Disaster Support Knowledge Extraction System

A project report submitted in partial fulfillment of the requirements for the degree of

Bachelor of Engineering
in
Computer Engineering

By

Nitin Chaube (8593)

Aniruddha Chaudhari (8594)

Justin Nadar (8621)

Under the guidance of
Prof.Supriya Kamoji

(Computer Science)

DEPARTMENT OF COMPUTER ENGINEERING
Fr. Conceicao Rodrigues College of Engineering, Bandra (W), Mumbai - 400050
University of Mumbai

(2021-2022)

This work is dedicated to my family.
I am very thankful for their motivation and support.

Internal Approval Sheet

CERTIFICATE

This is to certify that the project entitled **" Disaster Support Knowledge Extraction system"** is a bonafide work of **Nitin Chaube (8593), Justin Nadar(8621), Aniruddha Chaudhari (8594)** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **Bachelor Of Engineering** in **Computers.**

Prof. Supriya Kamoji
Supervisor/Guide

Dr. B.S. Daga	Dr.Srija Unnikrishnan
Head of Department	Principal

Approval Sheet

Project Report Approval

This project report entitled by **Disaster Support Knowledge Extraction system** by **Justin Nadar, Nitin Chaube, Aniruddha Chaudhari** is approved for the degree of Bachelor of Engineering.

Examiners

1.____________________________

2.____________________________

Date: 30/04/2022

Place: Mumbai

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Aniruddha Chaudhari	(Roll No. 8594) **(sign)**	______________
Justin Nadar	(Roll No. 8621) **(sign)**	______________
Nitin Chaube	(Roll No. 8593) **(sign)**	______________

Date: April 30, 2022

Acknowledgments

We take great pleasure in presenting this report on Analysis of growth and planning of urbanization and correlated changes in natural resources. We take this opportunity to express our sincerest thanks towards our project guide Prof. Supriya Kamoji, Fr. C.R.C.E, Bandra (W), Mumbai, for providing the technical guidelines, and the suggestions regarding the line of this work. We enjoyed discussing the work progress with him during our visits to department.

We thank Dr. B.S. Daga, Head of Computers Dept., Principal and the management of C.R.C.E., Mumbai for encouragement and providing necessary infrastructure for pursuing the project.

We also thank all non-teaching staff for their valuable support, to complete our project.

Nitin Chaube (Roll No. 8593)

Aniruddha Chaudhari (Roll No.8594)

Justin Nadar (Roll no. 8621)

Date: April 30, 2022

Table of Contents

Abstract

This research is conducted with objective of creating and presenting a system called FloodKnow, which is data extraction and visualization framework with capability of filtering, classifying and visualizing online social media data, to provide accurate information about ongoing events for decision making in disaster scenarios such as flood. Now a days social media platforms such as Twitter have received a great deal of attention as human sensors, revealing the information regarding the ongoing events such as natural disaster. Filtering and Analyzing disaster related situation updates from a large pool of data is still a challenge and has attracted a lot of research attention. Research works have been conducted in filtering and analyzing the online social media content for natural disaster but a complete system to provide a real-time a summarized information to the decision makers is still at its infant stage. In this work a complete framework is proposed which uses Twitter as a sensor to gather the information of ongoing disaster event such as flood, information is filtered and processed using natural language processing, Deep learning for classifying priority tweets and knowledge graph for real-time visual representation of gathered information for the decision makers. At the end of this process, the data is presented graphically which narrate the local and global storyline of the affected stakeholders and hotspot maps to depict the area in

which natural disaster are concentrated. Floodknow has proved to be on par with state of art disaster extraction systems, which helps in deep understanding the disaster development by visual approach. Among the disaster domain, we represent the knowledge graph for flooding event.

List of Figures

List of Table

Chapter 1

Introduction

Flood management (Disaster) involves three phases: Preparedness, Response, and Recovery. The preparedness phase comes into place when an emergency or a disaster is likely to take place. It corresponds to preparatory activities prior to a disaster to save lives and help response and rescue operations, such as stocking food and water, posting emergency contacts, and preparing evacuations. With plans and strategies developed beforehand, the response phase mainly puts them into action. Response activities happen during a disaster, usually involving evacuating threatened areas, search and rescue efforts, shelter management, and humanitarian assistance. After a disaster, the recovery phase refers to repair and reconstruction efforts to return to a normal or even better functionality level. Recovery actions usually include debris clean-up, precise damage assessment, and infrastructure reconstruction, as well as financial assistance from government agencies and insurance companies.

In social, political and even when tragedy happens, social media has played an important role in recent years, and it has also become the most influential contact media. In most situations, traditional communication fails during a disaster. During the tragedy, social media such as Facebook, Tweeter, Google, and so on made communication easier in this sort of situation. Gives the real-time information about the pre and post disaster/flood to avoid the losses. Messages and posts in the social media give more information to the people who are living in that area. This real time nature of social media makes it an attractive tool for disaster management, as both victims and officials can put their problems and solutions at the same place in real time. It provides the location and severity of the disaster so that the resource and rescue teams can be deployed to the stakeholders. Researchers have started to explore Twitter as a tool for disaster management.

The large amount of data generated during a short span of time during disaster. Even when there are resources available for help, many requests could go unnoticed as they are not rooted through proper channels. By collecting and analyzing such generated requests will help the government authorities and NGOs to deploy the rescue teams. [1]

During the time of disaster, people often post their real-time experiences and local news on social media to inform others. Many rescue agencies monitor this type of posts regularly to locate disasters and reduce the risk of lives. However, it is impossible for humans to manually check the large number of posts and identify disasters in real-time. For this purpose, many research works have been proposed to automate the filtering of huge mass of messages. [2]
Twitter is a micro blog where crowd approved knowledge such as text messages, photographs and audio clips are sent. Since users write small messages, they regularly send it and check for retweets from others. Twitter updates for disastrous events includes storms, heavy rainfall, earthquakes, fire, traffic jams etc. A lot of work has been done to detect events, both social as well as disastrous from Twitter messages. Most work for disastrous event detection systems are confined to detect whether a tweet is related to the disaster or not, based on textual and visual content. The disaster related tweets are further used to warn and inform people about precautionary measures. These tweets are viewed not only as an awareness platform, but a place where people can ask for help during disaster. The tweets asking for help need to be separated from other tweets related to the disaster. These tweets then can be used to guide the rescue personnel.[3]

Many studies show that huge amount of accurate and real time data is obtained from Twitter than collected through traditional data collection methods. Despite of its privileges, this real-time information provides interesting challenges, due to very high volume of messages, lack of metadata, verifying the veracity of the posted information, and missing geolocations. [4] Disaster respondents do not have enough bandwidth and time to manually monitor huge amount posts collected on twitter during the situations like disaster and take appropriate action. Understanding the purpose of the posts such as is it about requesting for rescue operation, reporting about the situation, calling for donation, infrastructure damage, praying for affected people and understanding the exact location of the people who need help is the biggest challenge. [5] This huge amount of information gathered from twitter has to be classified into relevant and non-relevant information. Relevant class of information can be further categorized into high priority information such as injured and help required and low priority such as praying for affected.

Social media such as Twitter provides three location information fields for sharing a user's location: (1) User location; (2) Place name and (3) Geo-coordinate. The user location field has 280-character in which the user can write their home location information while creating their profile. This field is compulsory to the user and the user can write any arbitrary words or leave it blank. In many instances, user does not want to reveal their location so they write meaningless words that might not refer to any location name.[7] It is analyzed that 34% of users do not want to reveal their "user location" Information [8]. Users use city level or below city level location names in their user location field. However, this field cannot be treated as the current location of the user as it is entered at the time of creating their profile as it is most of the time not updated by the users regularly. The second field is for the "place name," which can be attached to a tweet message when it is posted. The place name is represented by a location name with an array of the latitude-longitude pair in the form of the location's boundary coordinates. But these place names are predefined on the Twitter database, but it does not provide granular location information [7]. Research found that only 47.33% of tweets contain place names. However, 12% of those place names are not correct in terms of their spatiotemporal information [9]. The third field provided by Twitter is for the "geo-coordinates" (geographical footprints of latitude and longitude) which can be attached at the time of posting a tweet using a GPS- (Global Positioning System) enabled device. Most of the researchers have considered geo-coordinates as the most explicit information [10][11][12]. However, most users do not tweet with geo-coordinate information. Some researchers determined that only 0.42%, 3.17%, and 7.90%of tweets respectively are geo-tagged [6][7][13][14]. Researcher further reported that although geo-coordinates are the most precise location information, they are not always authentic in terms of their spatiotemporal information if the tweet is posted from some other location [7]. Hence, all three location information fields, available in twitter and user profiles, have their own constraints and cannot be completely relied on.

Now a days many users provide location information in tweet texts which is vitally important and authentic source of geographic evidence as it represents the location information of any event or user during emergencies. Most commonly used methods for location extraction are gazetteer-based approach [15][16][17][18] and the Named Entity Recognition (NER) based approach [19] [20] [21]. Gazetteer is a location names corpus (e.g., GeoNames, http://geonames.org). In the gazetteer-based approach, the words of tweets are searched upon in the gazetteer to find the location names. However, there are some problems with this approach: (i) the unavailability of gazetteers for all the locations and (ii) a location name mentioned in the tweet may have some other non-geographic meaning in the context of a text e.g., the word "Pentagon" may refer to a location name in US or it may also be used in another context. The other problem with this approach is the geo-ambiguity (distinct locations have the same name, e.g., Chamba in India is a town in Himachal Pradesh and Uttarakhand also. The second approach is Named Entity Recognition (NER). The NER technique generally tokenizes the tweet into different entities using language-specific part-of-speech tagging.[7]

There are numerous cooperation teams which are prepared to help the people at the time of disaster, but the cooperation among different teams is weak. Besides, literature emphasizes the establishment of centralized cooperative emergency system, which can serve every needy request during disaster. For this a graphical representation which can visualize every detail of disaster information so that no request can go unnoticed and can have cooperation among all the rescue teams [22]. Some study shows that disaster include heterogeneous data and lack interoperability. In particular, the case of social media data related to disasters, there are several issues, where the source and format of data are different because huge amount of data are collected by different organizations. This study proposes a visualization tool, knowledge graph to resolve the heterogeneity among various disaster data. Knowledge graph is used to assist, solve, and manage disaster problems [23].

Knowledge graph can be designed for a data extraction system with the capability of filtering social media data, to improve community resilience and decision-making in disaster scenarios. knowledge graphs (KG) can be used to connect insights, possible to generate real-time visual information about such disasters and affected stakeholders, to better the crisis management process [24].

1.1 Aim

To better the crisis management process, by disseminating such information to both relevant authorities and population alike, this generated Knowledge graph can provide real-time visual information about disasters like flood, earthquake and wildfire.

1.2 Motivation

Learning to code is incredibly rewarding but can also be difficult and frustrating. The strongest assets you can have as a student are a desire to build, a problem-solving mind, and persistence in the face of setbacks. At any given time and place, if a disaster occurs, the platform where it's first mentioned is twitter. This data could be used to develop a Model which can be used by various organizations to lend support to the affected as early as possible.

1.3 Objective

- To identify a tweet whether the tweet is related to disaster or not
- To do a multiclassification on a dataset related to different classes
- To use different models and find the best one giving high accuracies
- To make a model for NER to extract useful information from a tweet e.g., locations
- To extract tweets from twitter related to disasters from Indian regions
- To Create Knowledge Graph for better visualization of all the information

Chapter 2

Literature Review

2.1 Flood management during disaster:

During disasters, social media such as Twitter get flooded with messages for aid and help. These messages can provide valuable information on people stranded in a region and the amount of relief material needed. Along with these messages, people also post images that reflect the damage caused. These images are a very good source to discover the gravity of the damage caused or the amount of help needed in a particular area. Information obtained from these images can help assist relief work. Floods can cause tremendous damage to life while destroying property and infrastructure. Early response to floods becomes important to rescue the people stranded in floods and provide necessary aid to the people. Images from remote sensing technologies such as drones and social media can play a great role in helping in the early response. Pictures of inundated areas can help in ascertaining the severity of floods and identify the scale of damage to property. Tweets from Twitter can be used to identify human stranded and help rescue providers reach them or provide aid to them.

2.1.1 Flood Analysis using Twitter (social media):

Social media is a platform to express one's view in real time. This real time nature of social media makes it an attractive tool for disaster management, as both victims and officials can put their problems and solutions at the same place in real time. Both academia and industry have started to explore Twitter as a tool for disaster management.

Dhanya et al. (2021) have analyzed that sheer amount of data generated during a short span of time during disaster. Even when there are resources available for help, many requests could go unnoticed. So, by collecting the generated requests for help and resource availability and plot the location in the map. Request data shall be analyzed using three machine learning algorithms called initial filtering and will be passed through natural language processing to match needs and offers within a given geographic boundary. The system is working with 90% accuracy [1].

Maulana et al. (2021) aims to classify these tweets into "flooded" and "not flooded" predictions with the tweets and geospatial features. The model proposed for classifying is BERT-MLP. Bidirectional Encoder from Transformers (BERT) is used in the pre-trained model to classify these tweets and Multi-Layer Perceptron (MLP) is used to classify geospatial features. The scenario designed for the model focuses on the pre-processing of tweets as follows without stop word removal, without stemming, with both, and without both. *Once classified, the tweet will be visualized into a two-dimensional interactive map. The best scenario results have an accuracy of 82% in scenarios without stemming and with stop word removal [25].

Chanda et al. (2021), in this research work, they have explored the efficacy of BERT embeddings on predicting disaster from Twitter data and compare these to traditional context-free word embedding methods (GloVe, Skip-gram, and FastText). They have used both traditional machine learning methods and deep learning methods for this purpose. Both quantitative and qualitative results for this study. The results show that the BERT embeddings have the best results in disaster prediction task than the traditional word embeddings [2].

