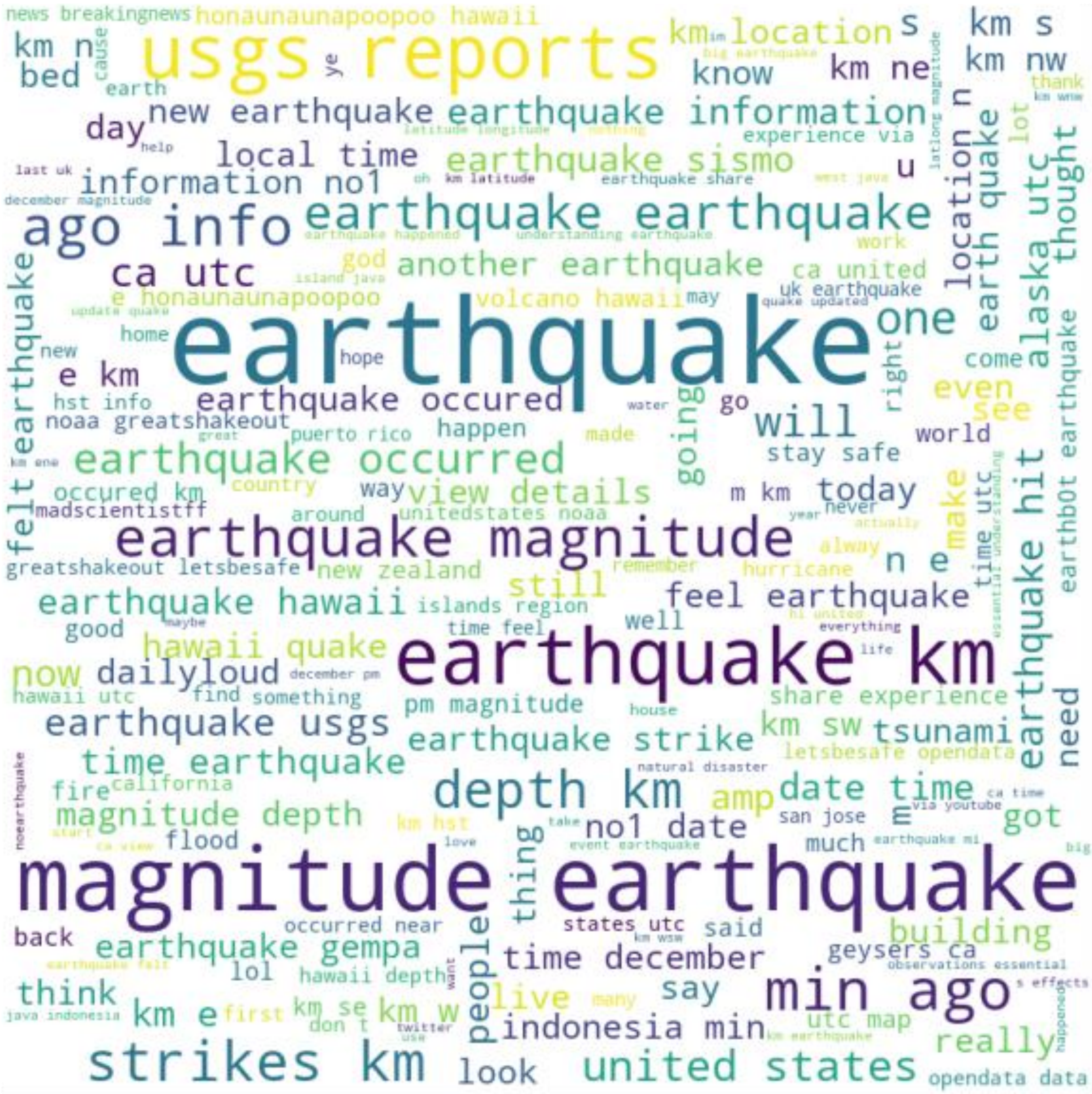


Figure 10 : After Remove the Punctuations

After applying the above-attached figure to our tweet, we got the tokenized result in the below figure.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

```
dfNew['Tweet_tokenized'].head(10)
```

```
0          [apparently, thank, you, very, much, is, something, elvis, said, a, lot, what, a, culture]
1    [i, was, watching, that, live, in, my, dorm, room, and, we, heard, al, michaels, say, the, word,...
2    [i, said, thank, you, very, much, to, the, waitress, at, the, sushi, place, and, she, replied, y...
3    [, magnitude, earthquake, occurred, at, esmeralda, nv, united, states, on, utc, map, earthbt, ea...
4    [, magnitude, earthquake, occurred, at, ottoboni, ridge, rd, cloverdale, ca, united, states, on,...
5                [i, thought, an, earthquake, was, suppose, to, break, cali, off, and, it, float, away]
6    [, magnitude, earthquake, occurred, at, sabodan, rd, bakersfield, ca, united, states, on, utc, m...
7    [last, week, our, hola, teammates, hosted, a, meet, greet, event, at, our, hq, to, raise, money,...
8                                                                               [geist, daddyferance]
9    [on, october, at, am, people, throughout, bc, will, practice, how, to, drop, cover, and, hold, o...
Name: Tweet_tokenized, dtype: object
```

Figure 11 : Tokenized Tables of the Values

Remove Stopwords

Before applying to remove the words first of all we have to import removed words and download nltk for stopwords. After downloading the packing of the stopwords and then another thing should be applied to the stopwords and by using NLTK to convert the corpus and the stopwords into English.

After doing the above part, now we apply the condition for stop words on text. To remove the stopwords, first of all, we have declared the method for removing stopwords, and inside the remove stopwords function, we have given the text.

After declaring the method, we checked each of the words used if the word is not in the stop words function, and when they split the word, they return the words to the declared string value X. When we have done the above, conditions apply to all conditions on our tweet tokenized column.

When we applied our above logic to the Tweet tokenized column, then that was the time we created another new column named Tweet nonstop and saved the result in this column. After removing the word, the result is the below figure.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

```
dfNew.head(10)
```

username	place	tweet	language	mentions	...	near	geo	source	reply_to	retweet_date	translate	location	Tweet_punct	Tweet_tokenized	Tweet_nonstop
no_earthquake	{}	apparently thank you very much is something elvis said a lot? what a culture	en	[]	...	NaN	NaN	NaN	[]	NaN	NaN	_NA_	apparently thank you very much is something elvis said a lot what a culture	[apparently, thank, you, very, much, is, something, elvis, said, a, lot, what, a, culture]	[apparently, thank, much, something, elvis, said, lot, culture]
voize_of_reazon	{}	I was watching that live (in my dorm room) and we heard Al Michael's say the word 'earthquake' b...	en	[]	...	NaN	NaN	NaN	[{'screen_name': 'SpiroAgnewGhost', 'name': 'Spiro's Ghost', 'id': '2323448533'}]	NaN	NaN	_NA_	I was watching that live in my dorm room and we heard Al Michaels say the word earthquake before...	[i, was, watching, that, live, in, my, dorm, room, and, we, heard, al, michaels, say, the, word,...	[watching, live, dorm, room, heard, al, michaels, say, word, earthquake, feed, cut, gulp]
no_earthquake	{}	i said thank you very much to the waitress at the sushi place and she replied you're welcome elvis	en	[]	...	NaN	NaN	NaN	[]	NaN	NaN	_NA_	i said thank you very much to the waitress at the sushi place and she replied you're welcome elvis	[i, said, thank, you, very, much, to, the, waitress, at, the, sushi, place, and, she, replied, y...	[said, thank, much, waitress, sushi, place, replied, welcome, elvis]

Figure 12: Remove Stopwords of the Values

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

Stemming and Lemmatization

For the use of the above two things, they need to create some important functions. We have created a function named stemming where we have taken text inside the function and applied it to their step words. After doing the above function, we then applied this function to our whole dataset on tweet datasets. After applying the above function, we got some results that we have added to the below figure.

```
dfNew['Tweet_stemmed'] = dfNew['Tweet_nonstop'].apply(lambda x: stemming(x))
dfNew.head()
```

user_id	username	place	tweet	language	mentions	...	geo	source	reply_to	retweet_date	translate	location	Tweet_punct	Tweet_tokenized	Tweet_nonstop	Tweet_stemmed
1316897066430451714	no_earthquake	{}	apparently thank you very much is something elvis said a lot? what a culture	en	[]	...	NaN	NaN	[]	NaN	NaN	_NA_	apparently thank you very much is something elvis said a lot what a culture	[apparently, thank, you, very, much, is, something, elvis, said, a, lot, what, a, culture]	[apparently, thank, much, something, elvis, said, lot, culture]	[appar, thank, much, someth, elvi, said, lot, cultur]
44883029	voize_of_reazon	{}	I was watching that live (in my dorm room) and we heard Al Michael's say the word 'earthquake' b...	en	[]	...	NaN	NaN	[{'screen_name': 'SpiroAgnewGhost', 'name': 'Spiro's Ghost', 'id': '2323448533'}]	NaN	NaN	_NA_	I was watching that live in my dorm room and we heard Al Michaels say the word earthquake before...	[i, was, watching, that, live, in, my, dorm, room, and, we, heard, al, michaels, say, the, word,...	[watching, live, dorm, room, heard, al, michaels, say, word, earthquake, feed, cut, gulp]	[watch, live, dorm, room, heard, al, michael, say, word, earthquak, feed, cut, gulp]
1316897066430451714	no_earthquake	{}	i said thank you very much to the waitress at the sushi place and she replied you're welcome elvis	en	[]	...	NaN	NaN	[]	NaN	NaN	_NA_	i said thank you very much to the waitress at the sushi place and she replied you're welcome elvis	[i, said, thank, you, very, much, to, the, waitress, at, the, sushi, place, and, she, replied, y...	[said, thank, much, waitress, sushi, place, replied, welcome, elvis]	[said, thank, much, waitress, sushi, place, repli, welcom, elvi]

Figure 13: Stemming and Lemmatization Values

Some issues occur during the course of which we need to install another library that is from NLTK, and that's WordNet. After downloading WordNet from NLTK, we then added another library from NLTK, which is OWM-1.4, and also added another important library for the lemmatizer, which is WordNetLemmatizer.