Madichetty et al. (2021) Fine-tuned BERT model is used to extract the linguistic, syntactic and semantic features which help deep understanding of the informative text present in the multi-modal tweet. On the other hand, the fine-tuned DenseNet model is used to extract the sophisticated features from the image. Different experiments are performed on a vast number of data-sets such as Hurricane Harvey, Hurricane Irma, Hurricane Maria, California Wildfire, Sri Lanka floods and Iraq–Iran Earthquake. Experimental results demonstrate that the proposed method outperforms the state-of-the-art method on different parameters. It is the first attempt, to the best of our knowledge, to detect multi-modal informative tweets using the combination of fine-tuned BERT and DenseNet models, where at least any text or image is informative during the disaster [26].

Kersten J. et al. (2020), study focuses on the design and evaluation of a generic workflow for Twitter data analysis that leverages the additional information to characterize crisis events more comprehensively. The workflow covers data acquisition, analysis and visualization, and aims at the provision of a multifaceted and detailed picture of events that happen in affected areas. This is approached by utilizing agile and flexible analysis methods providing different and complementary views on the data. Utilizing state-of-the-art deep learning and clustering methods, they are interested in the question, whether their workflow is suitable to reconstruct and picture the course of events during major natural disasters from Twitter data. Experimental results obtained with a data set acquired during hurricane Florence in September 2018 demonstrate the effectiveness of the applied methods but also indicate further interesting research questions and directions [27].

Singh Jyoti Prakash et al. (2019) The developed system takes tweets as inputs and categorizes them into high or low priority tweets. User location of high priority tweets with no location information is predicted based on historical locations of the users using the Markov model. The system is working well, with its classification accuracy of 81%, and location prediction accuracy of 87% [3].

Madichetty et al. (2018) used a re-ranking feature selection algorithm based on the α2-statistic feature selection algorithm and maximum frequency algorithm for identifying the available and requirement of resources during a disaster. Their method is compared with the baseline methods such as α2-statistic feature selection algorithm, maximum frequency algorithm individually and BOW model on Nepal earthquake. Their method outperforms the baselines by using the precision parameter [28].

Yang et al. (2018), These researchers have proposed a Twitter data credibility framework to identify trustworthy events from massive social media posts under a disaster management scenario. Individual events with disaster situation awareness topics including power, shelter, and damage were identified. To complement the small proportion of geo-tagged tweets, location information was extracted from tweets by constructing a local gazetteer and merged events occurring in a specific spatiotemporal range. Firstly, the credibility score for each tweet was calculated based on the information contained in its text and URL. Secondly, the accumulated credibility score for each event was calculated based on the number of tweets and retweets associated with the same event [29].

Utsu et al. (2017) developed a web application prototype for the collection of tweets and

discriminating them into disaster-related and non-disaster related. In addition, to automatically collect disaster-related information by focusing on tweets posted immediately after retweeting the news post based on keyword-based method to distinguish whether the tweets are related to the disaster [30].

Kwon et al. (2016) quantified the risk level of disaster occurrences in Seoul by analyzing tweet text. The usage frequency of keyword - flood, inclusion of disaster signword, and degree of adverbs present in tweets were used to quantify the risk levels. They also proposed tools to visualize these risk levels based on tweet locations with the help of a time series.

2.1.2 Flood Mapping using Knowledge Graph:

Xu, H et al. (2021), analyzed the knowledge map of intelligent emergency and visualized in this study. There are numerous cooperation teams, but the cooperation among different teams is weak. Besides, foreign literature emphasizes the establishment of intelligent emergency systems, the improvement of emergency industry technology and the intelligent emergency education. There are some deficiencies in these aspects, which can be improved in the future development [22].

Son et al (2021), authors have studied that Disaster data depend on the domain by disaster type and include heterogeneous data and lack interoperability. In particular, the case of open data related to disasters, there are several issues, where the source and format of data are different because various data are collected by different organizations. Moreover, the vocabularies used for each domain are inconsistent. This study proposes a knowledge graph to resolve the heterogeneity among various disaster data and provide interoperability among domains. Among disaster domains, they describe the knowledge graph for flooding disasters using Korean open datasets and cross-domain knowledge graphs. Furthermore, the proposed knowledge graph is used to assist, solve, and manage disaster problems [23].

Boné et al. (2020), researchers aimed at creating and presenting DisKnow, a data extraction system with the capability of filtering and abstracting tweets, to improve community resilience and decision-making in disaster scenarios. Nowadays most people act as human sensors, exposing detailed information regarding occurring disasters, in social media. Through a pipeline of natural language processing (NLP) tools for text processing, convolutional neural networks (CNNs) for classifying and extracting disasters, and knowledge graphs (KG) for presenting connected insights, it is possible to generate real-time visual information about such disasters and affected stakeholders, to better the crisis management process, by disseminating such information to both relevant authorities and population alike. DisKnow has proved to be on par with the state-of-the-art Disaster Extraction systems, and it contributes with a way to easily manage and present such happenings [24].

Ni et al. (2019), have explored the problem of generating storyline from huge amount of web information and propose a knowledge graph-based disaster storyline generating framework. The deep learning technique and the semi-supervision method are first used to extract two kinds of triples from disaster news. Then the evolution of disasters is summarized in a way of graph-based which can clearly show the sketch of disaster development. Finally, the location entities are extracted by the name entity recognized model. Compared with commonly used text summarization storyline, the proposed framework can provide a better user situational awareness and deeper understanding on disaster events by the presentation of graph [32].

Chapter 3
Problem Statement

Problem Statement is to create a Knowledge Graph to represent the different natural or man-made disasters occurring in India in a dynamic graphical representation so that government bodies or any social bodies could get all the information about the disaster and provide necessary relief to the affected people. We are making our system model using tweets of twitter on the disasters.

Our Problem statement consists of making of three deep learning model for classifying the tweets as informative or not, to do the To Create Knowledge Graph for better visualization of all the information multiclass classification on the tweets and to extract the important entities from tweets using NER model

After processing the tweets and extracting all the necessary information, this information is automatically fed to Neo4j graph using python's py2neo library which then can be queried as per our needs to get the information that we need.

Chapter 4

Proposed System

The block diagram of the proposed framework is depicted in figure 1. Each block is briefly described next:

Our main contributions are summarized as follows:

(1) Using the keywords such as Flood, Water logging, disaster, calamity etc. scraped the Indian Flood tweets.

(2) A text preprocessing module to remove noisy textual features.

(3) Processed Tweets are classified into Binary (relevant and non-relevant) and multiclass classification (affected individual, caution and advice, donation and volunteering, Structure and utility damage, sympathy and support and non-humanitarian)

(4) An NER-based geoparsing strategy (toponym extractor) where location of the tweets is extracted.

(4) A Geocoder to query Google Maps API with each toponym, thus presenting results in latitude and longitude values.

(5) A Knowledge Graph is used to graphically depict the overall local and global storyline from processed tweets.

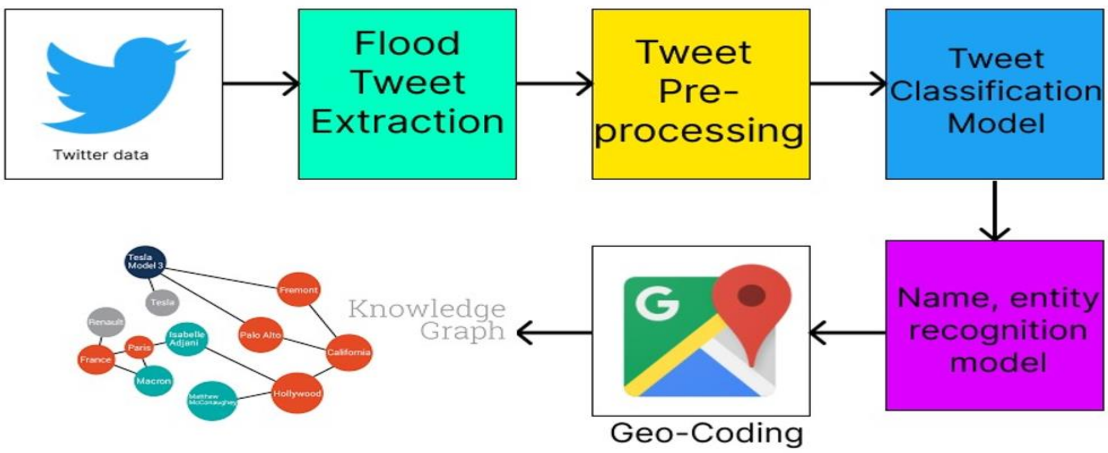

Figure 1: Framework Architecture

The system consists of the following modules, which have been described in the subsequent sections.

– Data collection

– Data pre-processing

– Event classification (Binary and Multi class)

– Location estimation

– Knowledge Graph

4.1 Data collection:

We have used the consolidate eight human-annotated datasets and provide 166.1k and 141.5k tweets for informativeness and humanitarian classification tasks, respectively. These consolidated datasets will help train more sophisticated models [33]. Eight datasets that were labeled for different disaster response classification tasks and whose labels can are informativeness and humanitarian information type classification. Brief overview of dataset used for consolidation are:

CrisisLex is one of the largest publicly available datasets, which consists of two subsets, i.e., CrisisLexT26 and CrisisLexT6 [34]. CrisisLexT26 comprises data from 26 different crisis events that took place in 2012 and 2013 with annotations for informative vs. not informative as well as humanitarian categories (six classes) classification tasks among others. CrisisLexT6, on the other hand, contains data from six crisis events that occurred between October 2012 and July 2013 with annotations for related vs. not-related binary classification task.

CrisisNLP is another large-scale dataset collected during 19 different disaster events that happened between 2013 and 2015, and annotated according to different schemes including classes from humanitarian disaster response and some classes related to health emergencies (Imran, Mitra, and Castillo 2016) [35].

SWDM2013, SCRAM2013, Disaster Response Data (DRD), Disasters on social media (DSM), CrisisMMD and AIDR are the additional five more datasets that are used for the consolidation. These datasets size is not as big as first two datasets [33].

4.2 Data pre-processing:

Tweets scraped directly from twitter using the keywords contain different types of noise and redundancies, such as emoticons, user mentions, Internet links etc. An efficient data pre-processing is needed to use these tweets for any meaningful purpose.

4.3 Event classification:

Hash tags and keywords in tweets help to extract tweets related to a target event. Consolidated Dataset by compiling 8 different datasets is prepared however, some of these tweets may be referring to general information such as "Floods have become a regular occurrence in Kerala". The above tweet refers to natural disaster such as floods, which may be a target event, but it does convey real time update of the event all the time. Hence, tweet classification system is required to filter out the type of tweet related to the event. The compiled dataset is classified into two categories Binary class (informative and non-informative) and multiclass category (6 classes).

4.3.1 Binary Classification:

When tweets are collected from a huge pool of twitter data using keywords of related flood still, we get tweets which are not related to the event hence the binary classification of informative and non-informative is required.

Multiclass classification:

Dataset is also classified into multiple classes as given below:

Affected individual: Message regarding the victims who are affected. E.g., 10 people are struck in water logging near Kalina.

Caution and advice: Warning given about a related incident e.g. Flooded neighbours in Bandra and its approaching near high tide

Donation and volunteering: Tweets which are offering help. E.g., Shelter and food are arranged at shiva Hall near highway

Infrastructure and utility damage: Information regarding the damage caused to the buildings and resources. Building collapsed in Mumbra near station

1. **Help and support:** Required help to the victims. E.g., @Maharastra Government please send volunteers to the Kalina west some people are stuck.
2. **Non-humanitarian:** Messages not related to flood e.g. My heart is flooded with lots of love.

4.3.2 Models and Architectures used for classification:

Deep neural networks (DNNs) are ideally suited for classifying a crisis-related tweets. They are usually trained with large pool of data and have the elasticity to learn and modify from new batches of labeled data without requiring retraining from the beginning. Due to their distributed word representation, they generalize well and make better use of the previously labeled data from other events to speed up the learning process. DNNs prevent the need of manually crafting features and automatically learn hidden features as distributed dense vectors, which have shown to advantage various NLP tasks. [36]

4.3.3 Convolutional Neural Network:

This classification model is based on the CNN architecture. We used similar architecture as proposed by (Nguyen et al. 2017) [36]. Convolutional Neural Network Figure 2 shows CNN model for classifying tweets into *binary class and multi class* for a crisis event. The architecture of our model is similar to the one proposed in (Nguyen et al. 2017) [36]. For distributed representation of words, we first construct a vocabulary V from the training set by selecting T most frequent words. Each word in the vocabulary is then represented by aD dimensional vector in a shared look-up table $L \in \mathrm{R}\ |V|\times D$, which is considered a model parameter to learn. We can initialize L randomly or using pretrained word embedding vectors like word2vec. Given an input tweet $\mathbf{s} = (w1, \cdot\ \cdot\ \cdot, wT)$, we first transform it into a feature sequence by mapping each word token $wt \in \mathbf{s}$ to an index in L. The look-up layer then creates an input vector xt ∈ RD for each token w_t, which are passed through a sequence of convolution and pooling operations to learn high-level feature representations.

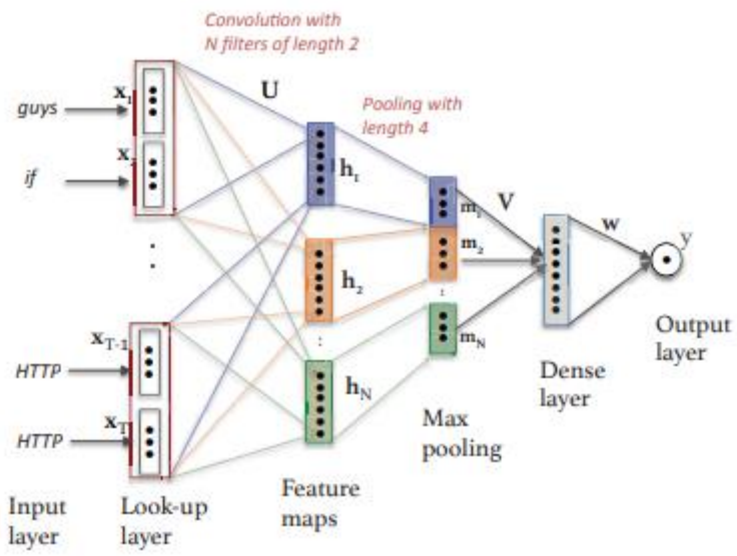

Figure 2: Convolutional neural network on a tweet: "guys if know any medical emergency around balaju area you can reach umesh HTTP doctor at HTTP"

The convolutional layer is the core building block of a CNN, and it is where the majority of computation occurs. It requires a few components, which are input data, a filter, and a feature map.

Pooling layers, also known as down sampling, conducts dimensionality reduction, reducing the number of parameters in the input. Similar to the convolutional layer, the pooling operation sweeps a filter across the entire input, but the difference is that this filter does not have any weights. Instead, the kernel applies an aggregation function to the values within the receptive field, populating the output array.

Fully-Connected Layer performs the task of classification based on the features extracted through the previous layers and their different filters. While convolutional and pooling layers tend to use ReLu functions, FC layers usually leverage a SoftMax activation function to classify inputs appropriately, producing a probability from 0 to 1.

Transformer models: pre-trained model as shown in figure 3 have achieved state-of-the-art performance on natural language processing tasks and have been adopted as feature extractors for solving down-stream tasks such as question answering, and sentiment analysis. Though the pre-trained models are mainly trained on non-Twitter text, we hypothesize that their rich contextualized embeddings would be beneficial for the disaster domain [33].

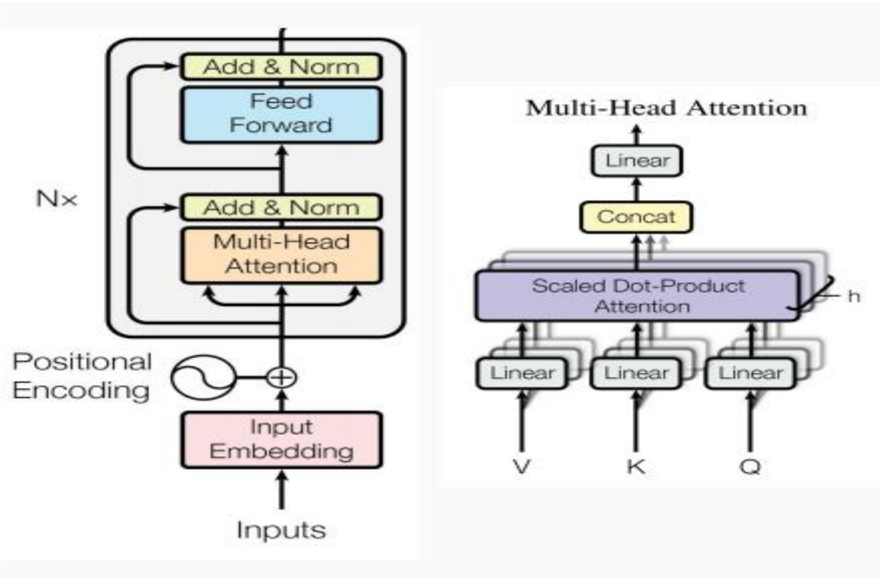

figure 3: Pretrained Transformer Encoder

BERT is basically an Encoder stack of transformer architecture. A transformer architecture is an encoder-decoder network that uses self-attention on the encoder side and attention on the decoder side. $\text{BERT}_{\text{BASE}}$ has 12 layers in the Encoder stack while $\text{BERT}_{\text{LARGE}}$ has *24 layers in the Encoder stack*. These are more than the Transformer architecture described in the original paper (*6 encoder layers*). BERT architectures (BASE and LARGE) also have larger feedforward-networks (768 and 1024 hidden units respectively), and *more attention heads (12 and 16 respectively)* than the Transformer architecture suggested in the original paper. It contains *512 hidden* units and 8 attention heads. BERTBASE contains 110M parameters while BERTLARGE has 340M parameters.

4.4 Location estimation:

In this work, we are particularly interested in the locations described in the content of tweets.

While both the news and literature told us that people used Twitter and other social media plat-forms to request for help and share information, we still do not know how specifically people describe locations in social media messages during this natural disaster. Manually analyzing the more than 100k tweets is tweets is practically impossible. We randomly select 1,00 tweets for testing.

BERT NER is a fine-tuned BERT model as shown in figure 4 that is ready to use for Name Entity Recognition and achieves good for the NER task. It has been trained to recognize four types of entities location (LOC), organizations (ORG), person (PER) and Miscellaneous (MISC).

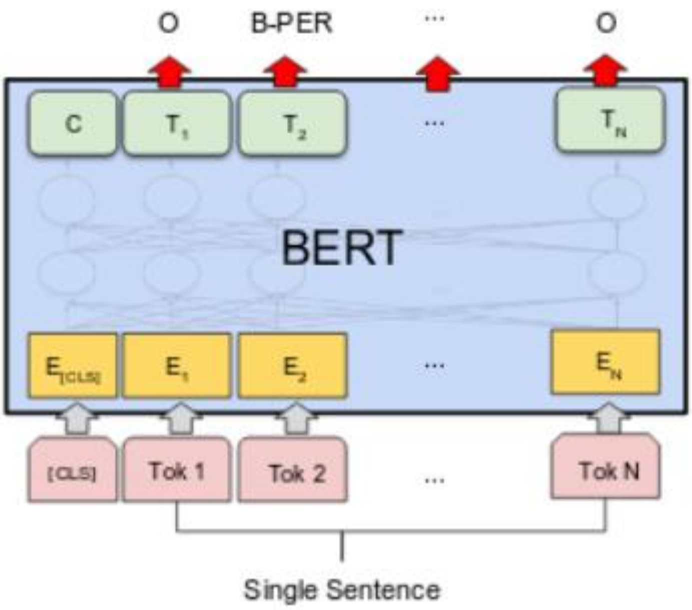

Figure 4: NER BERT MODEL

When several tweets are collected, are fed into the classification model to obtain a series of predicted entities ˆy. Prior to the conversion of predictions into useful toponyms, words classified with O and PER tags are discarded (their presence is required in the training stage to capture entity tag transitions at the CRF output layer, but they are not needed for toponym identification). Furthermore, those classified as LOC and ORG are identified and joined to form a sentence, consequently creating a toponym, which is used to request a Google API location. Responses from Google are geocoded in JSON format to form an address with geographic coordinates.

4.5 Knowledge Graph:

The knowledge base consists of a dynamic knowledge graph [37]. This allows us to structure entities as a knowledge graph in which new nodes and relationships are constantly being added as new feed is detected, to expand and upgrade its knowledge in real-time. This knowledge is disaster-centered, as disaster occurrences are represented by specific nodes which then connect to locations, dates, people, organizations, etc. In addition to having several node types, such as the date and locations, each node also has several attributes. This allows us for expanding in real time when dealing with various types of data associated with entities. As for its management, our knowledge base uses the Neo4j (https://neo4j.com/) graph database management system [4].

Lastly, the classified tweets must be included in our knowledge graph. This step involves all validations regarding the previous existence of our gathered data, as well as the creation of new nodes and relationships between them, to express the identity of

the extracted disasters in a way that facilitate human consumption and future system integrations. To further deal with social media unreliability, all disaster nodes have an attribute which is incremented when the same disaster is detected from different tweets. This attribute can then be used as a detection threshold, to allow for better filtering of relevant tweets. If new information regarding an already existing disaster is extracted from the previous steps, it is also validated and then connected with that event. This mechanic allows for our system to be continually learning and updating, expanding its knowledge, and bettering the knowledge it already has.

Chapter 5

Technologies Used

5.1 Software Requirements

5.1.1 Python

Python is a powerful high-level, object-oriented programming language created by Guido van Rossum. It has simple easy-to-use syntax, making it the perfect language for someone trying to learn computer programming for the first time. Python is a general-purpose language. It has wide range of applications from Web development (like: Django and Bottle), scientific and mathematical computing (Orange, SymPy, NumPy) to desktop graphical user Interfaces (Pygame, Panda3D). The syntax of the language is clean and length of the code is relatively short. It's fun to work in Python because it allows you to think about the problem rather than focusing on the syntax.

5.1.2 Jupyter Notebook

The notebook extends the console-based approach to interactive computing in a qualitatively new direction, providing a web-based application suitable for capturing the whole computation process: developing, documenting, and executing code, as well as communicating the results.

5.1.3 Neo4j

Neo4j is an open-source, NoSQL, native graph database that provides an ACID-compliant transactional backend for your applications that has been publicly available since 2007.Neo4j is a native graph database, which means that it implements a true graph model all the way down to the storage level. The data isn't stored as a "graph abstraction" on top of another technology, it's stored just as your whiteboard it. This is important because it's the reason why Neo4j outperforms other graphs and stays so flexible. Beyond the core graph, Neo4j provides what you'd expect out of a database; ACID transactions, cluster support, and runtime failover. This stability and maturity are why it's been used in production scenarios for large enterprise workloads for years.

Following are the notable features of Neo4j −

- Data model (flexible schema) − Neo4j follows a data model named native property graph model. Here, the graph contains nodes (entities) and these nodes are connected with each other (depicted by relationships). Nodes and relationships store data in key-value pairs known as properties. In Neo4j, there is no need to follow a fixed schema. You can add or remove properties as per requirement. It also provides schema constraints.
- ACID properties − Neo4j supports full ACID (Atomicity, Consistency, Isolation, and Durability) rules.
- Scalability and reliability − You can scale the database by increasing the number of reads/writes, and the volume without effecting the query processing speed and data integrity. Neo4j also provides support for replication for data safety and reliability.
- Cypher Query Language − Neo4j provides a powerful declarative query language known as Cypher. It uses ASCII-art for depicting graphs. Cypher is easy to learn and can be used to create and retrieve relations between data without using the complex queries like Joins.
- Built-in web application − Neo4j provides a built-in Neo4j Browser web application. Using this, you can create and query your graph data.
- Drivers − Neo4j can work with −
    - REST API to work with programming languages such as Java, Spring, Scala etc.
    - Java Script to work with UI MVC frameworks such as Node JS.
    - It supports two kinds of Java API: Cypher API and Native Java

API to develop Java applications. In addition to these, you can also work with other databases such as MongoDB, Cassandra, etc.

- Indexing − Neo4j supports Indexes by using Apache Lucence.

5.1.4 Libraries Used

5.1.4.1 PyTorch

PyTorch is an open-source machine learning library for Python that was developed mainly by Facebook's AI research group. PyTorch supports both CPU and GPU computations and offers scalable distributed training and performance optimization in research and production. Its two high-level features include tensor computation (similar to NumPy) with GPU acceleration and deep neural networks built on a tape-based autodiff system.

5.1.4.2 Torchvision

The torchvision package consists of popular datasets, model architectures, and common image transformations for computer vision.

5.1.4.3 NumPy

NumPy is a library for the Python programming language, adding support for large, multidimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

5.1.4.4 Scikit-learn

Scikit-learn (formerly scikits.learn) is a free software machine learning library for the

Python programming language. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy. Clustering of unlabeled data can be performed with the module sklearn.cluster. From sklearn.cluster we can import KMean which is used in our application.

5.1.4.5 Pickle

The Pickle module is a library provided by Python. The pickle module implements a fundamental, but powerful algorithm for serializing and de-serializing a Python object structure. "Pickling" is the process whereby a Python object hierarchy is converted into a byte stream, and "unpickling" is the inverse operation, whereby a byte stream is converted back into an object hierarchy.

5.1.4.6 NLTK

NLTK is a leading platform for building Python programs to work with human language data. It provides easy-to-use interfaces to over 50 corpora and lexical resources such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active discussion forum.

5.1.4.7 Tensorflow

TensorFlow is an end-to-end open-source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML-powered applications.

5.1.4.8 Gensim

Gensim is a Python library for topic modelling, document indexing and similarity retrieval with large corpora.

5.1.4.9 Py2neo

Py2neo is a client library and toolkit for working with Neo4j from within Python applications and from the command line. The library supports both Bolt and HTTP and provides a high-level API, an OGM, admin tools, an interactive console, a Cypher lexer for Pygments, and many other bells and whistles.

5.1.4.10 Snscrape

Snscrape is a scraper for social networking services (SNS). It scrapes things like user profiles, hashtags, or searches and returns the discovered items, e.g., the relevant posts.
The following services are currently supported:

- Facebook: user profiles, groups, and communities (aka visitor posts)
- Instagram: user profiles, hashtags, and locations
- Reddit: users, subreddits, and searches (via Pushshift)
- Telegram: channels
- Twitter: users, user profiles, hashtags, searches, tweets (single or surrounding thread), list posts, and trends

5.1.4.11 Tweepy

Tweepy is an open-source Python package that gives you a very convenient way to access the Twitter API with Python. Tweepy includes a set of classes and methods that represent Twitter's models and API endpoints, and it transparently handles various implementation details, such as:

- Data encoding and decoding
- HTTP requests
- Results pagination
- OAuth authentication
- Rate limits
- Streams

If you weren't using Tweepy, then you would have to deal with low-level details having to do with HTTP requests, data serialization, authentication, and rate limits. This could be time consuming and prone to error. Instead, thanks to Tweepy, you can focus on the functionality you want to build. Almost all the functionality provided by Twitter API can be used through Tweepy. The only current limitation, as of version 3.7.0, is that Direct Messages don't work properly due to some recent changes in the Twitter API.