After stemming, then we are going to lemmatize the stemmed tweets. For this purpose, we have created another function that is named lemmatizer. Inside the function of the lemmatizer, take the text to apply this text on the lemmatizer function on each word, get the results, and return the tweets. We have added a function below showing how we have applied our function finally and got some results.

So, after applying the above function to all our tweets, we got some results and added these results below in figure form.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

```
dfNew.head()
```

user_id	username	place	tweet	language	mentions	...	source	reply_to	retweet_date	translate	location	Tweet_punct	Tweet_tokenized	Tweet_nonstop	Tweet_stemmed	Tweet_lemmatized
30451714	no_earthquake	{}	apparently thank you very much is something elvis said a lot? what a culture	en	[]	...	NaN	[]	NaN	NaN	_NA_	apparently thank you very much is something elvis said a lot what a culture	[apparently, thank, you, very, much, is, something, elvis, said, a, lot, what, a, culture]	[apparently, thank, much, something, elvis, said, lot, culture]	[appar, thank, much, someth, elvi, said, lot, cultur]	[apparently, thank, much, something, elvis, said, lot, culture]
44883029	voize_of_reazon	{}	I was watching that live (in my dorm room) and we heard Al Michael's say the word 'earthquake' b...	en	[]	...	NaN	[{'screen_name': 'SpiroAgnewGhost', 'name': 'Spiro's Ghost', 'id': '2323448533'}]	NaN	NaN	_NA_	I was watching that live in my dorm room and we heard Al Michaels say the word earthquake before...	[i, was, watching, that, live, in, my, dorm, room, and, we, heard, al, michaels, say, the, word,...	[watching, live, dorm, room, heard, al, michaels, say, word, earthquake, feed, cut, gulp]	[watch, live, dorm, room, heard, al, michael, say, word, earthquak, feed, cut, gulp]	[watching, live, dorm, room, heard, al, michael, say, word, earthquake, feed, cut, gulp]
30451714	no_earthquake	{}	i said thank you very much to the waitress at the sushi place and she replied you're welcome elvis	en	[]	...	NaN	[]	NaN	NaN	_NA_	i said thank you very much to the waitress at the sushi place and she replied you're welcome elvis	[i, said, thank, you, very, much, to, the, waitress, at, the, sushi, place, and, she, replied, y...	[said, thank, much, waitress, sushi, place, replied, welcome, elvis]	[said, thank, much, waitress, sushi, place, repli, welcom, elvi]	[said, thank, much, waitress, sushi, place, replied, welcome, elvis]
			1.90									magnitude				

Figure 14: Word Tokenized Values

After doing all these steps, now we want to check. Now we have added another function named clean_text(text). In this function, we have two different ways of the text. First, we checked the text in lowercase, which checks the text in words that are not in string punctuations and removes the punctuation. The second thing is to check the text from 0 to 9 and compare it with lowercase.

After passing the above conditions, we checked the tokenization and split the text of the word one by one and stored it in the W+ variable.

After completing the above tokenization steps, the second thing again was that we removed the stop words and stemming and returned all results in the text form.

After applying the logic, then the next steps are to check for the count vectorizer and for CountVectorizer, first we have to import CountVectorizer from sklearn.feature_extraction.text, and then CountVectorizer applies to the text to analyze the clean text. And then also check the count Vector fit to transform from the dataset tweet column and get the number of tweets and words.

The next step is to check the count vector in array form and also the count vectorizer. And we have got the feature names.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

```
count_vect_df = pd.DataFrame(countVector.toarray(), columns=countVectorizer.get_feature_names())
count_vect_df.head()
```

```
/usr/local/lib/python3.8/dist-packages/sklearn/utils/deprecation.py:87: FutureWarning: Function get_feature_names is deprecated; get_feature_na
  warnings.warn(msg, category=FutureWarning)
```

		aab	aaexox	aardbew	aata	aaya	ab	abc	abcdino	abduct	...	—	ALERT	MUSIC	M5	M6	NEW	*Issued*	*on*	07	67
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	2	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
4	2	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0

5 rows × 7253 columns

Figure 15: Table of the Count Vectors Values

After doing the above step, we now want to check the dataset size and shape of both datasets, original and updated, and check the difference between both datasets.

For this purpose, we have shape and column functions.

And also check the column for both datasets and check the difference between the datasets. For the original dataset column's shape:

For the new and updated dataset column's shape And after the above, all steps then also save these into a CSV form named cleaned.csv.

Find Location Inside the tweets

Above, we have explained how we found the location address by using latitude and longitude. We have done another work where we have found the location from inside the tweets.

We have checked all the tweets, and if any tweets have mentioned any country or any state/province or any city name that exists, then we have taken it from it. To obtain such a type of location from tweets. First of all, we have to install a location tagger.

As we know, the location tagger is used to detect and extract the locations of countries, states/regions, and cities from given text or their URLs. They also help us find the relationship among cities, provinces/states/regions, and countries.

After applying our logic location tagger and all NLTK, then we applied logic for the locations of the countries, states/regions/provinces, and city addresses to obtain success.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

```
dfNew['Tweet_punct'].apply(lambda txt:locationtagger.find_locations(text = txt))
```

```
0       <locationtagger.locationextractor.LocationExtractor object at 0x7f7a55eddf40>
1       <locationtagger.locationextractor.LocationExtractor object at 0x7f7b2d95f790>
2       <locationtagger.locationextractor.LocationExtractor object at 0x7f7b1de55e80>
3       <locationtagger.locationextractor.LocationExtractor object at 0x7f7a55edd160>
4       <locationtagger.locationextractor.LocationExtractor object at 0x7f7b28b4c700>
                                        ...
5665    <locationtagger.locationextractor.LocationExtractor object at 0x7f7a548ea1c0>
5666    <locationtagger.locationextractor.LocationExtractor object at 0x7f7a548ea5b0>
5667    <locationtagger.locationextractor.LocationExtractor object at 0x7f7a50079820>
5668    <locationtagger.locationextractor.LocationExtractor object at 0x7f7a50079dc0>
5669    <locationtagger.locationextractor.LocationExtractor object at 0x7f7a50079a60>
Name: Tweet_punct, Length: 3979, dtype: object
```

Figure 16: Table of the Location Extractor

After successfully finding out the mentioned country and city from inside the tweets, we have done this work because of this reason. The reason is to extract the country or city, and any reason is that there may be chances that many people tweet from any other place; maybe their location is taken from them. And they mention any county, city, and region name in their tweet. So, that's why we have extracted the mention of the country and city from the text of the tweets.

After the extracted countries, cities, and regions are taken into another new column. After finding these countries, cities, and regions from tweets, our next step is to find all those important and special items from tweets. And we have checked some important items like food, tents, matches, hygiene packs, camping gear, snacks, batteries, flashlights, etc. that are important items according to our thinking during any disaster time.

So, by mentioning a few items above, we have extracted these items the same as the country extracted from tweets, we have extracted all these items from tweets and saved them in another new item column on our database.

Our last work is we have been trying to find such types of words from tweets where these words have contained some intention. For example, need, needed, require, required, important, urgent, urgently, important etc. And the same as many more such types of words extracted from our tweets in which some intention exists. After finding these important words, then we are easily know what tweets are important, what is important inside the tweets, where their focus is, and what they want. After successfully found out all these important things. Then finally we created a table by combining all the above things.

Below we have mentioned the item table.

Items Table

No #	Item Count	Items
1	38	Backup battery
2	24	barbecue

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

3	8	battery
4	9	Battery Powered
5	64	book

Table 1: ITEMS TABLE

Earthquake Addresses and Time Tables

No#	Time Stamp	Tweet Clean	Region	Country	City
1	2022-08-01 04:59:06	70 magnitude earthquake hobipalooza alone	Nan	Nan	Nan
2	2022-08-01 04:58:07	brandyljensen noearthquake peligrietzer don't ...	Nan	Nan	Nan
3	2022-08-01 04:57:50	usgs report m198 earthquake 7 km w orca washin...	Washington	Nan	Washington
4	2022-08-01 04:57:50	usgs report m12 earthquake 65 km se denali nat...	Alaska	Nan	Nan
5	2022-08-01 04:57:38	transdano scared caused whole earthquake	Nan	Nan	Nan

Table 2: EARTHQUAKE ADDRESSES AND TIMES

And this is our features table, consisting of three main thin ones in the country and second items, and the third is the frequency of the table, and finally, we are still working on labels. Now, we have created the probability of the item with the tweets and addresses figure added here.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

	timestamp	tweet_clean	region	country	city	backup battery	barbecue	battery	battery powered	book	...	sleeping bag_probabiliy	snack_probabiliy	sturdy shoe_probabiliy	tent_probabiliy	warm blanket_probabiliy	water_probabiliy
0	2022-06-01 [illegible]	7.0 magnitude earthquake [illegible]	NaN	NaN	NaN	1	1	1	1	1	...	0.000025	0.000025	0.000025	0.000024	0.000025	0.000024
1	2022-06-01 [illegible]	[illegible]	NaN	NaN	NaN	1	1	1	1	1	...	0.000025	0.000025	0.000025	0.000024	0.000025	0.000024
2	2022-06-01 [illegible]	usgs report [illegible] earthquake 7 km [illegible]	Washington	NaN	Washington	1	1	1	1	1	...	0.000025	0.000025	0.000025	0.000024	0.000025	0.000024
3	2022-06-01 [illegible]	usgs report [illegible] earthquake 65 km se denali nat...	Alaska	NaN	NaN	1	1	1	1	1	...	0.000025	0.000025	0.000025	0.000024	0.000025	0.000024
4	2022-06-01 [illegible]	[illegible] caused whole earthquake	NaN	NaN	NaN	1	1	1	1	1	...	0.000025	0.000025	0.000025	0.000024	0.000025	0.000024

5 rows × 143 columns

waterproof_probabiliy	whistle_probabiliy	wipe_probabiliy	wrench_probabiliy
0.000025	0.000025	0.000025	0.000025
0.000025	0.000025	0.000025	0.000025
0.000025	0.000025	0.000025	0.000025
0.000025	0.000025	0.000025	0.000025
0.000025	0.000025	0.000025	0.000025

Figure 17 : Item Probability data values

So, there are some issues during the labeling of the data, so that's why, for this purpose, we have used USGS earthquake data from 2019 onward.
We combined both our data, and then we considered the week-wise earthquake that happened and the city it was located at. Then we got the final table.