4.2 Hardware Requirements

For training the model, a GPU is preferable. The GPU used to develop this model is Nvidia GTX 1080. Minimum of 4GB RAM is required to run the code snippets. 8GB RAM was used to train this model

Chapter 6

Results and Conclusion

6.1 Results

6.1.1 Datasets

Python's snscrape module is being used to create datasets by extracting tweets from twitter. Each dataset contains anywhere around 5000-30000 tweets related to the particular event that had occurred in the past. These datasets will be used to train the deep learning model. The datasets contain the following columns- Username, Location, Tweet, Date, Followers count, Co-ordinates, Friends count, Protected status, verified status, Reply count, Like count, Retweet count

This is the code used for scraping the tweets-

```
import snscrape.modules.twitter as sntwitter
import csv
maxTweets = 30000
loc = ' 19.663280, 75.300293, 500km'
csvFile = open('Flood.csv', 'a', newline='', encoding='utf8')

#Use csv writer

csvWriter = csv.writer(csvFile)
csvWriter.writerow(['username','Location','co-ordinates','Follower
Count','Friends count','Protected','Verified','date','tweet','reply
count','like count','retweet count',''])

for i,tweet in enumerate(sntwitter.TwitterSearchScraper('(location name) +
since:2016-06-01 until:2016-08-01 -filter:links -
filter:replies').get_items()):
        if i > maxTweets :
            break

csvWriter.writerow([tweet.user.username,tweet.user.location,tweet.coordinat
es,tweet.user.followersCount,tweet.user.friendsCount,tweet.user.protected,t
weet.user.verified, tweet.date,
tweet.content,tweet.replyCount,tweet.likeCount,tweet.retweetCount])
csvFile.close()
```

The following datasets were extracted using snscrape module-

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	username	Location	co-ordinat	Follower Count	Friends count	Protected	Verified	date	tweet	reply cour	like count	retweet count		
2	vipulendr	Varanasi, India		39	179	FALSE	FALSE	2019-09-30	#BIHARfloods I request to all, pls don't take jibe a	0	0	0		
3	vipulendr	Varanasi, India		39	179	FALSE	FALSE	2019-09-30	I request to all, pls don't take jibe at #BiharFlood.	0	0	0		
4	SyedAzha	Warangal		8155	243	FALSE	FALSE	2019-09-15	SIO volunteers continues its works in flood	0	20	2		donation_and_volunteering
5	trinayon_	Moranhat, India		42	49	FALSE	FALSE	2019-09-09	Happy birthday padman @akshaykumar	0	1	0		
6	Ankurkrtw	Bengaluru, India		196	196	FALSE	FALSE	2019-08-12	Can someone please provide one-one POCâ€™s	2	1	0		donation_and_volunteering
7	Nizar____	à´¹à´°à´¿à´¹àµ¼ à´¨à´–		10621	2141	FALSE	FALSE	2019-08-11	à´ªà´² à´¸à´¸à´¸àµà´¥à´¾à´¨à´™àµà´™à´³à´¿à´²àµà´,	0	22	2		
8	saycoolr	New Delhi, India		32	57	FALSE	FALSE	2019-08-09	#AssamFlood was just a news but was wnever a h	0	0	0		not_humanitarian
9	JakirBlogg	Dalgaon, Darang (Ass		44	8	FALSE	FALSE	2019-08-02	Financial condition bahut down chal rha hai	0	0	0		
10	GhyFoodi	Guwahati,	Coordinat	448	167	FALSE	FALSE	2019-07-29	Man Vs Wild is so cool ... But if you have the	0	0	0		
11	rozeanne	1997		1264	2116	FALSE	FALSE	2019-07-28	SrBachchan sir Bihar ke sath partiality kyu	0	0	0		
12	NewsWor	India		18042	81	FALSE	TRUE	2019-07-28	Total of 1,716 villages in 56 revenue circles of 18	0	0	0		infrastructure_and_utilities_damage
13	SeniorPar	à¦®à§°à§‹à§±à¦¾, à¦		650	839	FALSE	FALSE	2019-07-27	No need to give any funds in Assam Flood. This	0	0	0		
14	JohnDeka	Nalbari,In	Coordinat	235	957	FALSE	FALSE	2019-07-26	Thanks To CM of Uttarakhand @tsrawatbjp to the	0	2	0		
15	airnews_g	Guwahati, India		24883	134	FALSE	TRUE	2019-07-26	à¦‡à¦£à§à¦¡à¦¿à¦¯à¦¼à¦¾à¦¨	0	2	0		
16	kattar_hindu05			2021	2287	FALSE	FALSE	2019-07-25	Chamki bukhar me hospital ke ICU se live reportir	0	0	0		
17	arc_iam	Guwahati, India		1	91	FALSE	FALSE	2019-07-24	Big thanks to @akshaykumar for support us and o	0	0	0		
18	ShashaSri	Mumbai, India		2323	26	FALSE	FALSE	2019-07-24	à¤¶à¤¾à¤® à¤¢à¤² à¤—à¤¯à¤¾ à¤µà¥‹	0	0	0		
19	Imravi44790404			23	409	FALSE	FALSE	2019-07-23	#assamflood #helpassamðŸ˜,ðŸ˜,ðŸ˜,	0	2	0		
20	arunaruz			4558	884	FALSE	FALSE	2019-07-23	à´Žà´,.à´ªà´¿ à´•àµà´•àµâ€Œ à´•à´¾à´±àµâ€Œ	0	6	0		
21	JunaidAhe	à¤‰à¤¤à¥à¤¤à¤° à¤ª		1	77	FALSE	FALSE	2019-07-23	Flood situation in our state Assam is very critical,	1	2	0		affected_individual
22	GDwayn_	à´µà´¯à´¨à´¾à´Ÿàµ		232	492	FALSE	FALSE	2019-07-23	à´¸à´¹à´¾à´¯à´¿à´•àµà´•àµ,	0	0	0		
23	suhasman	Thrissur		6004	218	FALSE	FALSE	2019-07-23	Make Assam Awesome #AssamIsInPain #Assam #.	0	2	2		
24	sandeeps	newdelhi		439	1683	FALSE	FALSE	2019-07-23	Due to the flood,Assam situation is	0	0	0		
25	_Mind_ha	Qatar ,à´•àµ‹à´´à´¿à´•		2752	1051	FALSE	FALSE	2019-07-23	à´†à´¸à´¾à´, à´Žà´¨àµà´¤à´¾	0	8	1		
26	jaimejohn300			5166	2361	FALSE	FALSE	2019-07-22	The weeps of Assam not in Indiaâ€™s trends	0	0	0		
27	vincenzo_	Land of Hc	Coordinat	5024	273	FALSE	FALSE	2019-07-22	à´¸àµ‡à´¨à´¾ à´¸à´¹à´¾à´¯à´,	0	12	1		
28	Great_Cha	Russia		5755	660	FALSE	FALSE	2019-07-22	100 à´°àµ‚à´ª à´•àµŠà´Ÿàµà´¤àµà´¤à´¾à´²àµà´, à´…à	0	11	4		
29	vincenzo_	Land of Hc	Coordinat	5024	273	FALSE	FALSE	2019-07-22	à´ªà´•àµà´•à´¿à´Ÿàµ† à´†à´¨àµà´±à´¿ à´†à´¸à´¾à´®à´	0	17	3		
30	vincenzo_	Land of Hc	Coordinat	5024	273	FALSE	FALSE	2019-07-22	à´±àµ€ à´¬à´¿àµ½à´¡àµâ€Œ à´•àµ‡à´°à´³à´¾	0	11	2		
31	kanimang	à´²àµ‹à´•à´®àµ‡ à´¤à´		5018	591	FALSE	FALSE	2019-07-22	à´šàµ†à´±àµà´¤àµ†à´¨àµà´¨àµ à´•à´°àµà´¤à´¿ à´†à´°	0	70	21		
32	DeepuGe	Kozhikode, India		3978	3139	FALSE	FALSE	2019-07-22	à´ªàµà´¤à´¿à´¯ à´•à´³à´•àµà´·àµ»	2	47	40		
33	LekshmiR	Kochi, Kerala		1053	194	FALSE	FALSE	2019-07-22	à´Žà´¨àµà´±àµ† à´•àµà´Ÿàµà´,à´¬à´,	0	4	0		
34	pappettar	à´‡àµ— à´…à´£àµà´Ÿà		3755	495	FALSE	FALSE	2019-07-22	Assam may not be able to raise fund like "WE"	0	12	9		
35	jokerboy9	Palakkad		603	505	FALSE	FALSE	2019-07-22	#assamFlood	0	0	0		
36	dhatrikutt	à´¬àµ†à´¸àµ¼à´ªàµà´		2847	178	FALSE	FALSE	2019-07-22	Please Pray âŒ	0	4	2		
37	Anurag22	Calicut		9	46	FALSE	FALSE	2019-07-22	Pray For #Assaam â˜º	1	1	1		

Table 1: Assam Flood 2017

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	username	Location	co-ordinat	Follower (	Friends cc	Protected	Verified	date	tweet	reply cour	like count	retweet count		
2	news24tv	Noida		627277	187	FALSE	TRUE	2019-10-2	à¤ªà¤Ÿà¤	1	21	0		Not useful
3	thehdnew	Patna		211	51	FALSE	FALSE	2019-10-2	à¤°à¤¾à¤	0	0	0		Not useful
4	writersaty	India		134	125	FALSE	FALSE	2019-10-1	à¤ªà¤Ÿà¤	0	4	0		Not useful
5	SarkarSip	New Delhi, India		56	198	FALSE	FALSE	2019-10-1	à¤…à¤—à	0	2	1		Not useful
6	MODIfied	Delhi, Purnea,india		1458	22	FALSE	FALSE	2019-10-1	à¤šà¤®à¥	0	1	0		Not useful
7	Now_Abhi82			354	623	FALSE	FALSE	2019-10-1	"Ink	0	0	0		not _humanitarian
8	krchandar	Delhi - Ghaziabad - E		2607	444	FALSE	FALSE	2019-10-1	Ink	0	0	0		not _humanitarian
9	ManishAr	Ranchi, India		302	323	FALSE	FALSE	2019-10-1	à¤¬à¤¿à¤	0	0	0		Not useful
10	IamKumar	Patna, India		355	282	FALSE	FALSE	2019-10-1	There is	0	1	1		Not useful
11	sitamarhil	Sitamarhi, India		91	13	FALSE	FALSE	2019-10-1	à¤…à¤ªà¤	0	1	0		Not useful
12	PawanRss	à¤—à¥‹à¤ªà¤¾à¤²à¤		1700	2767	FALSE	FALSE	2019-10-0	à¤¬à¤¾à¤	0	3	0		Not useful
13	airnews_p	Patna, India		21243	239	FALSE	TRUE	2019-10-0	#BiharRai	0	0	0		Not useful
14	airnews_p	Patna, India		21243	239	FALSE	TRUE	2019-10-0	#BiharRai	0	0	1		Not useful
15	airnews_p	Patna, India		21243	239	FALSE	TRUE	2019-10-0	#BiharRai	0	0	0		Not useful
16	LUCKYAG27051993			6	55	FALSE	FALSE	2019-10-0	#BIHARflo	0	0	0		sympathy_and_support
17	shakti_sin	Lucknow		92	354	FALSE	FALSE	2019-10-0	à¤¬à¤¿à¤¹	0	1	0		Not useful
18	Kavendra	à¤•à¤¾à¤¨à¤ªà¥à¤° à		83	450	FALSE	FALSE	2019-10-0	à¤¬à¤¿à¤	0	1	0		Not useful
19	ManishKu	saharsa(Bihar) | Nev		1184	792	FALSE	FALSE	2019-10-0	à¤•à¤¹à¤	0	6	4		Not useful
20	anshul_ur	à¤«à¤¤à¥‡à¤¹à¤ªà¥à		839	588	FALSE	FALSE	2019-10-0	à¤¬à¤¾à¤	0	6	2		Not useful
21	TOIPatna			6867	27	FALSE	TRUE	2019-10-0	Bihar:	1	2	0		caution_and_advice
22	Manoj_D_	Ahmedabad		2583	4787	FALSE	FALSE	2019-10-0	#à¤¬à¤¿à	0	3	0		not useful
23	PrinceSingh1509			16	59	FALSE	FALSE	2019-10-0	à¤¬à¤¿à¤¹	0	2	0		not useful
24	Suvasit			1018	1825	FALSE	FALSE	2019-10-0	à¤¨à¤—à¤	0	3	0		not useful
25	khabarser	Bihar, India		5286	248	FALSE	FALSE	2019-10-0	à¤•à¤¿à¤	1	4	3		not useful
26	LogicalNe	Lucknow, India		1007	83	FALSE	FALSE	2019-10-0	à¤¬à¤¿à¤	0	0	0		not useful
27	manojsing	Sasaram, Rohtas, Bih		1018	2117	FALSE	FALSE	2019-10-0	à¤ªà¤Ÿà¤	0	2	1		not useful
28	sbrsins	Kurnool, India		23	1210	FALSE	FALSE	2019-10-0	Bihar is pa	0	1	0		not useful
29	Ankitdu63	Murliganj, India		1393	1623	FALSE	FALSE	2019-10-0	I am	0	3	0		not useful
30	archu_pea	Vishakhapatnam, Inc		18	168	FALSE	FALSE	2019-10-0	#BIHARflo	2	3	0		not useful
31	nawaneet	Patna, India		20	124	FALSE	FALSE	2019-10-0	â€˜à¤¸à¥	0	0	0		not useful
32	Anandkun	New Delhi/Gorakhpu		5989	2882	FALSE	FALSE	2019-10-0	à¤¬à¤¿à¤	0	3	1		not useful
33	raviran22	Ranchi, India		433	940	FALSE	FALSE	2019-10-0	No big	0	0	0		not useful
34	Akdubey0	Noida, India		159	1081	FALSE	FALSE	2019-10-0	#à¤¬à¤¿à	0	0	0		not useful
35	Vishnut75276167			71	133	FALSE	FALSE	2019-10-0	#bihar flo	0	0	0		caution_and_advice
36	sujit2427	India		3412	2662	FALSE	FALSE	2019-10-0	à¤®à¤¾à	0	9	10		Not useful
37	SkymetW	India		94004	488	FALSE	TRUE	2019-10-0	(1/n) #We	1	5	1		caution_and_advice