Week Wise Earthquake Table

No#	timestamp	depth	magnitude	City	year	week
1770	2019-01-01 03:47:49.850000+00:00	50.000	3.32	Charlotte	2019	1
11768	2019-01-01 04:19:17.680000+00:00	46.030	4.60	Aurora	2019	1

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

11753	2019-01-01 11:55:02.180000+00:00	43.760	4.90	Banda Aceh	2019	1
11742	2019-01-01 14:37:02.540000+00:00	224.370	4.10	Seattle	2019	1
11735	2019-01-01 17:17:08.780000+00:00	108.390	4.10	San Lorenzo	2019	1
...	...	...	...	...	...	...
35	2022-12-10 10:44:37.380000+00:00	13.620	2.57	Nice	2022	49
33	2022-12-10 11:36:51.100000+00:00	10.290	3.33	San Juan	2022	49
25	2022-12-10 16:13:59.602000+00:00	6.310	2.60	Van	2022	49
12	2022-12-11 01:32:08.026000+00:00	134.910	4.50	San Pedro	2022	49
0	2022-12-11 09:07:41.641000+00:00	44.586	3.10	Charlotte	2022	49

Table 3: WEEK WISE EARTHQUAKE TABLE

After this we got the final feature, and then we labeled all places that had an earthquake as 1 and applied that the place had a quake and finally attached a figure of our final labeling data.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

data_final

	timestamp_x	tweet_clean	region	country	city	backup battery	barbecue	battery	battery powered	book	...	water	waterproof	whistle	wipe	wrench	year	week	timestamp_y	depth	magnitude
0	2022-08-01 04:57:50	usgs report m198 earthquake 7 km w orca washin...	Washington	NaN	Washington	1	1	1	1	1	...	1	1	1	1	1	2022	31	2022-08-04 13:22:39.400000+00:00	21.54	3.06
1	2022-08-01 04:57:03	wifegeist noearthquake peligrietzer yeah got c...	NaN	NaN	Christmas	1	1	1	1	1	...	1	1	1	1	1	2022	31	NaT	NaN	NaN
2	2022-08-01 04:55:40	cyno design leaked omfg guaranteed shakws as h...	NaN	NaN	Hard	1	1	1	1	1	...	1	1	1	1	1	2022	31	NaT	NaN	NaN
3	2022-08-01 04:53:50	usgs report m089 earthquake 30 km ne amboy was...	Washington	NaN	Washington	1	1	1	1	1	...	1	1	1	1	1	2022	31	2022-08-04 13:22:39.400000+00:00	21.54	3.06
4	2022-08-01 04:53:50	usgs report m089 earthquake 30 km ne amboy was...	Washington	NaN	Amboy	1	1	1	1	1	...	1	1	1	1	1	2022	31	NaT	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
8827	2021-10-01 04:47:30	bsbwolverine it's total package madison awesom...	NaN	NaN	Madison	1	1	1	1	1	...	1	1	1	1	1	2021	39	NaT	NaN	NaN
8828	2021-10-01 04:46:16	usgs report m22 earthquake 31 km se mina nevad...	Nevada	NaN	Nevada	1	1	1	1	1	...	1	1	1	1	1	2021	39	NaT	NaN	NaN
8829	2021-10-01 04:45:54	earthquake sismo m22 strike 75 km se hawthorne...	Nevada	NaN	Nevada	1	1	1	1	1	...	1	1	1	1	1	2021	39	NaT	NaN	NaN
8830	2019-09-01 04:47:56	billkarins nhcatlantic sister north carolina h...	Virginia	NaN	Virginia	1	1	1	1	1	...	1	1	1	1	1	2019	35	NaT	NaN	NaN
8831	2019-09-01 04:47:56	billkarins nhcatlantic sister north carolina h...	North Carolina	NaN	Virginia	1	1	1	1	1	...	1	1	1	1	1	2019	35	NaT	NaN	NaN

8832 rows × 79 columns

Figure 18: Final Labeled from Features Data

Now, data is prepared to apply the Models of our finally labeled data.

Prepare Dataset for Modeling

After successfully merging our new data with the USG's dataset, Then by the comparing both city-wise and month-wise by using embedding And we have got tweets embedding earthquake depth, earthquake magnitude, and also checking if the earthquake is in 0 and 1 form.

If 0, it means no earthquake. 1 mean earthquake occurs. We have attached below all related results in table and image forms.

	Timestamp	Tweets	Tweet Embedding	Earthquake Depth	Earthquake Depth	Earthquake Magnitude	City	Earthqua
0	2022-08-01 04:59:06	magnitude earthquake hobipalooza alone	[-0.05080826, -0.0025520804, 0.047641296,	-0.0...	0.00	0.00	UNKOWN	0

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

1	2022-08-01 04:58:07	brandyljensen noearthquake peligrietzer don't w...	[-0.10833041, 0.027527222, -0.0060223234,	- 0.0...	0.00	0.00	UNKOWN	0
2	2022-08-01 04:57:50	usgs reports m earthquake km w orcas Washington	[-0.0071552075, 0.0252599, -0.007953779,	- 0.10...	18.57	2.58	Washington	0
3	2022-08-01 04:57:50	usgs reports m earthquake km w orcas Washington	[-0.0071552075, 0.0252599, -0.007953779,	- 0.10...	42.34	2.77	Washington	0
4	2022-08-01 04:57:50	usgs reports m earthquake km w orcas Washington	[-0.0071552075, 0.0252599, -0.007953779,	- 0.10...	11.53	2.86	Washington	0