Table 2: Bihar Flood 2019

username	Location	co-ordinat	Follower Count	Friends count	Protected	Verified	date	tweet	reply cour	like count	retweet count	
ceogsdma			430	71	FALSE	FALSE	2017-08-1	#Gujaratflood (24/07/17 to 14/08/17)	0	4	1	
BhumikaS	Gujarat, India		20535	1308	FALSE	FALSE	2017-08-0	#GujaratFlood proved who is true leader!	0	3	1	
Harsh4Ue	Bengaluru, India		3798	3653	FALSE	FALSE	2017-08-0	MLA à¤¸à¤«à¤° kr rhe hai or unki jnta Suffer	0	0	0	
divyagupt	Vadodara		702	2649	FALSE	FALSE	2017-08-0	There is huge need of Mosquito Nets to pro	0	3	1	caution_and_advice
parveensa	New Delhi & Bhiwan		1103	437	FALSE	FALSE	2017-08-0	à¤®à¤¾à¤¨à¤¨à¥€à¤¯	0	7	9	
ceogsdma			430	71	FALSE	FALSE	2017-08-0	#gujaratflood VMC will send one Supersuck	0	2	1	sympathy_and_support
rajinandagopal			219	336	FALSE	FALSE	2017-08-0	Our so-called secular media ll never show	0	0	0	
ceogsdma			430	71	FALSE	FALSE	2017-08-0	#gujaratflood accessing of loss of affected p	0	2	1	
ceogsdma			430	71	FALSE	FALSE	2017-08-0	#gujaratflood electricity restoration in #bar	0	1	1	
premjourn	Tamil Nadu		2303	2173	FALSE	FALSE	2017-08-0	à®•à¯à®œà®°à®¾à®¤à¯	0	0	0	
ceogsdma			430	71	FALSE	FALSE	2017-07-3	#gujaratflood relief material sent to morbi,	0	2	0	
Dilip_Desl	Dist.kutch,Village.bi		15063	882	FALSE	FALSE	2017-07-3	àª-àª¾àªœàªª àª¬à«‹àª²à«àª¯à«àª, àªàª•	1	31	10	
Sundayhe	Noida, India		795	66	FALSE	FALSE	2017-07-3	#GujaratFlood: à¤…à¤¬ à¤¤à¤• 60,000 à¤²à¥	0	0	0	
ceogsdma			430	71	FALSE	FALSE	2017-07-3	#gujaratflood #teams NDRF, SDRF, ARMY, A	0	3	1	sympathy_and_support
With_Sura	New Delhi, India		423	862	FALSE	FALSE	2017-07-3	Congress party is only organization which	0	1	0	
ceogsdma			430	71	FALSE	FALSE	2017-07-3	Extensive fogging, sanitation & cleani	0	2	0	
chintanjda	Gandhinagar, Gujara		188	133	FALSE	FALSE	2017-07-3	Solution of #IndianPoliticians on #GujaratF	0	0	0	
chintanjda	Gandhinagar, Gujara		188	133	FALSE	FALSE	2017-07-3	Wow! I am sure watching videos of @vijayr	0	0	0	
simuntwe	India		1524	165	FALSE	FALSE	2017-07-3	And @aamir_khan asking people to	0	0	0	
padmanab	Bhubaneswar		46	83	FALSE	FALSE	2017-07-3	#GujaratCongress enjoying in resorts when	0	0	0	
ceogsdma			430	71	FALSE	FALSE	2017-07-3	Chlorine tablets & ORS packets availab	0	2	1	
singhratar	mumbai		252	134	FALSE	FALSE	2017-07-2	à¤•à¤¾à¤‚à¤—à¥à¤°à¥‡à¤¸à¥€	0	0	0	
ceogsdma			430	71	FALSE	FALSE	2017-07-2	#gujaratflood keep umbrellas &bamb	0	1	1	caution_and_advice
ceogsdma			430	71	FALSE	FALSE	2017-07-2	Restoration of power, roads & cleanin	0	1	1	
Aashkajan	rajkot		843	235	FALSE	FALSE	2017-07-2	Glimpses of seva for flood relief work by	0	1	0	
Hope_bel	#Ahmedabad #India		937	830	FALSE	FALSE	2017-07-2	â›ˆâ›ˆâ›ˆ	0	0	0	
rawal_jay	Ahmedabad | Udaip		10651	1478	FALSE	FALSE	2017-07-2	à¤—à¥à¤œà¤°à¤¾à¤¤ à¤®à¥‡à¤‚, à¤¬à¤¾à¤	0	0	0	
pandiyank	Chennai, India		733	1407	FALSE	FALSE	2017-07-2	Let's see how much fund the center is goin	0	0	0	
sanulovex6			335	1830	FALSE	FALSE	2017-07-2	#GujaratFlood @Naveen_Odisha bjp only v	1	0	0	
INeelrajsinh			228	206	FALSE	FALSE	2017-07-2	Dharoi dam 50000 cusec water channeled ir	0	0	0	caution_and_advice
mihir_kan	ahmedabad,india		71	94	FALSE	FALSE	2017-07-2	àª¬àª¹à« àªµàª°àª¸à«àª¯à«‹ àª¹à«‡	0	1	0	
zeesalaam	Delhi, india		6993	52	FALSE	TRUE	2017-07-2	à¤¸à¤¾à¤¬à¤°à¤®à¤¤à¥€ à¤¨à¤¦à¥€ à¤®à¥‡à	0	0	0	
Zee_Hind	Essel Studio, FC-19, S		58668	127	FALSE	TRUE	2017-07-2	à¤¸à¤¾à¤¬à¤°à¤®à¤¤à¥€ à¤¨à¤¦à¥€ à¤®à¥‡à	0	2	1	
ashishkel	India		253	381	FALSE	FALSE	2017-07-2	We live within the same boundaries but #A	0	0	0	
BhaveshM	Surat, Gujarat		2592	1290	FALSE	FALSE	2017-07-2	àª…àª®àª¦àª¾àªµàª¾àª¦àª®àª¾àª‚,	0	1	0	
shalavkela	India		219	468	FALSE	FALSE	2017-07-2	Pls help the flood victims. If the person car	0	0	0	

Table 3: Gujarat Flood 2017

username	Location	date	tweet	
jikoni_for	Pune, Indi	2021-09-1	#BackToSchool is a fundraiser for raising 1.5 lacs for SPM Urdu High School, Village Rajewadi. To support DM or WA us on 9359176814. Please amplify. #MaharashtraFloods #MissionRajewadi	donation_and_volunteering
rons1212	Pune	2021-09-0	Can we have 140 people supporting 140 families in village Rajewadi for their basic household items. You can contribute 1000 inr per family for no of families you want to support. #MissionRajewadi #MaharashtraFloo	donation_and_volunteering
NukteVivi	ðŸŒ MH -	2021-08-3	It's bn raining	caution_and_advice
rons1212	Pune	2021-08-3	Spare a thought for the 168 families we are fundraising to send them mattresses, bedsheets and some basic things for this house which was washed away by floods last month. You can contribute 1000 inr per family	caution_and_advice
JaiMahara	Mumbai, I	2021-08-2	#BIGNEWS Raju	Not useful
rons1212	Pune	2021-08-2	Our fundraiser for #MaharashtraFloods are still ongoing since we are short of 170,000 for 170 families for their basic items to rebuild their homes once again. The floods has been devastating and all the houses in vill	structure_and_utlities_damage
JaiMahara	Mumbai, I	2021-08-2	#BIGNEWS Raju	Not useful
RohitKo68	Thane, Inc	2021-08-1	à¤…à¤«à¤—à¤	Not useful
jikoni_for	Pune, Indi	2021-08-1	Thank you for choosing @jikoni_for_all as your charity of choice. Together weâ€™ve changed many lives affected by floods in Maharashtra. The families we have support with your help now lives with a new hope. #	Not useful
rons1212	Pune	2021-08-1	Konkan is still having heavy rains. Do support the families where flood have devastated their homes. #MaharashtraFloods	sympathy_and_support
rons1212	Pune	2021-08-1	We are short of 12 families for our target of 100 families #MissionRajewadi The need for mattresses and bedsheets are urgent. To contribute to #JikoniFloodFund to enable us to support families to get back to their	donation_and_volunteering
Shambhav	India	2021-08-0	When #Maharashtra has faced a lot of tragedy as floods. No bollywood artists offered any kind of help also when they work here and earn their livelihood... Shame on such so called 'Stars' #Bollywood #Maharashtra	Not useful
jikoni_for	Pune, Indi	2021-08-0	Jikoni	donation_and_volunteering
rons1212	Pune	2021-08-0	Jikoni	donation_and_volunteering
rons1212	Pune	2021-08-0	Can we have 40 people supporting 40 families of Rajewadi to get mattresses & gas stoves? Each family needs 2000 inr to get back to their normal lives. Let us do it. Please share widely. #MissionRajewadi #Maha	donation_and_volunteering
humanaid	à¤¨à¤¨ à¤	2021-08-0	#HAI is thankful to @CAFIndia_tweets for supporting our #flood response in Maharashtra. #MaharashtraFloods	Not useful
jikoni_for	Pune, Indi	2021-08-0	Call for request to corporates, socially responsible brands, entrepreneurs, start ups if helping a village of 700 families synergises with your CSR focus, please get in touch with our team on 9359176814 or email jikonif	donation_and_volunteering
VijayWad	Bramhapu	2021-08-0	à¤¨à¥‹à¤-à¤¿à¤¯à¤¾ à¤à¤¾à¤²à¥‡à¤²à¥à¤¯à¤¾ à¤¾à¤¸ à¥§à¥§,à¥«à¥¦à¥¦ à¤•à¥‹à¤Ÿà¥€ à¤°à¥à¤ªà¤¯à¤¾à¤‚à¤šà¥à¤¯à¤¾	
karishmas	MUMBAI,I	2021-08-0	#MaharashtraFloods #Mumbaikar To all celebrities n Mumbaikar Maharashtra ppl need ur help.	Not useful
SolankitTa	Maharash	2021-08-0	à¤¸à¤°à¥à¤µà¤	Not useful
Raj_abd79795754		2021-08-0	Dear	Not useful
Yogeshfas	Paithan, I	2021-08-0	à¤ªà¥‚à¤°à¤—à	Not useful
Dev_Fada	Maharash	2021-08-0	Dev_Fadnavis:	Not useful
Dev_Fada	Maharash	2021-08-0	Dev_Fadnavis:	Not useful
Dev_Fada	Maharash	2021-08-0	Dev_Fadnavis:	Not useful
sanjay_go	NOIDA	2021-08-0	#à¤®à¤¹à¤¾à¤°	Not useful
Shrimant	Nagpur	2021-08-0	à¤•à¥à¤£à¥€ à¤—à¤¾à¤¨ à¤®à¤¾à¤°à¤²à¥€ à¤®à¥à¤¨à¥€à¤,à¤¨ à¤†à¤ªà¤£ à¤µà¤¾à¤¸à¤°à¥, à¤®à¤¾à¤°à¤£à¥à¤¨ à¤¾à¤šà¥€ à¤—à¤°à¤œ à¤¨à¤¾à¤¹à¥€. à¤°à¤¾à¤œà¥à¤¯à¤ªà¤¾à¤² #BhagatsinghKoshyari	
Sanketsku	Sangli. Ma	2021-08-0	à¤ªà¥…à¤•à¥‡à	Not useful
PrabhakarAshis		2021-08-0	à¤®à¤¹à¤¾à¤°à	Not useful
rons1212	Pune	2021-08-0	Any masala brand / startup on my TL who would like to donate / support a basic masala kit for 700 families of Rajewadi who are starting their lives on a fresh note? Please amplify. #MaharashtraFloods #JikoniFloodF	donation_and_volunteering
deshkar_a	Nagpur, Ir	2021-08-0	#Maharashtra	donation_and_volunteering
JaiMahara	Mumbai, I	2021-08-0	#BIGNEWS	donation_and_volunteering
rons1212	Pune	2021-08-0	Reaching out to corporate donors and Entrepreneurs for supporting #JikoniFloodFund. Do support @jikoni_for_all to re-build lives of people affected in Rajewadi & Birwadi. Do drop us an email@at jikoniflood	donation_and_volunteering
jikoni_for	Pune, Indi	2021-08-0	Jikoni Foundationâ€™s Disaster Relief & Rehabilitation efforts in supporting villages of Mahad has been going on progressively. We are giving you an update from our recent field visit to villages of Rajewadi, Bi	donation_and_volunteering
TOIMumb	Mumbai	2021-08-0	#Maharashtra cabinet clears relief package of Rs 11,500 crore for flood and landslide-affected Raigad, Ratnagiri, Satara, Sangli and Kolhapur district. #MaharashtraRains #MaharashtraFloods	donation_and_volunteering
JaiMahara	Mumbai, I	2021-08-0	#BIGNEWS	not useful