Table 4: Embedding Earthquake depth, Magnitude and cities

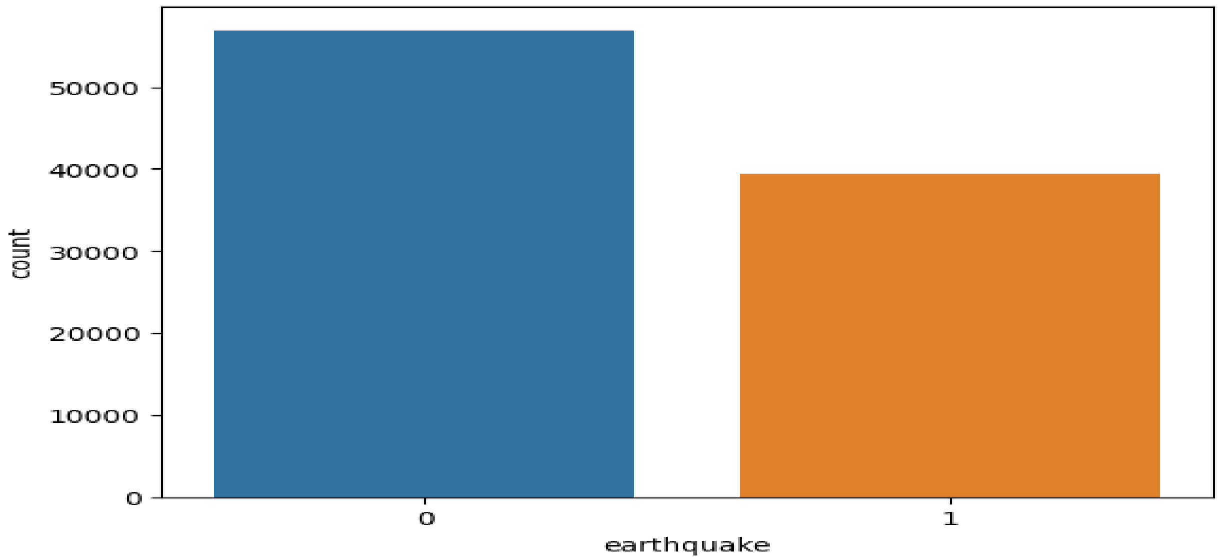


Figure 19: Earthquakes with Count figure

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

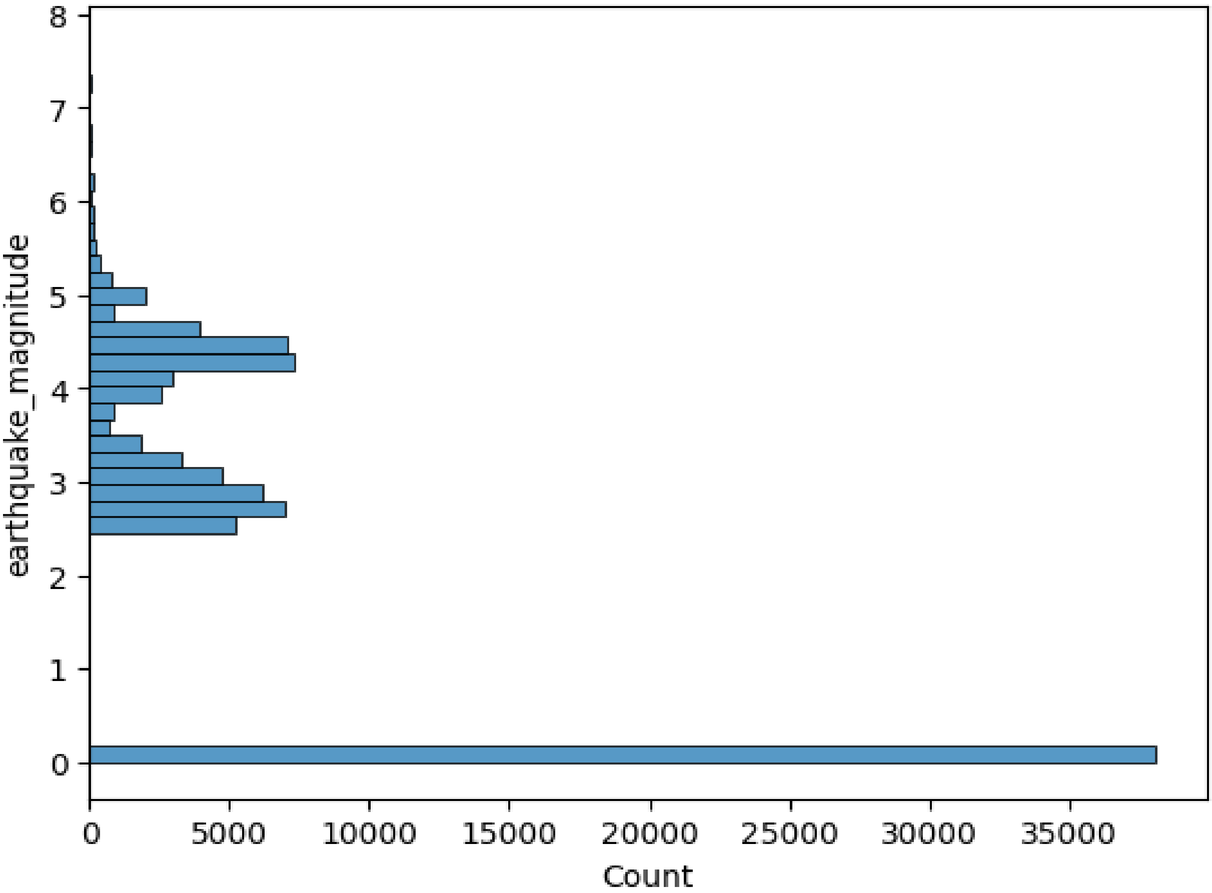


Figure 20: Earthquake Magnitude and Count

Another major contribution to this research work is a dataset of tweets on earthquakes between the years 2019 and 2022 along with details about earthquakes, including their depth, magnitude, and the city in which they occurred. And this allowed us to label the data with binary classes of an earthquake or not an earthquake along with frequencies of items most commonly needed during the earthquake. This dataset could further be used to train machine learning models that can predict an earthquake given a tweet as well as items needed during the earthquake.

Another major contribution to this research work is a dataset of tweets on earthquakes between the years 2019 and 2022 along with details about earthquakes, including their depth, magnitude, and the city in which they occurred.

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

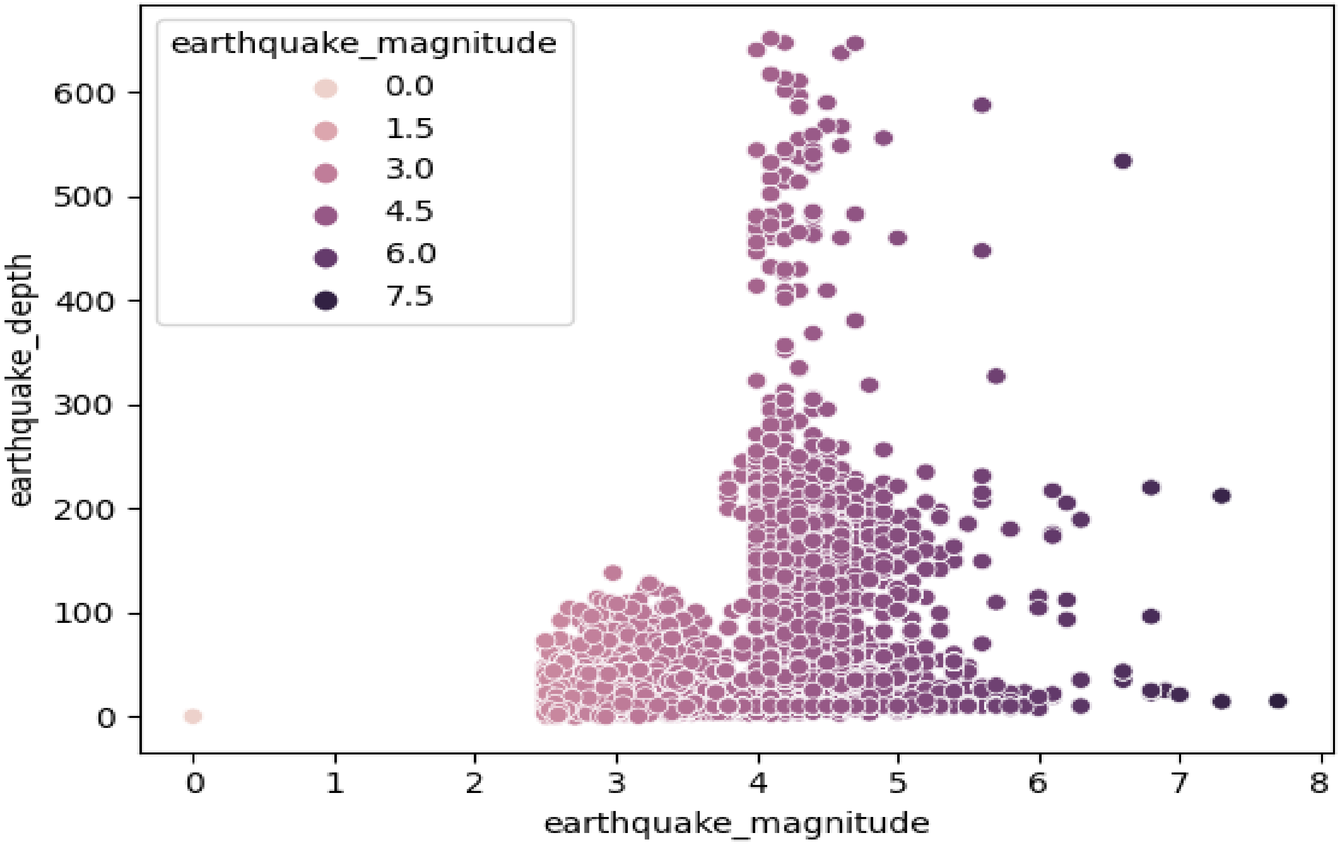


Figure 21: Earthquake Depth and Earthquake Magnitude

	tweet_embedding	earthquake_depth	earthquake_magnitude
0	[-0.05080826, -0.0025520804, 0.047641296, -0.0...	0.00	0.00
1	[-0.10833041, 0.027527222, -0.0060223234, -0.0...	0.00	0.00
2	[-0.0071552075, 0.0252599, -0.007953779, -0.10...	18.57	2.58
3	[-0.0071552075, 0.0252599, -0.007953779, -0.10...	42.34	2.77
4	[-0.0071552075, 0.0252599, -0.007953779, -0.10...	11.53	2.86
...	...	...	...
96402	[-0.0057738493, 0.05571529, 0.019276842, -0.00...	0.00	0.00
96403	[-0.09227878, 0.17164165, -0.116587184, 0.0580...	0.00	0.00
96404	[-0.034281936, -0.011746239, -0.0359722, -0.07...	0.00	0.00
96405	[-0.06891152, -0.044658184, 0.0034773592, 0.07...	0.00	0.00
96406	[0.07713166, 0.061372153, -0.18260805, 0.04686...	0.00	0.00

96407 rows × 3 columns

Table 1: Tweet Embedding, Earthquake Depth, and Earthquake Magnitude

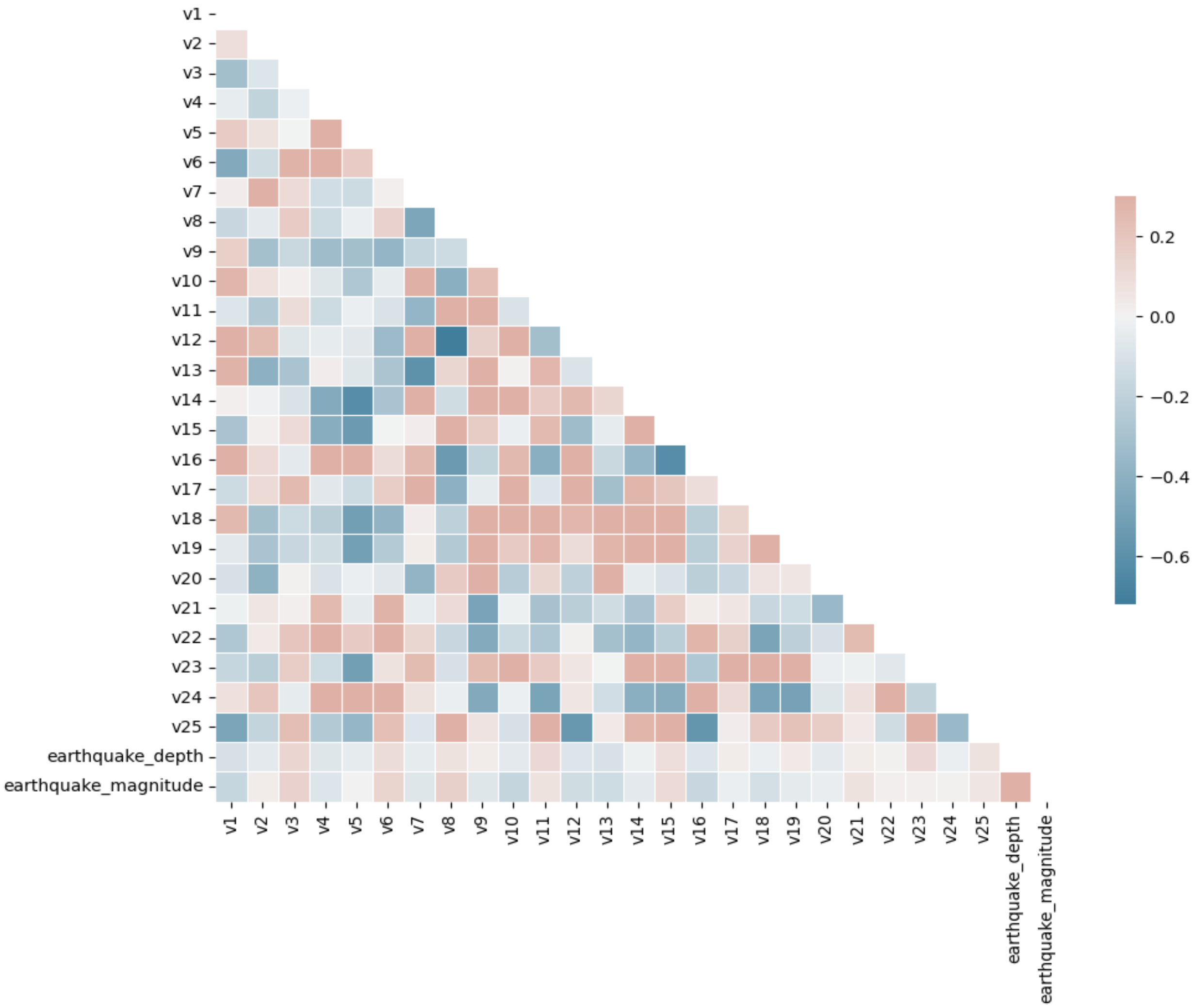


Figure 22: Subplot Earthquake Depth and Earthquake Magnitude

Final Classification

After fully preparing our dataset for modeling, we applied 8 different models.

1. Neural Net
2. Nearest Neighbors
3. Linear SVM
4. Logistic Regression
5. Naïve Bayes
6. Decision Tree
7. Random Forest
8. AdaBoost

The test and train size is 0.25, and the random state is 45.

Result: Neural Net

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

```
Training: Neural Net
Score: 1.0
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     14389
           1       1.00      1.00      1.00      9713

    accuracy                           1.00     24102
   macro avg       1.00      1.00      1.00     24102
weighted avg       1.00      1.00      1.00     24102

(array([1., 1.]), array([1., 1.]), array([1., 1.]), array([14389,  9713]))
[[14389     0]
 [    0  9713]]
```

Result: Nearest Neighbors

```
Training: Nearest Neighbors
Score: 0.9987552900174259
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     14389
           1       1.00      1.00      1.00      9713

    accuracy                           1.00     24102
   macro avg       1.00      1.00      1.00     24102
weighted avg       1.00      1.00      1.00     24102

(array([0.99930454, 0.99794302]), array([0.99861005, 0.99897045]), array([0.99895717, 0.99845647]), array([14389,  9713]))
[[14369    20]
 [   10  9703]]
```

Result: Linear SVM

```
Training: Linear SVM
Score: 1.0
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     14389
           1       1.00      1.00      1.00      9713

    accuracy                           1.00     24102
   macro avg       1.00      1.00      1.00     24102
weighted avg       1.00      1.00      1.00     24102

(array([1., 1.]), array([1., 1.]), array([1., 1.]), array([14389,  9713]))
[[14389     0]
 [    0  9713]]
```

Result: Logistic Regression

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

Vol. 3 No. 3 (March) (2025)

```
Training: Logistic Regression
Score: 1.0
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     14389
           1       1.00      1.00      1.00      9713

    accuracy                           1.00     24102
   macro avg       1.00      1.00      1.00     24102
weighted avg       1.00      1.00      1.00     24102

(array([1., 1.]), array([1., 1.]), array([1., 1.]), array([14389,  9713]))
[[14389     0]
 [    0  9713]]
```

Result: Naïve Bayes

```
Training: Naive Bayes
Score: 1.0
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     14389
           1       1.00      1.00      1.00      9713

    accuracy                           1.00     24102
   macro avg       1.00      1.00      1.00     24102
weighted avg       1.00      1.00      1.00     24102

(array([1., 1.]), array([1., 1.]), array([1., 1.]), array([14389,  9713]))
[[14389     0]
 [    0  9713]]
```

Result: Decision Tree.

```
Training: Decision Tree
Score: 1.0
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     14389
           1       1.00      1.00      1.00      9713

    accuracy                           1.00     24102
   macro avg       1.00      1.00      1.00     24102
weighted avg       1.00      1.00      1.00     24102

(array([1., 1.]), array([1., 1.]), array([1., 1.]), array([14389,  9713]))
[[14389     0]
 [    0  9713]]
```

Result: Random Forest

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

```
Training: Random Forest
Score: 0.9698365280889553
              precision    recall  f1-score   support

           0       0.97      0.98      0.97     14389
           1       0.97      0.95      0.96      9713

    accuracy                           0.97     24102
   macro avg       0.97      0.97      0.97     24102
weighted avg       0.97      0.97      0.97     24102

(array([0.96691729, 0.97434544]), array([0.9831121 , 0.95016988]), array([0.97494745, 0.96210581]), array([14389,  9713]))
[[14146   243]
 [  484  9229]]
```

Result: AdaBoost

```
Training: AdaBoost
Score: 1.0
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     14389
           1       1.00      1.00      1.00      9713

    accuracy                           1.00     24102
   macro avg       1.00      1.00      1.00     24102
weighted avg       1.00      1.00      1.00     24102

(array([1., 1.]), array([1., 1.]), array([1., 1.]), array([14389,  9713]))
[[14389     0]
 [    0  9713]]
```