Table 4: Maharashtra Flood 2021

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
22	RatikantTripa10			1	4	FALSE	FALSE	2020-11-1	Sir take the alignment of Gopalpur-Digha Coastal Highway in betw	0	0	0		caution_and_advice			
23	dpradhanl	Delhi/Odisha/India		1602389	1695	FALSE	TRUE	2020-11-1	Thank Hon. HM Shri @AmitShah for approving â‚¹4381.88 crore of a	10	939	120		sympathy_and_support			
24	TNITweet	Bhubaneswar. New I		1878	462	FALSE	FALSE	2020-11-1	#WestBengal , #Odisha , #Maharashtra , #Karnataka , #MadhyaPrad	0	2	0		sympathy_and_support			
25	bhartijain	New Delhi, Delhi		74916	1346	FALSE	TRUE	2020-11-1	Panel chaired by HM Amit Shah approves additional Central assista	1	169	21		sympathy_and_support			
26	Sambad_E	Odisha		65115	4	FALSE	FALSE	2020-11-1	High-Level Committee under Chairmanship of Union Home	0	11	0		sympathy_and_support			
27	raykasturi	Bhubanes	Coordinat	4231	745	FALSE	TRUE	2020-11-1	Additional funds for Odisha!	1	23	5		Not useful			
28	timesofin	New Delhi		13943584	470	FALSE	TRUE	2020-11-1	High-Level Committee under Chairmanship of Union Home Minist	2	62	5		sympathy_and_support			
29	IndiaAhea	India		24944	158	FALSE	TRUE	2020-11-1	High-Level Committee under Chairmanship of Union Home Minist	0	0	0		sympathy_and_support			
30	TOICitiesNews			25148	104	FALSE	TRUE	2020-11-1	#UnionHomeMinistry approves Rs 4,381.88 crore additional assista	0	26	6		sympathy_and_support			
31	GulistanN	Jammu And Kashmir		13840	4	FALSE	FALSE	2020-11-1	High-Level Committee under Chairmanship of Union Home Minist	0	4	0		sympathy_and_support			
32	sudhakarc	N. Delhi		2725	1735	FALSE	FALSE	2020-11-1	High Level Committee chaired by HM @AmitShah approves Rs.	0	7	1		sympathy_and_support			
33	Dynamite	New Delhi, India		16252	64	FALSE	TRUE	2020-11-1	High-Level Committee under Chairmanship of Union Home Minist	0	0	0		sympathy_and_support			
34	ANI	India		6109154	0	FALSE	TRUE	2020-11-1	High-Level Committee under Chairmanship of Union Home Minist	10	453	14		sympathy_and_support			
35	sanashakil_TNIE			4649	3618	FALSE	FALSE	2020-11-1	Union Home Minister @AmitShah approves Rs. 4,381.88 crore of a	0	1	0		sympathy_and_support			
36	journo_jit	New Delhi, India		6987	169	FALSE	TRUE	2020-11-1	High Level Committee under Chairmanship of Union HM	0	13	2		sympathy_and_support			
37	sudhira_jena			3	83	FALSE	FALSE	2020-10-3	The above pictures show the condition of RD road connecting	0	2	0		affected_individual			
38	otvnews	Bhubaneswar, India		859123	256	FALSE	TRUE	2020-10-2	Aware that you have not been able to meet me; will visit once #CC	1	39	1		sympathy_and_support			
39	SatwikDas	Orissa, India		68	588	FALSE	FALSE	2020-10-2	Our HCM Shri @Naveen_Odisha helps the Telangana flood	0	1	1		sympathy_and_support			
40	DeoSudha	Cuttack-Odisha-Indi		517	1031	FALSE	FALSE	2020-10-2	Request to govt of odisha to supply ground nut seed in flood affec	0	0	1		donation_and_volunteering			
41	otvnews	Bhubaneswar, India		859123	256	FALSE	TRUE	2020-10-2	#Odisha Special Relief Commissioner, Pradeep Jena speaks to con	0	91	2		caution_and_advice			
42	otvnews	Bhubaneswar, India		859123	256	FALSE	TRUE	2020-10-2	CM Naveen Patnaik announces a contribution of Rs 5 crore toward	6	542	14		sympathy_and_support			
43	alpharay6	LIFE...		1596	1637	FALSE	FALSE	2020-10-2	Hon CM Sri #naveenpatnaik announced a contribution of Rs 5 cror	0	5	0		sympathy_and_support			
44	iam_Moh	Hyderabad, India		108	1268	FALSE	FALSE	2020-10-2	Odisha chief minister Naveen Patnaik announces Rs 5 crore	0	0	0		sympathy_and_support			
45	ArshIsmai	Berhampur / Bhubne		478	57	FALSE	FALSE	2020-10-2	BREAK: #Odiaha CM @Naveen_Odisha announces contribution of I	0	1	1		sympathy_and_support			
46	Nandigho	Bhubaneshwar, Indi		55852	260	FALSE	FALSE	2020-10-2	CM @Naveen_Odisha announced a contribution of Rs 5 crore	0	6	2		sympathy_and_support			
47	Dibyajyoti	Bhubaneshwar, Indi		879	1627	FALSE	FALSE	2020-10-2	CM Naveen patnaik announced a contribution of Rs 5 crore toward	0	0	0		sympathy_and_support			
48	SujitBisoy	Bhubanes	Coordinat	8206	666	FALSE	TRUE	2020-10-2	CM @Naveen_Odisha announces Rs 5 crore towards relief efforts		1	17	0		sympathy_and_support			
49	Vkpandiar	Bhubaneswar		4855	2	FALSE	FALSE	2020-10-2	Hon CM @Naveen_Odisha sir announced contribution of Rs 5 crore	1	18	1		sympathy_and_support			
50	SecyChief	Bhubaneshwar, Indi		178286	57	FALSE	TRUE	2020-10-2	Hon CM Sri Naveen patnaik announced a contribution of Rs 5 crore	26	505	22		sympathy_and_support			
51	babrumar	Bengaluru, India		0	19	FALSE	FALSE	2020-10-2	thanks for giving grocery items to flood affected people of	0	0	0		donation_and_volunteering			
52	mano_soh	India		267	409	FALSE	FALSE	2020-10-1	1 supersonic helicopter invented, which takes 17 minutes to comp	0	1	0		sympathy_and_support			
53	RaoBagusetty			0	5	FALSE	FALSE	2020-10-1	We from Bhamini village, Ramanna guda block, Rayagada dist.,Odi	0	0	0		infrastructure_and_utility_damage			
54	otvnews	Bhubaneswar, India		859123	256	FALSE	TRUE	2020-10-1	In view of the low pressure & heavy rainfall forecast on Octol	0	58	3		caution_and_advice			
55	partha_papu			10	20	FALSE	FALSE	2020-10-1	I survived flood,tsunami,cyclone,flu and stupid people just to be k	0	1	0		affected_individual			
56	Swapnil4I	ðŸ‡®ðŸ‡³ #Akhand_E		372	1882	FALSE	FALSE	2020-10-1	Neither Cyclones,floods,	1	0	0		Not useful			
57	Akkifg3	India		651	566	FALSE	FALSE	2020-10-1	Odisha flood ~ 1 Crore.. #à¤°à¤¾à¤·à¥à¤Ÿà¥à¤°à¤µà¤¾à¤¦à¥€_à¤¹à¥	0	0	1		Not useful			
58	Sassy_104	Hyderabad, India		379	631	FALSE	FALSE	2020-10-1	Facebook will never be able to introduce a "marked safe in XYZ	0	10	0		Not useful			

Table 5: Odisha Flood 2019

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	username	Location	co-ordinat	Follower	Friends cc	Protected	Verified	date	tweet	reply cour	like count	retweet count						
2	m00n_s33d			144	859	FALSE	FALSE	2014-09-2	Why not/Will you allow Int. Orgs to help flood affected Ka	1	0	0		not_humanitarian				
3	abrarzarga	Kashmir		551	89	FALSE	FALSE	2014-09-2	Delighted to say @927BIGFM #Srinagar is back,it's time to	0	0	0		Not useful				
4	szaffariqb	Jammu and Kashmir		25444	928	FALSE	FALSE	2014-09-2	So 'darbar' will move to Jammu as usual on 'right' time- lea	0	0	1		Not useful				
5	Ak07Khan	Aurangabad, India		2316	2646	FALSE	FALSE	2014-09-2	Rahul Gandhi is on his way to the relief camps run by IYC a	0	0	0		symapthy_and_support				
6	Shayaq_A	Kashmir		394	163	FALSE	FALSE	2014-09-2	Srinagar ravaged by floods.. Lal chowk looks like war torn k	0	0	1		infrastrucrure_and_utilities_damage				
7	ekguru1			74	6	FALSE	FALSE	2014-09-2	#GurmeetRamRahimSinghJiInsan distributes relief in Srina	0	1	1		Not useful				
8	bukharish	Srinagar		53264	1631	FALSE	TRUE	2014-09-2	Shifting RPO Srinagar to Jammu is to add to sufferings of fl	0	0	4		donation_and_volunteering				
9	szaffariqb	Jammu and Kashmir		25444	928	FALSE	FALSE	2014-09-2	At UN PM Modi talks about Kashmir floods n helping Pak d	6	0	3		Not useful				
10	PasbaneH	Pakistan		2286	763	FALSE	FALSE	2014-09-2	Diverting water from river Neelum into local lake created	0	0	0		caution_and_advice				
11	basiitzarg	SRINAGAR		32842	143	FALSE	TRUE	2014-09-2	TIMES NOW J&K lensman	0	0	0		Not useful				
12	leonbakh	http://twitter.com/l		4852	3678	FALSE	FALSE	2014-09-2	No chopper sorties were carried out to evacuate the	1	1	1		not_humanitarian				
13	ians_india	New Delhi		67700	45	FALSE	TRUE	2014-09-2	17,360 tonnes of flood garbage removed from Srinagar	0	1	1		Not useful				
14	how2mak	srinagar kashmir ind		369	497	FALSE	FALSE	2014-09-2	Cleaning of products............ cleaning	0	0	0		Not useful				
15	AbuMuha	ðŸ‡¸ðŸ‡¦ Ø§Ù„Ù…Ù.		120	257	FALSE	FALSE	2014-09-2	12 lakh families hit in JK Flood	0	0	0		affected_individual				
16	A123LAD	Bat Cave		1605	1540	FALSE	FALSE	2014-09-2	Had ordered seeds for Father from Amazon a month back,	0	0	0		Not useful				
17	Fairoz_JK	Ramban Jammu & Ka		51301	726	FALSE	TRUE	2014-09-2	#NSUI Medical Relief Camp in Srinagar and flood effected	0	1	0		donation_and_volunteering				
18	shafkat81	Jammu And Kashmir		239	138	FALSE	FALSE	2014-09-2	In #kashmir floods at least did one good work by clearing t	0	0	0		Not useful				
19	FarahBashir			7301	1391	FALSE	TRUE	2014-09-2	Watching #Harud & weeping over #Srinagar city that's	1	2	1		Not useful				
20	omni_pur	Kolkata, India		1671	1032	FALSE	FALSE	2014-09-2	#Srinagar's business hubs lie desolate after catastrophic fl	0	0	0		Not useful				
21	Bilalahme	SRINAGAR,J & K.		140	734	FALSE	FALSE	2014-09-2	Srinagar Floods:SMHS Hospital without sufficient drinking	0	0	0		Not useful				
22	Bilalahme	SRINAGAR,J & K.		140	734	FALSE	FALSE	2014-09-2	Srinagar Floods: Day 17th Residency Road commercial cent	0	0	0		infrastructure_and_utility_damage				
23	Bilalahme	SRINAGAR,J & K.		140	734	FALSE	FALSE	2014-09-2	Srinagar Floods: Sad Govt not doing enough,day 17th and s	0	0	0		not_humanitarian				
24	Bilalahme	SRINAGAR,J & K.		140	734	FALSE	FALSE	2014-09-2	Srinagar Floods: friend got two shirts for me.	1	0	0		Not useful				
25	Bilalahme	SRINAGAR,J & K.		140	734	FALSE	FALSE	2014-09-2	Srinagar Floods:Raj Bagh still under 4/5 ft water	0	0	0		infrastructure_and_utility_damage				
26	TalismanC	New Delhi, India		405	80	FALSE	FALSE	2014-09-2	Mechanics in #Srinagar are charging hefty sum of money fc	0	1	1		not_humanitarian				
27	javedabba	kashmir		92	166	FALSE	FALSE	2014-09-2	available at Srinagar for relief to flood effected families	0	0	0		infrastructure_and_utility_damage				
28	Karma_Pa	New Delhi		17112	1978	FALSE	TRUE	2014-09-2	Militants attack a CRPF bunker in Srinagar and kill a Jawan	2	2	12		Not useful				
29	satisharma09			4364	3238	FALSE	FALSE	2014-09-2	#KashmirFloods;As per media reports local politicians play	0	0	0		Not useful				
30	aslam_sai	Mumbai,India		269	336	FALSE	FALSE	2014-09-2	Borderless World Foundation Kashmir Relief Fund	0	0	0		donation_and_volunteering				
31	NikitaNaresh			16	43	FALSE	FALSE	2014-09-2	You need not go to Srinagar to see floods, just come to bar	0	0	0		Not useful				
32	bhatkhurs	Srinagar		0	41	FALSE	FALSE	2014-09-2	Sir plz plz help me becose I am flood victom from Srinagar	0	0	0		affected_individual				
33	aam868	UK / India		369	209	FALSE	FALSE	2014-09-2	1) #KashmirFloods Google has uploaded high res. satellite	1	0	0		Not useful				
34	CA_global	Worldwide		10072	3126	FALSE	FALSE	2014-09-2	Floods have destroyed homes & left Srinagar #Kashm	0	0	2		Not useful				
35	927srinag	Srinagar		3700	19	FALSE	FALSE	2014-09-2	#KashmirFloods Now, JK seeks outside help to pump out	0	0	0		Not useful				
36	iamjunny	srinagar jammu & ka		123	117	FALSE	FALSE	2014-09-2	I got emotional after seeing the condition of flood in	0	0	0		Not useful				
37	bhartendu	à¤¹à¤¿à¤®à¤¾à¤²à¤¯		2783	1240	FALSE	FALSE	2014-09-2	Aoa RT @milligazette: Kashmir and Assam Floods. Media f	0	0	0		Not useful				