Apply Confusion Matrix and Results.
Neural Net

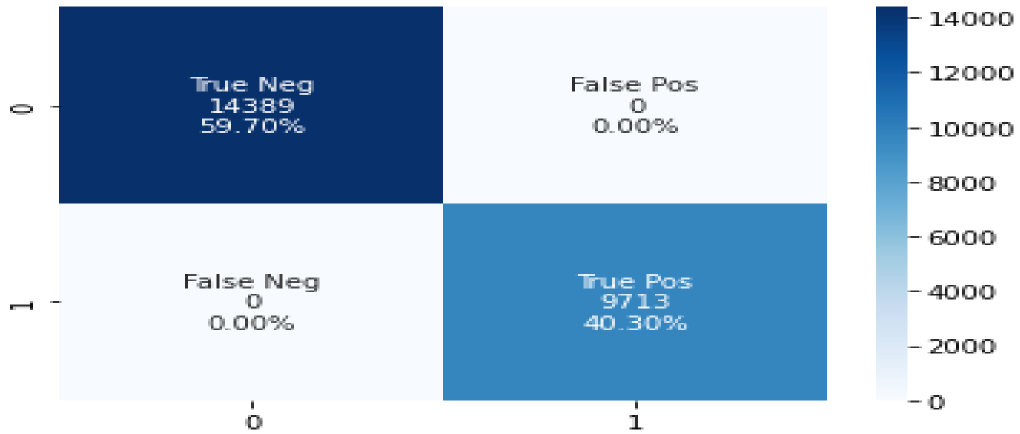


Figure 23: Final Neural Net

Near Forest Neighbors

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

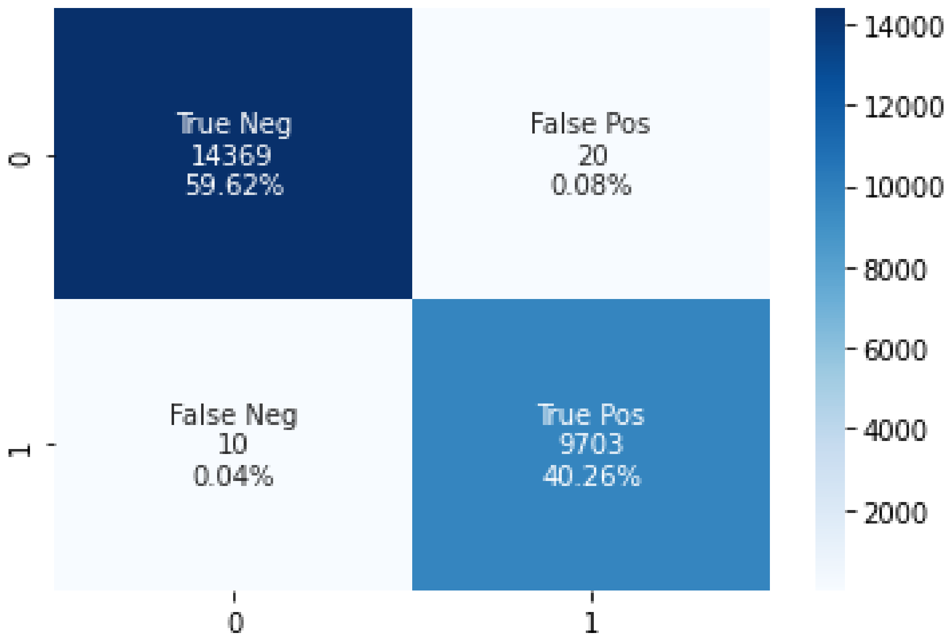


Figure 24: Final Nearest Neighbors

Linear SVM Image

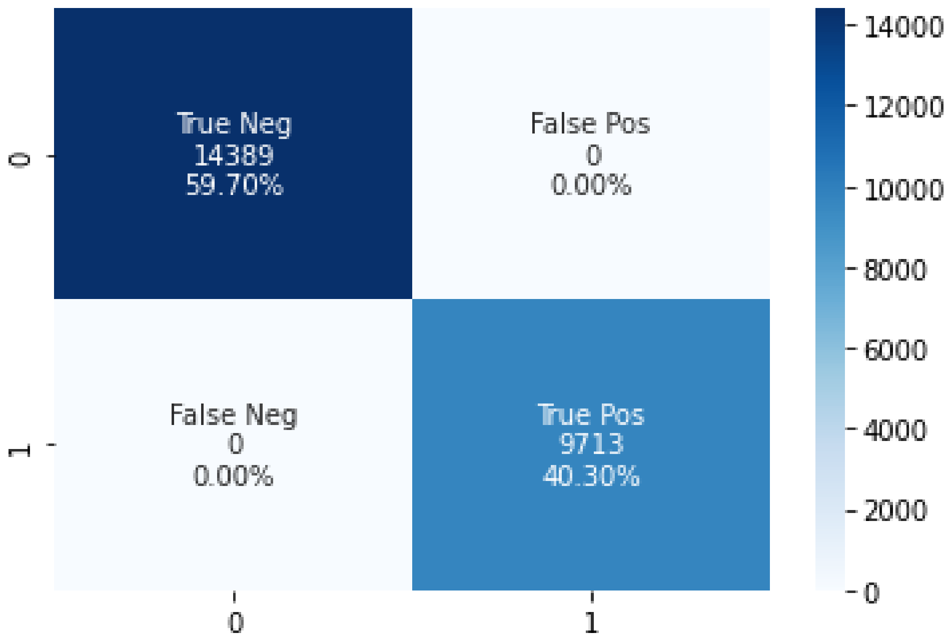


Figure 25: Linear SVM

Logistic Regression Image

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

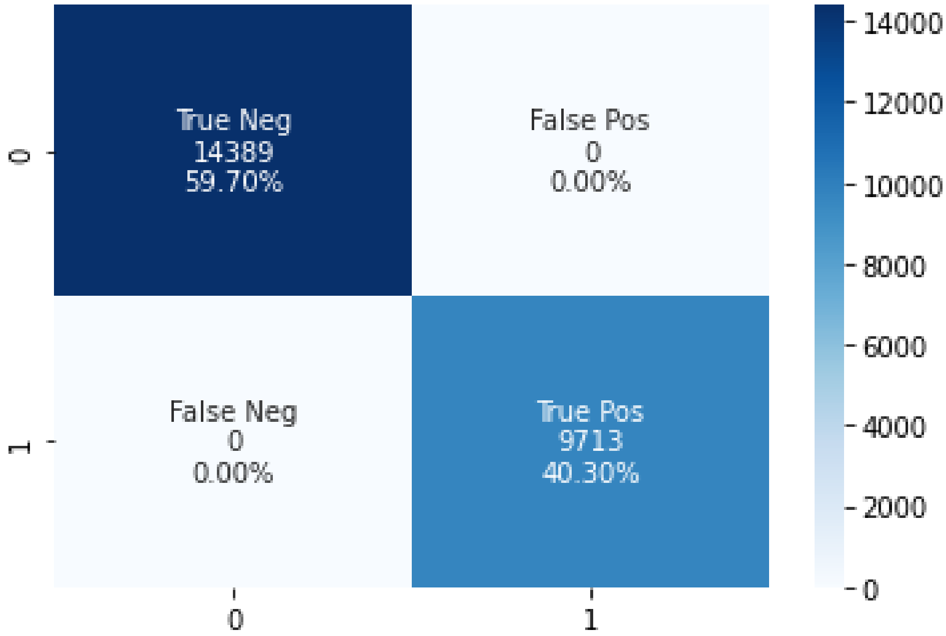


Figure 26: Logistic Regression

Naïve Bayes Image

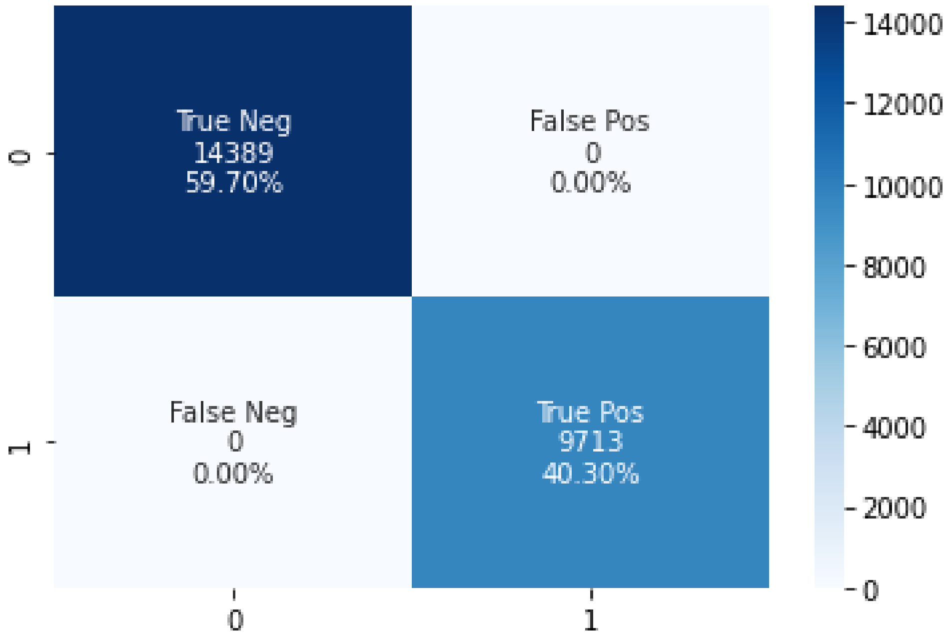


Figure 27: Naive Bayes

Decision Tree image

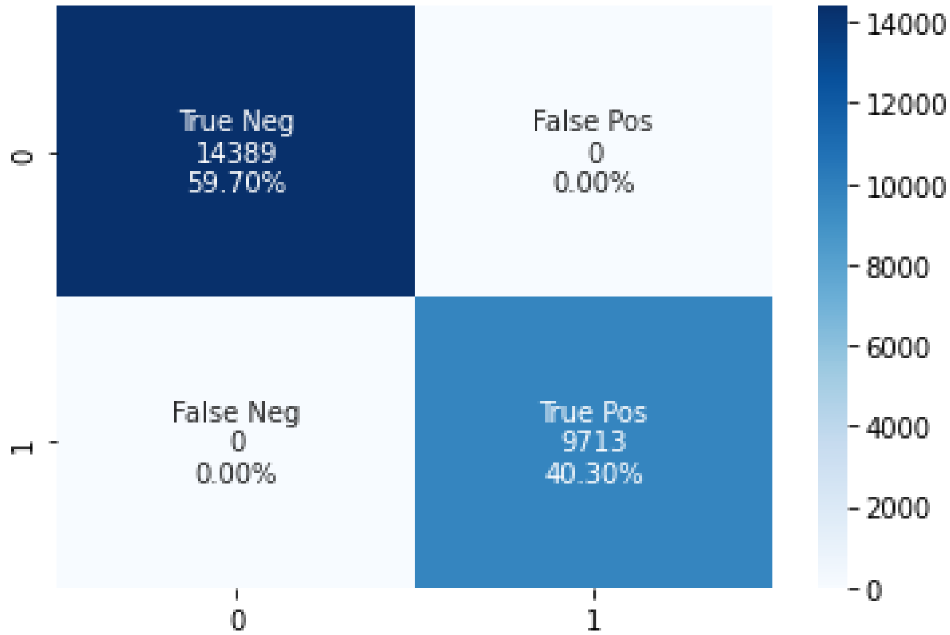


Figure 28: Decision Tress

Random Forest Image

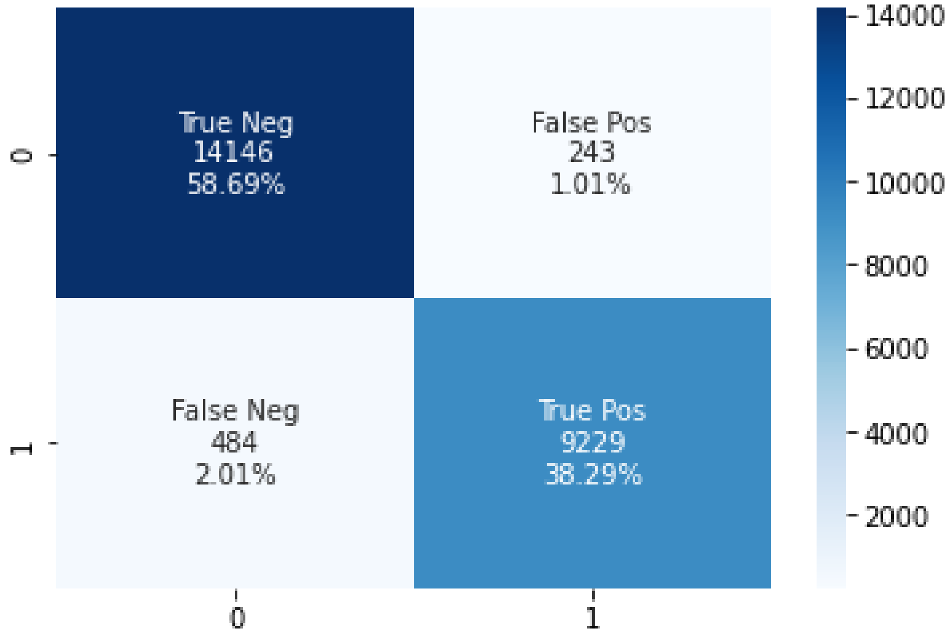


Figure 29: Random Forest

AdaBoost Image

Dialogue Social Science Review (DSSR)
www.thedssr.com

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

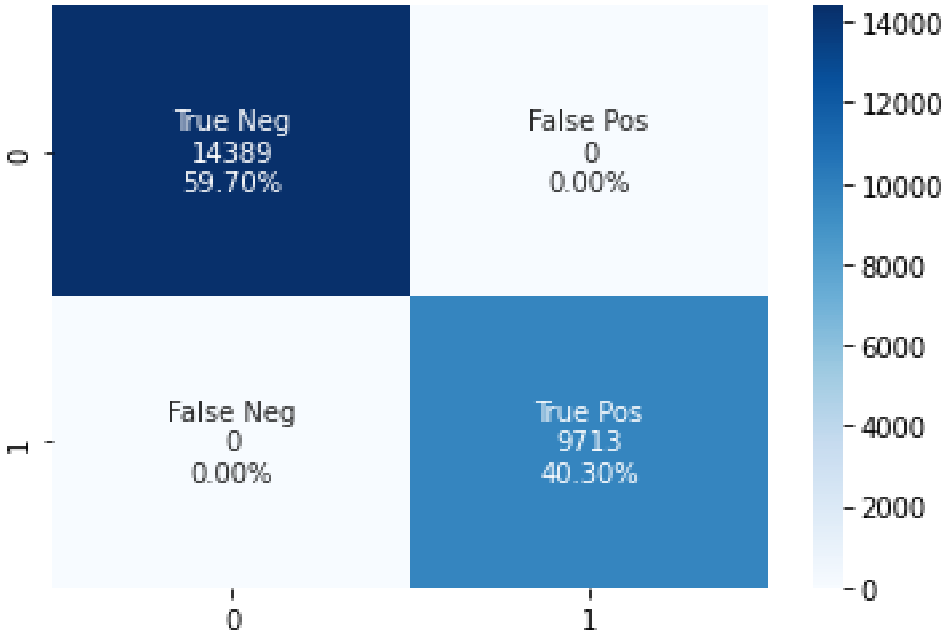


Figure 30: AdaBoost

Conclusion

In this thesis, we want social media data-driven analytics to be studied for improving disaster response efficiency, and they investigate the public sentimental characteristics via machine learning techniques used for the detection of disasters and managing the quick response. Social media platforms like Facebook and Twitter receive an overwhelming quantity of situational responsive information. Aimed at present disasters, emergency response understanding is insignificant; the organization of human and artificial intelligence can improve disaster response efforts. We have divided it into two ways. First, we collect the disaster data from social media (e.g., Twitter and Facebook). Secondly, we use different machine techniques and respond to any disaster and easily control the disaster.