Table 6: Srinagar Flood 2014

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
Manjunaa	Planet Earth		832	289	FALSE	FALSE	2016-01-2	#Jallikattu	0	2	0								
Rktest4			2	34	FALSE	FALSE	2016-01-2	Chennai r	0	0	0								
gsurya	Malmo, Sweden		41723	2937	FALSE	FALSE	2016-01-2	Despite lo	0	2	4		not_humanitarian						
SwetaScribbles			164	285	FALSE	FALSE	2016-01-2	Almost tw	0	0	1								
amrigiri			83	209	FALSE	FALSE	2016-01-2	Relief for	0	0	0								
thameesu	Not in Mars		361	332	FALSE	FALSE	2016-01-2	à®Ÿà®¿à®	0	0	0								
ImSunilCh	India		1690	228	FALSE	FALSE	2016-01-2	#RamChar	0	1	2		donation_and_volunteering						
Harish_Ch	Born in INDIA,Made		4120	270	FALSE	FALSE	2016-01-2	#RamChar	0	3	3		donation_and_volunteering						
itzSHAFI	Andhra Pradesh, Ind		6059	488	FALSE	FALSE	2016-01-2	#RamChar	1	3	3		donation_and_volunteering						
SunithaSa	Chennai		25908	96	FALSE	FALSE	2016-01-2	If yr pets s	0	6	1								
HDFCBankNews			41219	93	FALSE	TRUE	2016-01-2	.@HDFC_E	0	0	0								
Oracle_In	India		14100	670	FALSE	TRUE	2016-01-2	Providing	0	0	0								
VivegamN	Coimbato	Coordinat	6528	1030	FALSE	FALSE	2016-01-2	à®µà¯†à®	0	2	2								
rniruphan			164	250	FALSE	FALSE	2016-01-2	After cher	0	1	0								
MosurAnna			1019	1023	FALSE	FALSE	2016-01-2	Not belitt	0	0	1								
Mettur_Rajesh			4170	461	FALSE	FALSE	2016-01-2	Almost 96	0	12	13								
Mettur_Rajesh			4170	461	FALSE	FALSE	2016-01-2	Governm	0	2	2								
JodiLogik	India	Coordinat	547	242	FALSE	FALSE	2016-01-2	#Republic	0	1	0								
tkaran50	India		3963	1193	FALSE	FALSE	2016-01-2	I want a n	0	0	0								
Madrassa	Bengaluru/Chennai/		31061	833	FALSE	TRUE	2016-01-2	TN Govt h	0	3	12								
TheVishnu	Madras-Sirkazhi,Indi		152	124	FALSE	FALSE	2016-01-2	after #Che	0	1	0								
bhat_shri	India		678	159	FALSE	FALSE	2016-01-2	donated 1	0	0	0		donation_and_volunteering						
krishthep	Coimbatore, Tamilna		951	875	FALSE	FALSE	2016-01-2	@DDneel	0	0	0								
Prabhadxl	Chennai		145	549	FALSE	FALSE	2016-01-2	So many c	0	1	0								
HDFCBankNews			41219	93	FALSE	TRUE	2016-01-2	.@HDFC_E	0	3	0								
ind_narra	Chennai,India		3	174	FALSE	FALSE	2016-01-2	First sign	0	0	0								
CSR_RT	Germany		4635	50	FALSE	FALSE	2016-01-2	RT @ HDF	0	0	0		sympathy_and_support						
vishnu_63	Tamil Nadu, India		358	1015	FALSE	FALSE	2016-01-2	Raining fo	0	0	0								
fairtellind	Chennai, Tamil Nadu		0	0	FALSE	FALSE	2016-01-2	We exten	0	0	0		sympathy_and_support						
Ganeshku	Tamil Nadu, India		135	594	FALSE	FALSE	2016-01-2	It is rainin	1	1	0		caution_and_advice						
sampath2	Chennai		161	399	FALSE	FALSE	2016-01-2	I didn't ge	0	0	0								
CSR_RT	Germany		4635	50	FALSE	FALSE	2016-01-2	RT @ push	0	0	0								
mohan711	India		327	442	FALSE	FALSE	2016-01-2	Cash disco	1	0	0		not_humanitarian						
dhwanigiri			113	281	FALSE	FALSE	2016-01-2	"Ola and U	0	1	0								
dhwanigiri			113	281	FALSE	FALSE	2016-01-2	People fro	0	1	1		sympathy_and_support						
nansense_	United States of Anx		5118	158	FALSE	FALSE	2016-01-2	"It is easy	0	1	1								
dhwanigiri			113	281	FALSE	FALSE	2016-01-2	"My field	0	1	1								

ChennaiFloods2015

Table 7: Chennai Flood 2015

In order to train and validate our model, sufficient tweets related to an event are needed, which should reflect the realistic scenario of that event. We used Twitter API to capture live tweets related to floods in southern and eastern states of India. The data collection was done using snscrape's twitter python library. A total of 32,400 tweets were collected with keywords "flood", "water". The collected tweets were in English, Hindi, and some other regional languages. For this study, we concentrated only on tweets in English and Hindi languages. One of the major problems with data collected from Twitter is that it may contain a lot of irrelevant tweets such as advertisements. There are many spammers, also known as 'spambots', sending huge number of tweets. Finding spammers is a very difficult task and a number of researchers (Benevenuto et al. 2010; Gayo-Avello 2013; Li and Du 2014; Yardi et al. 2010) are focusing on fixing this issue. In our case, spamming does not pose a significant problem, as we were collecting tweets originating from mobile phones only. The rationale behind this is that hand-held devices are used as personal devices, and they are hardly used for mass tweet dissemination. To filter out the tweets coming from hand-held devices, the source field of the tweets is used. To further reduce the effect of spambots, only tweets from users having a ratio of the number of followers to the number of those following less than one was stored

Event Name	Affected In	Caution &adv	donation & volunteering	infrastructure &ut	not Hummanita	sympathy & Support
Assam Flood	61	52	65	48	35	39
Bihar Flood	16	14	15	7	5	23
Gujarat Flood	62	37	41	34	12	77
Maharahtra Flood	32	10	34	1	4	19
Srinagar Flood	76	29	34	30	9	45
Odisha Flood	59	72	15	31	19	70
Chennai Flood	1219	1475	1703	1005	913	1385

Table 8: Dataset Labelling

	A	B	C	D	E	F
1	Event Name	Event type	Total Tweets	Date Range	Duplicate Tweets	Other Language
2	Assam Flood	Floods	3300	30-09-2019-2017-04-01	2900	100
3	Bihar Flood	Floods	725	2019-10-29 -2019-09-28	637	200
4	Gujarat Flood	Floods	2713	14-08-2017-2017-06-03	2300	150
5	Maharahtra Flood	Floods	832	2021-07-23-2021-07-22	354	300
6	Srinagar Flood	Floods	1172	29-09-2014-2014-09-03	740	200
7	Odisha Flood	Floods	1795	2020-11-29-2019-09-04	1250	279
8	Chennai Flood	Floods	30002	2016-03-24-2015-12-05	17500	4800
9						

Table 9: Events occurred

6.1.2 Models

Train	Test	Accuracy(%)	P(%)	R(%)	F1(%)
CrisisLex(6C)	CrisisLex(6C)	77	83	83	78
CrisisNlp(10C)	CrisisNlp(10C)	67	72	72	66
Consolidated(11C)	Consolidated(11C)	77	75	77	76

Table 10: Results of CNN model on Humanitarian datasets

Train	Test	Accuracy(%)	P(%)	R(%)	F1(%)
CrisisLex(2C)	1- CrisisLex 2- CrisisNlp 3- Consolidated	94 69 81	94 69 80	95 68 80	94 69 81
CrisisNlp(2C)	1- CrisisNlp 2- CrisisLex 3- Consolidated	82 71 72	82 80 76	82 71 73	82 70 72
Consolidated(2C)	1- Consolidated 2- CrisisLex 3- CrisisNlp	87 82 83	88 83 84	90 84 84	89 82 83

Table 11: Results of CNN model on Informativeness datasets

Train	Test	Accuracy	P(%)	R(%)	F1(%)
CrisisLex(6C)	CrisisLex(6C)	93.156	93.251	93.156	92.798
CrisisNLP(10C)	CrisisNLP(10C)	88.071	87.071	88.071	85.977
Consolidated(11C)	Consolidated(11C)	94.116	94.315	94.116	93.487

Table 12: Results of BERT model on humanitarian datasets

Train	Test	Accuracy	P(%)	R(%)	F1(%)
CrisisLex(2C)	1-CrisisLex 2-CrisisNLP 3-Consolidated	98.5375 71.329 82.177	98.559 73.905 82.160	98.537 71.329 82.177	98.534 69.189 81.926
CrisisNLP(2C)	1-CrisisLex 2-CrisisNLP 3-Consolidated	78.575 95.218 77.546	83.703 95.221 79.289	78.575 95.218 77.546	78.502 95.219 77.767
Consolidated(2C)	1-CrisisLex 2-CrisisNLP 3-Consolidated	85.325 86.624 96.389	85.125 86.612 96.398	85.325 86.624 96.389	84.930 86.513 96.380

Table 13: Results of BERT model on Informativeness datasets

6.1.2.1 Comparison between CNN and BERT

As BERT is bidirectional encoders representation from transformer, the accuracy which we got is higher than CNN model. We ran both model on same training and testing dataset and acquired greater accuracy on BERT. As shown in the above tables, the precision, recall and F1 scores of BERT model is higher than the CNN model. So, it can be said that BERT model is more effective than CNN model on tweets.

6.1.2.2 Named Entity Recognition Results

for named entity recognition(ner) we are using BiLstm (bidirectional LSTM) A Bidirectional LSTM, or biLSTM, is a sequence processing model that consists of two LSTMs: one taking the input in a forward direction, and the other in a backwards direction. We have trained our model on a tweeter dataset having tagged entities. The accuracy which we are getting is around 94 % for all the entities.

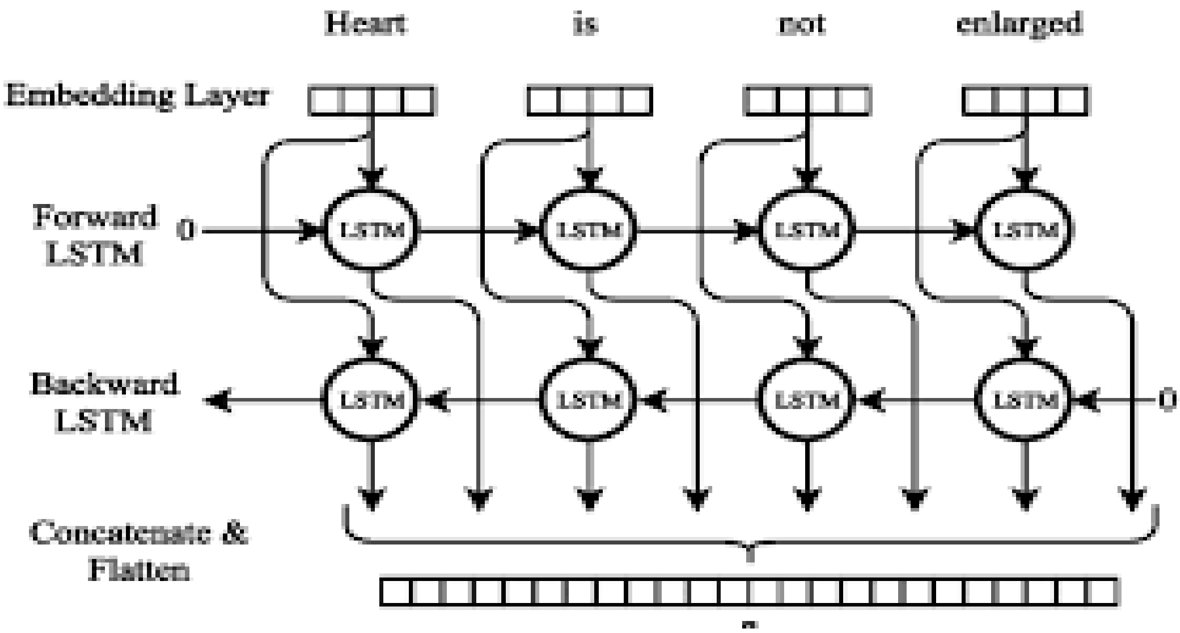

Figure 5 : Bi LSTM

```
-------------------- Train set quality: --------------------
processed 105778 tokens with 4489 phrases; found: 4544 phrases; correct: 4396.

precision:  96.74%; recall:  97.93%; F1:  97.33

            company: precision:   97.24%; recall:   98.60%; F1:   97.92; predicted:   652

           facility: precision:   91.89%; recall:   97.45%; F1:   94.59; predicted:   333

            geo-loc: precision:   97.81%; recall:   98.80%; F1:   98.30; predicted:  1006

              movie: precision:   94.12%; recall:   94.12%; F1:   94.12; predicted:    68

        musicartist: precision:   96.55%; recall:   96.55%; F1:   96.55; predicted:   232

              other: precision:   96.33%; recall:   97.09%; F1:   96.71; predicted:   763

             person: precision:   98.54%; recall:   98.87%; F1:   98.70; predicted:   889

            product: precision:   97.52%; recall:   98.74%; F1:   98.12; predicted:   322

         sportsteam: precision:   96.35%; recall:   97.24%; F1:   96.79; predicted:   219

             tvshow: precision:   80.00%; recall:   82.76%; F1:   81.36; predicted:    60

                    test set quality:
```

Table 14: Result of Bi LSTM model on train Dataset

```
-------------------- test set quality: --------------------
processed 118614 tokens with 5026 phrases; found: 4979 phrases; correct: 4591.

precision:  92.21%; recall:  91.35%; F1:  91.77

            company: precision:   91.38%; recall:   92.24%; F1:   91.81; predicted:   754

           facility: precision:   88.61%; recall:   91.67%; F1:   90.11; predicted:   360

            geo-loc: precision:   95.07%; recall:   93.87%; F1:   94.46; predicted:  1095

              movie: precision:   83.12%; recall:   85.33%; F1:   84.21; predicted:    77

        musicartist: precision:   93.85%; recall:   88.08%; F1:   90.87; predicted:   244

              other: precision:   90.58%; recall:   90.69%; F1:   90.64; predicted:   839

             person: precision:   95.57%; recall:   90.88%; F1:   93.17; predicted:   949

            product: precision:   90.86%; recall:   90.34%; F1:   90.60; predicted:   350

         sportsteam: precision:   91.14%; recall:   91.14%; F1:   91.14; predicted:   237

             tvshow: precision:   64.86%; recall:   77.42%; F1:   70.59; predicted:    74
```

Table 15: Result of Bi LSTM model on test dataset

6.1.3 Graph database

Knowledge graph are a way of structuring information in graph form by representing entities as nodes and relationship between entities as edges.