Another major contribution to this research work is a dataset of tweets on earthquakes between the years 2019 and 2022 along with details about earthquakes, including their depth, magnitude, and the city in which they occurred. And this allowed us to label the data with binary classes of an earthquake or not an earthquake along with frequencies of items most commonly needed during the earthquake.

This dataset could further be used to train machine learning models that can predict an earthquake given a tweet as well as items needed during the earthquake.

Another major contribution to this research work is a dataset of tweets on earthquakes between the years 2019 and 2022 along with details about earthquakes, including their depth, magnitude, and the city in which they occurred.

And this allowed us to label the data with binary classes of an earthquake or not an earthquake along with frequencies of items most commonly needed during the earthquake.

After experiments, results in the thesis show that another major contribution to

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

this research work is a dataset of tweets on earthquakes.

The future of working is by using different machine learning models and techniques, deep learning, and NLP. We got some expected results in terms of form.

First, we have created our own dataset manually by using tweets on public data from 2019 to 2022. Then we have the data and remain only English tweets and remove other all-language tweets, and in this tweet's city names exist; that's why we have used name entity recognition and included the cities. And remove other extra columns from our dataset.

And then normalize the tweets and clean them by using an NLP pipeline and removing stop words and also lemmatizing, and special characters are also removed from it.

Secondly, we have the USGS website and scraped the data and obtained data from 2019 to 2022 with the same time duration and downloaded earthquake data.

And 3rdly, we have merged both datasets on the basis of cities, and the cities of both datasets must be the same. And within 45 days, if the disaster is occurring, then we have added it in front of the tweets.

And finally, we have applied 8 different machine learning algorithms and analyzed them.

References

[1] Zhijie Sasha Dong, Linqyu, Lauren Christenson & Lawrence Fulton, "Social media information sharing for natural disaster response." In Springer on Natural Hazard, 2021, pp. 2077–2104.

[2] Khatoon, Shaheen and Asif, Amna and Hasan, MdMaruf and Alshamari, Majed, "Social Media-Based Intelligence for Disaster Response and Management in Smart Cities" Disaster Response and Management in Smart Cities" in Springer, pp. 2022, 211–235.

[3] Firoj Alam, Ferda Ofli, Muhammad Imran, Michael Aupetit, "A Twitter Tale of Three Hurricanes: Harvey, Irma, and Maria" in arXiv preprint arXiv: 1805.05144, 2018, pp. 1805–05144.

[4] Alam, Firoj and Ofli, Ferda and Imran, Muhammad, "Crisismmd :Multimodel twitterdatasets from natural disasters" in Twelfth international AAAI conference on web social media, 2018, pp. 2018–00.

[5] Chuanjie Yang, Guofeng Su, Jianguo Chen, "Using big data to enhance crisis response and Disaster resilience for a smart city " in IEEE, 2017, pp. 504–507.

[6] Ramchurn, Sarvapali D and Huynh, Trung Dong and Wu, Feng and Ikuno, Yukki and Flann,Jack and Moreau, Luc and Fischer, Joel E and Jiang, Wenchao and Rodden, Tom and Simpson, Edwin and others, "A disaster response system based on human-agent collections" in Journal of Artificial Intelligence Research, 2016, pp. 661–708.

[7] Muhammad Imran, Carlos Castile, Ji Locus, Patrick Meier, Sarah Viewing, "AIDR: artificial intelligence for disaster response."InWWW '14 Companion: Proceedings of the 23rd International Conference on World Wide WebApril, 2014, pp. 159–162.

[8] M. Imran, C. Castillo, J. Lucas, M. Patrick, and J. Rogstadius."Coordinating human and machine intelligence to classify microblog communications in crises." Proc. of ISCRAM, 2014, pp. 2014–00.
[9] M. Imran, S. Elbassuoni, C. Castillo, F. Diaz, and P. Meier."Practical extraction of disaster-relevant information from social media". In Proc. of Workshop on Social Media Data for Disaster Management, WWW '13 Companion pages 1021–1024. ACM/IW3C2, 2013, pp. 1021–1024.
[10] ZongshengYue, JianwenXie, Qian Zhao, DeyuMeng, "Semi-Supervised Video Deraining with Dynamical Rain Generator,"in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 642–652.
[11] S.Vieweg,"Microblogged contributions to the emergency arena: Discovery, interpretation, and implications." In Proc. of CSCW, February 2010, pp. 515–516.
[12] Ansari, Mohdzeeshan and Aziz, MB and Siddiqui, MO and Mehra, H and Singh, KP."Analysis of political sentiment orientation on twitter" In Procedia Computer Science on Elsevier, 2020, pp. 1828–1828.
[13] Jang Won Bae, Kyohong Shin, Hyun-Rok Lee, Hyun Jin Lee, Taesik Lee, Chu Hyun Kim, Won-Chul Cha, and II-Chul Moon, "Evaluation of Disaster Response System Using Agent-Based Model with Geospatial and Medical Details" in IEEE Transaction On System, Man, and Cybernetics Systems, 2017, pp. 1454–1469.
[14] Kuo, Yong-Hong and Leung, Janny MY and Meng, Helen M and Tsoi, Kelvin KF, "A real-time decision support tool for disaster response: a mathematical programming approach" in IEEE International Congress on Big Data, 2015, pp. 639–642.
[15] Nugroho, Lukito Edi and Rakhman, Arkham Zahri and Lazuardi, Lutfan and others, "A refugee tracking system in dCoST-ER: Disaster command and support centre for emergency response." In 7th International Conference on Information Technology and Electrical Engineering (ICITEE), 2015, pp. 136–140.
[16] Kim, Seungun, and Matsubara, Masaki, and Morishima, Atsuyuki, "Analysis of Hard-drawn Maps of Places in Natural Disaster Pictures." In IEEE International Conference on Big Data (BigData), 2020, pp. 3076–2078.
[17] Lee, Min-Fan Ricky and Chein, Tzu-Wei, "Artificial intelligence and internet of things for robotic disaster response." In International Conference on Advanced Robotics and Intelligent Systems (ARIS), 2020, pp. 1–6.
[18] Soler, LS and Gon{\c{c}}alves, D and Gregorio, LT and Leal, P and Saito, S, "Fifth International Conference on Geo-Information Technology for Natural Disaster Management." In IEEE, 2013, pp. 42–51.
[19] Li, Linna and Ulaganathan, Manju Narmada, "Design and Development of a Crowdsourcing mobile app for disaster response." In 25th International Conference on Geoinformatics, 2017, pp. 1–4.
[20] Jahanian, Mohammad and Hasegawa, Toru and kawabe, Yoshinobu and koizumi, Yuki and Magdy, Amr and Nishigaki, Masakatsu and Ohki, Tetsushi and Ramakrishnan, kk, "Direct: Disaster response coordination with trusted volunteers." In IEEE, 2019, pp. 1–8.

[21] Cui, Yan and Li, Suju and Wang, Liying and shu, MoQuan and Shu, Yang, "Disaster event management based on Integrated Disaster Reduction and rapid Services Plateform." In IEEE, 2016, pp. 649–652.
[22] Wu, Jianium and Han, Xin_yi and Zhou, Yi and Yue, Peng and wang, Xiaoqing and lu, Jingxuan and jiand, Weiguo and Li, Jing and Tang, Hong and Wang, Futao and Others, "Disaster Monitoring and Emergency Response Services in China." In IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 3473–3476.
[23] Naik, Nitin, "Flooded streets—A crowedsourced sensing system for disaster response: A case study." In IEEE International Symposium on Systems Engineering (ISSE), 2016, pp. 1–3.
[24] Zhao, Qing-Zhan and Zheng, Xu-Rong and Li, Wei and Wang, Chuan-Jian and Zhang, Xian-Feng, "Framework design of natural disasters and emergency management system oriented to region." In Fifth International Conference on Geo-Information Technologies for Natural Disaster Management, 2013, pp. 108–113.
[25] Mudombi, Shakespear and Muchie, Mammo, "ICTs in development and disaster response: opportunities and challenges for Africa." In 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010, pp. 1–5.
[26] Bala, Hillol and Venkatesh, Viswanath and Venkatraman, Srinivasan and Bates, Jack, "If the worst happens: Five strategies for developing and leveraging information technology-enabled disaster response in healthcare." In IEEE journal of biomedical and health informatics, 2015, pp. 1545–1551.
[27] Dubois, David and Lepage, Richard, "Meeting remote sensing requirements for faster disaster response." In IEEE International Geoscience and Remote Sensing Symposium, 2012, pp. 2986–2989.
[28] Kryvasheyeu, Yury and Chen, Haohui and Obradovich, Nick and Moro, Esteban and Van Hentenryck, Pascal and Fowler, James and Cebrian, Manuel, "Rapid assessment of disaster damage using social media activity." In Science Advance, 2016, pp. 3.
[29] Laverdiere, Melanie and Yang, Lexie and Tuttle, Mark and Vaughan, Chris, "Rapid Structure Detection in Support of Disaster Response: A Case Study of the 2018 Kilauea Volcano Eruption." In IEEE, 2020, pp. 6826–6829.
[30] Seo, Kap-Ho and Park, Jeong Woo and Oh, Seungsub and Hahm, Jehun and Suh, Jinho, "Requirement analysis of simulator-based integration for disaster response robots." In IEEE, 2017, pp. 1295–1299.
[31] Madichetty, Sreenivasulu and Sridevi, M. "Improved classification of crisis-related data on Twitter using contextual representations." In Prodedia Computer Science on Elsevier, 2020, pp. 962–968.
[32] Gata, Windu and Amsury, Fachriand Wardhani, Nia Kusumaand Sugiyarto, Ipin and Sulistyowti, Daning Nur and Saputra, Irwansyah. "Informative Tweet Classification of the Earthquake Disaster Situation In Indonesia." In 2019 5th International Conference on Computing Engineering and Design (ICCED) on IEEE, 2019, pp. 1–6.
[33] George, Eldho Ittan and Abraham, Cerene Mariam. "Real-time earthquake detection by using Twitter tweets." In AIP Conference Proceedingson AIP Publishing LLC, 2022, pp. 030014.

[34] Lamsal, Rabindra and Kumar, TV Vijay. “Twitter-based disaster response using recurrent nets.” In Research Anthology on Managing Crisis and Risk Communications on IGI Global, 2022, pp. 613–632.
[35] Contreras, Diana and Wilkinson, Sean and Alterman, Evangeline and Herv{\'a}s, Javier. “Accuracy of a pre-trained sentiment analysis (SA) classification model on tweets related to emergency response and early recovery assessment: the case of 2019 Albanian earthquake.” In Natural Hazards on Springer, 2022, 1–19.
[36] Ruan, Tao and Kong, Qingkai and McBride, Sara K and Sethjiwala, Amatullah and Lv, Qin. “Cross-platform analysis of public responses to the 2019 Ridgecrest earthquake sequence on Twitter and Reddit.” In Scientific reports on Nature Publishing Group, 2022, pp. 1–14.
[37] Lagria, Raymond Freth and Jalao, Eugene Rex and Resurreccion, Joanna. “A Text Mining Framework for the Classification and Prioritization of Disaster-Related Tweets for Disaster Response.” In Philippine Engineering Journal, 2022, pp. 1.
[38] Kumar, Abhinav and Singh, Jyoti Prakash and Rana, Nripendra P and Dwivedi, Yogesh K. “Multi-Channel Convolutional Neural Network for the Identification of Eyewitness Tweets of Disaster.” In Information Systems Frontiers on Springer, 2022, pp. 1–16.
[39] Mukherjee, Shubhadeep and Kumar, Rahul and Bala, Pradip Kumar. “Managing a natural disaster: actionable insights from microblog data.” In Journal of Decision Systems on Taylor \& Francis, 2022, pp. 134–149.
[40] Munawar, Hafiz Suliman and Mojtahedi, Mohammad and Hammad, Ahmed WA and Kouzani, Abbas and Mahmud, MA Parvez. “Disruptive technologies as a solution for disaster risk management: A review.” In Science of the total environment on Elsevier, 2022, pp. 151351.
[41] Djoumessi, Yannick Fosso and Mbongo, Louis de Berquin Eyike. “An analysis of information Communication Technologies for natural disaster management in Africa.” In International Journal of Disaster Risk Reduction on Elsevier, 2022, pp. 102722.
[42] Sufi, Fahim K and Khalil, Ibrahim. “Automated Disaster Monitoring From Social Media Posts Using AI-Based Location Intelligence and Sentiment Analysis.” In IEEE Transactions on Computational Social Systems on IEEE, 2022.
[43] Poddar, Soham and Mondal, Mainack and Ghosh, Saptarshi and Jana, Arnab. “A Survey on Disaster: Understanding the After-Effects of Super-Cyclone Amphan, the Helping Hand of Social Media.” In Advances in Urban Design and Engineering on Springer, 2022, pp. 134–149.
[44] Khaleq, Abeer Abdel and Ra, Ilkyeun. “Cloud-based disaster management as a service: A microservice approach for hurricane twitter data analysis.” In 2018 IEEE Global Humanitarian Technology Conference (GHTC) on IEEE, 2018, pp. 48–61.
[45] Reynard, Darcy and Shirgaokar, Manish. “Harnessing the power of machine learning: Can Twitter data be useful in guiding resource allocation decisions during a natural disaster?.” In Transportation research part D: Transport and environment on Elsevier, 2019, pp. 449–463.

ISSN Online: 3007-3154
ISSN Print: 3007-3146

DIALOGUE SOCIAL SCIENCE REVIEW

Vol. 3 No. 3 (March) (2025)

[46] Chen, Xiannian and Elmes, Gregory and Ye, Xinyue and Chang, Jinhua}. "Implementing a real-time Twitter-based system for resource dispatch in disaster management." In GeoJournal on Springer, 2016, pp. 863–873.
[47] Hariharan, Kartick and Lobo, Ashley and Deshmukh, Sujata. "Hybrid Approach for Effective Disaster Management Using Twitter Data and Image-Based Analysis." In 2021 International Conference on Communication information and Computing Technology (ICCICT) on IEEE, 2021, pp. 1–6.
[48] Banujan, K and Kumara, Banage TGS and Paik, Incheon. "Twitter and online news analytics for enhancing post-natural disaster management activities." In 2018 9th International Conference on Awareness Science and Technology (iCAST) on IEEE, 2018, pp. 302–307.
[49] Ilyas, Andrew. "Microfilters: Harnessing twitter for disaster management." In IEEE Global Humanitarian Technology Conference (GHTC 2014) on IEEE, 2014, pp. 417–424.
[50] Nair, Meera R and Ramya, GR and Sivakumar, P Bagavathi. "Usage and analysis of Twitter during 2015 Chennai flood towards disaster management." In Procedia computer science on Elsevier, 2017, pp. 350–358.
[51] Laylavi, Farhad and Rajabifard, Abbas and Kalantari, Mohsen. "A multi-element approach to location inference of twitter: A case for emergency response." In ISPRS International Journal of Geo-Information on MDPI, 2016, pp. 56.
[52] Chowdhury, Jishnu Ray and Caragea, Cornelia and Caragea, Doina. "On identifying hashtags in disaster twitter data." In Proceedings of the AAAI Conference on Artificial Intelligence, 2020, pp. 498–506.
[53] Zou, Lei and Lam, Nina SN and Cai, Heng and Qiang, Yi. "Mining Twitter data for improved understanding of disaster resilience." In Annals of the American Association of Geographers on Taylor \& Francis, 2018, pp. 1422–1441.
[54] Mohanty, Somya D and Biggers, Brown and Sayedahmed, Saed and Pourebrahim, Nastaran and Goldstein, Evan B and Bunch, Rick and Chi, Guangqing and Sadri, Fereidoon and McCoy, Tom P and Cosby, Arthur. "A multi-modal approach towards mining social media data during natural disasters-A case study of Hurricane Irma." In International journal of disaster risk reduction on Elsevier, 2021, pp. 1422–1441.
[55] Ni, Wenjun and Shu, Jia and Song, Miao}. "Location and emergency inventory pre-positioning for disaster response operations: Min-max robust model and a case study of Yushu earthquake." In Production and Operations Management on Wiley Online Library, 2018, pp. 160–183.
[56] Asokan, G Vaithinathan and Vanitha, Asokan. "Disaster response under One Health in the aftermath of Nepal earthquake, 2015." In Journal of epidemiology and global health on Elsevier, 2017, pp. 91–96.
[57] Kankanamge, Nayomi and Yigitcanlar, Tan and Goonetilleke, Ashantha and Kamruzzaman, Md. "Determining disaster severity through social media analysis: Testing the methodology with South East Queensland Flood tweets." In International journal of disaster risk reduction on Elsevier, 2020, pp. 101360.
[58] Hara, Yusuke. "Behaviour analysis using tweet data and geo-tag data in a natural disaster." In Transportation Research Procedia on Elsevier, 2015, pp. 399–412.

[59] Algur, Siddu P and Venugopal, S. “Classification of disaster specific tweets-a hybrid approach.” In 2021 8th International Conference on Computing for Sustainable Global Development (INDIACom) on IEEE, 2021, pp. 774–777.
[60] Kumar, Shamanth and Barbier, Geoffrey and Abbasi, Mohammad and Liu, Huan. “Tweettracker: An analysis tool for humanitarian and disaster relief.” In Proceedings of the international aaai conference on web and social media, 2011, pp. 661–662.
[61] Suwaileh, Reem and Elsayed, Tamer and Imran, Muhammad and Sajjad, Hassan. “When a disaster happens, we are ready: Location Mention Recognition from crisis tweets.” In International Journal of Disaster Risk Reduction on Elsevier, 2022, pp. 103107.
[62] Choi, Seonhwa and Bae, Byunggul. “The real-time monitoring system of social big data for disaster management.” In Computer science and its applications on Springer, 2015, pp. 809–815.
[63] Mendon, Shalak and Dutta, Pankaj and Behl, Abhishek and Lessmann, Stefan, “Mendon, ShalakA Hybrid approach of machine learning and lexicons to sentiment analysis: enhanced insights from twitter data of natural disasters and Dutta, Pankaj and Behl, Abhishek and Lessmann, Stefan.” In Information Systems Frontiers on Springer, 2021, pp. 1145–1168.
[64] SpringerWang, Zheye and Lam, Nina SN and Obradovich, Nick and Ye, Xinyue. “Are vulnerable communities digitally left behind in social responses to natural disasters? An evidence from Hurricane Sandy with Twitter data.” In Applied geography on Elsevier, 2019, pp. 1–8.
[65] Truong, Brandon and Caragea, Cornelia and Squicciarini, Anna and Tapia, Andrea H. “Identifying valuable information from twitter during natural disasters.” In Proceedings of the American Society for Information Science and Technology on Wiley Online Library, 2014, pp. 1–4.
[66] Goswami, Saptarsi and Chakraborty, Sanjay and Ghosh, Sanhita and Chakrabarti, Amlan and Chakraborty, Basabi. “A review on application of data mining techniques to combat natural disasters.” In Ain Shams Engineering Journal on Elsevier, 2018, pp. 365–378.
[67] Mukkamala, Alivelu Manga and Beck, Roman. “The role of social media for collective behaviour development in response to natural disasters.” In 26th European Conference on Information Systems on AIS Electronic Library (AISeL), 2018, pp.
[68] Nagar, Seema and Seth, Aaditeshwar and Joshi, Anupam. “Characterization of social media response to natural disasters.” In Proceedings of the 21st international conference on world wide web, 2012, pp. 671–674.
[69] Ofli, Ferda and Imran, Muhammad and Alam, Firoj. “Using artificial intelligence and social media for disaster response and management: an overview.” In AI and Robotics in Disaster Studies on Springer, 2020, pp. 63–81.
[70] Nunavath, Vimala and Goodwin, Morten. “The role of artificial intelligence in social media big data analytics for disaster management-initial results of a systematic literature review.” In 2018 5th International Conference on information and communication technologies for disaster management (ICT-DM) on IEEE, pp. 1–4.