Following is an example of how the database increases in size as number of tweets increases-

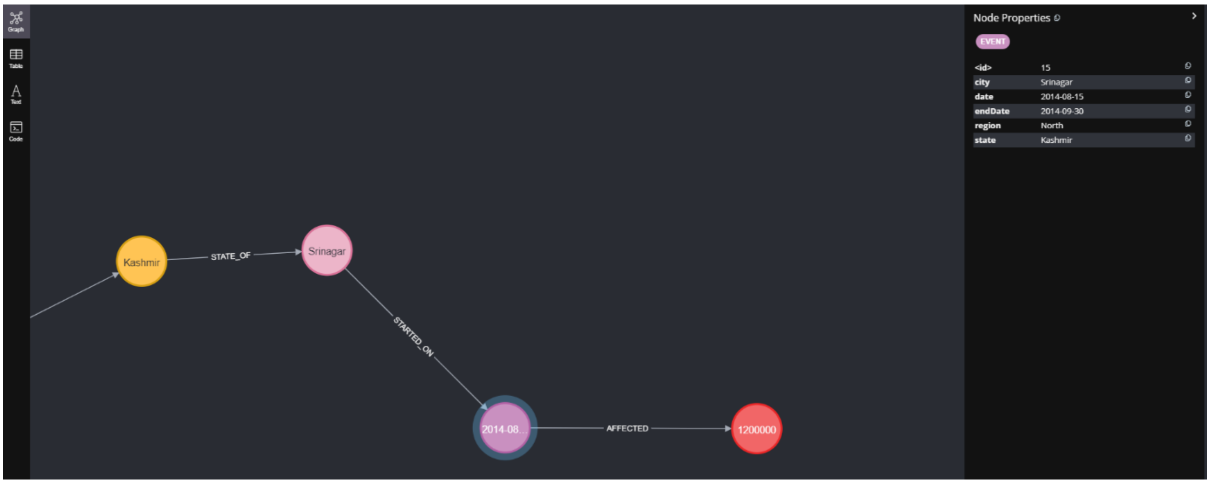

Figure 6: Graph after 15 Tweets

	A	B	C	D	E	F
1	Disaster	State	City	Started on	Affected	
2	Flood	Kashmir	Srinagar	15-08-2014	1200000	
3						
4						
5						
6						
7						

Table 16: CSV file after 15 tweets

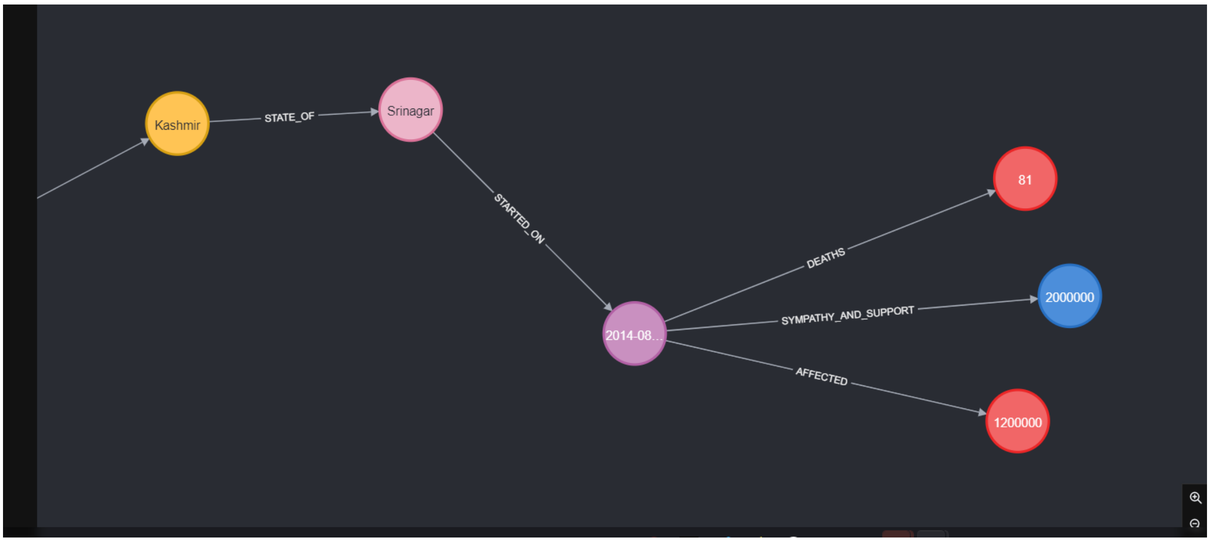

Figure 7: Graph after 50 Tweets

Disaster	State	City	Started on	Affected	Deaths	Donation
Flood	Kashmir	Srinagar	15-08-2014	1200000	81	AAP MPs MLAs will contribute Rs 20 lakh each for the Kashmir flood relief from their development fund

Table 17: CSV file after 50 Tweets

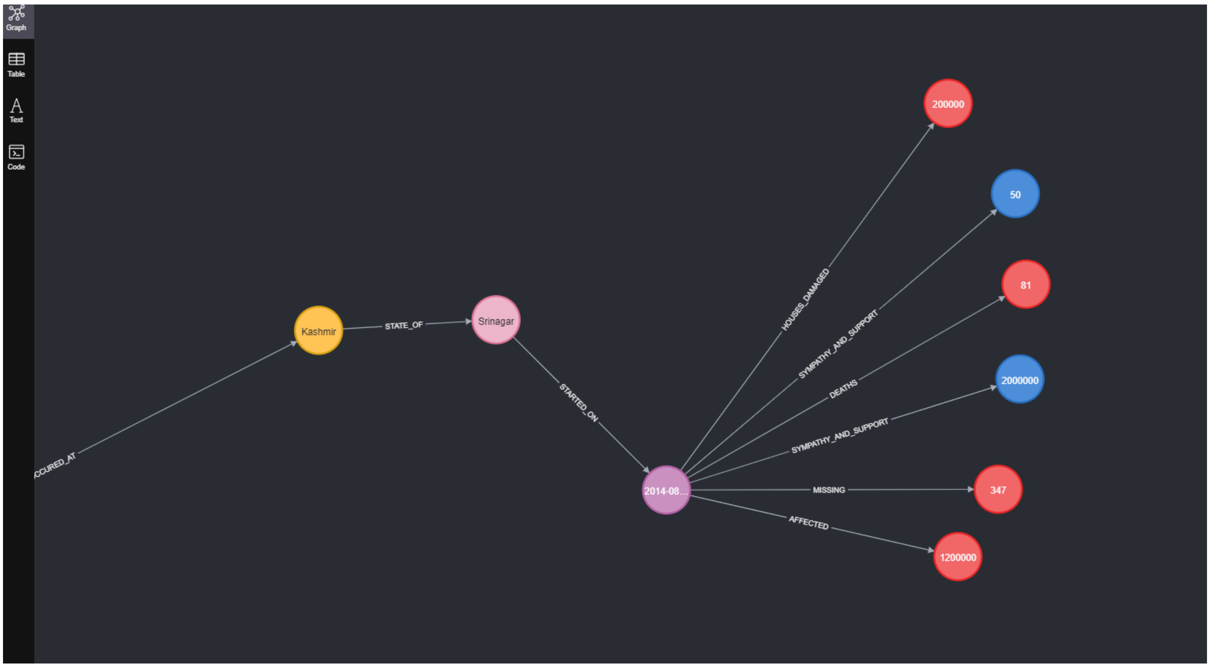

Figure 8: Graph after 100 Tweets

Disaster	State	City	Started on	Affected	Missing	Deaths	Houses Da	Donation	Donation
Flood	Kashmir	Srinagar	15-08-2014	1200000	347	81	200000	AAP MPs MLAs will contribute Rs 20 lakh each for t	Water survival boxes ready to be flown out to Srinagar Kashmir for flood rel

Table 18: CSV file after 100 tweets

The final graph that is formed can be queried in many ways. Some of the queries are given below-

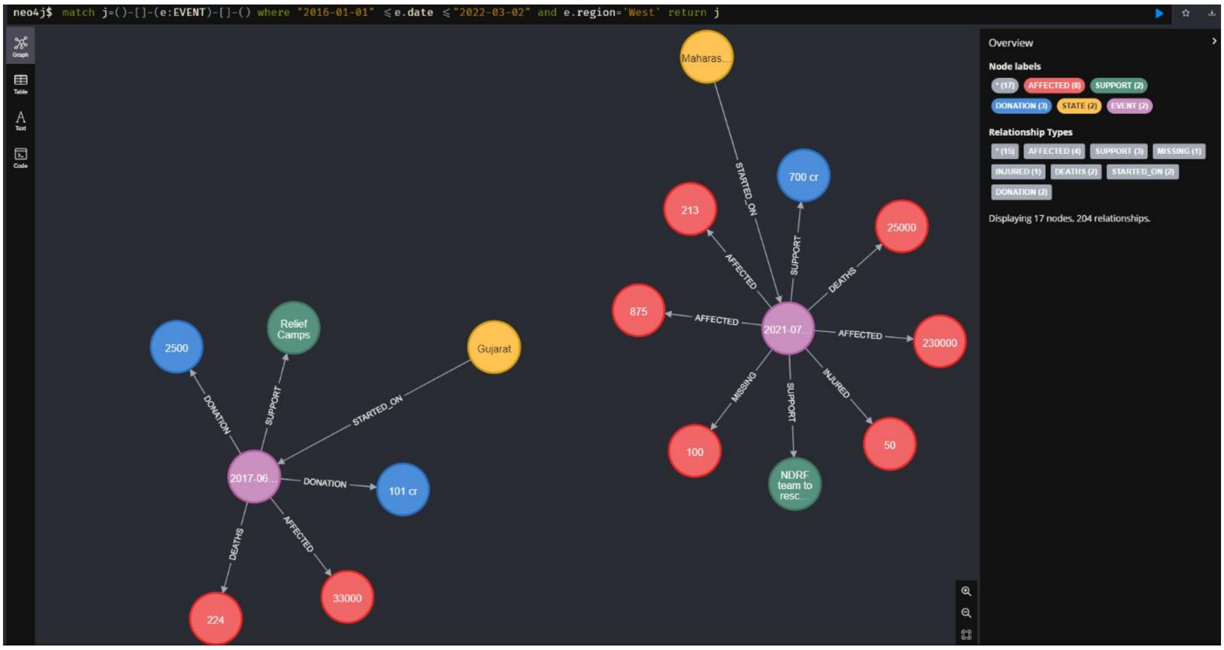

Figure 9: Graph which shows the floods that occurred in western region of India between 2016 and 2022.

Disaster	State	Started on	Affected	Missing	Deaths	Houses Da	Donation	Support	Villages aff	Injured	Animal dea	Donation
Flood	Maharasht	2021-07-15	230000	100	213	200000	Centre ap	34 NDRF te	81	50	25000	
Flood	Gujarat	01-06-2017	33000		224		101 cr don	12 Relief camps				2500 food packets donated by various organization

Table 19: CSV of Graph which shows the floods that occurred in western region of India/Local Storyline between 2016 and 2022.

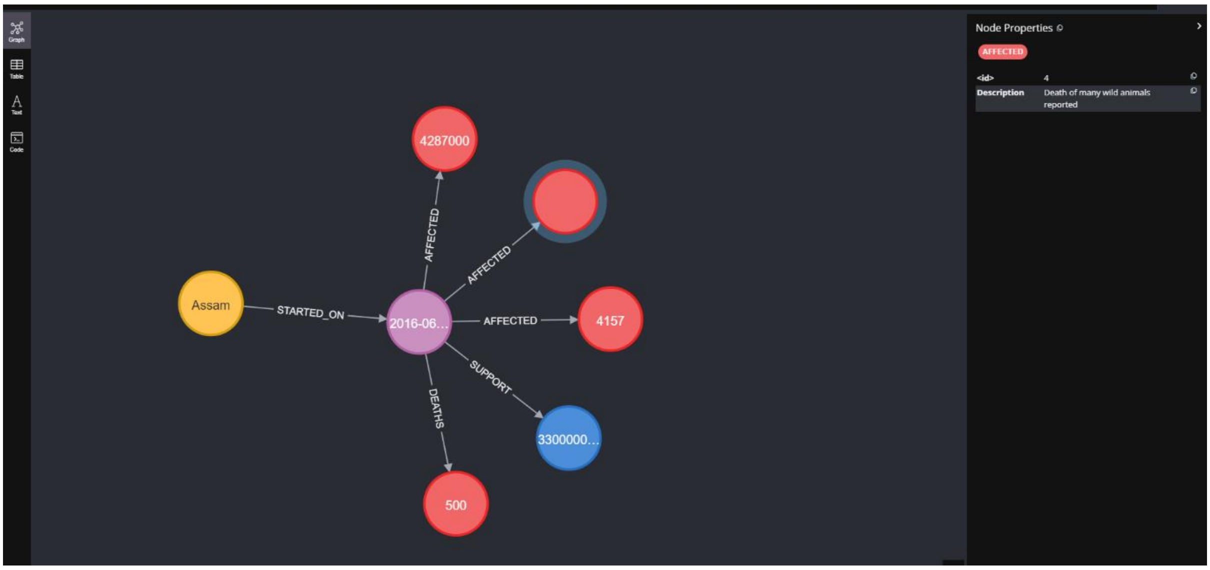

Figure 10: Graph that shows the flood that happened on a particular date (2017/03/16).

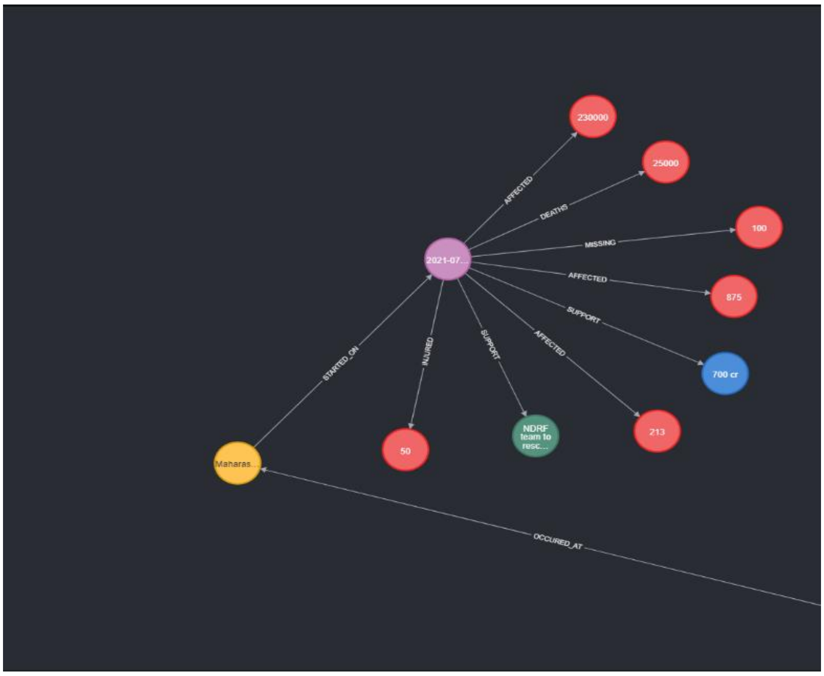

Figure 11: Local Storyline/Maharashtra

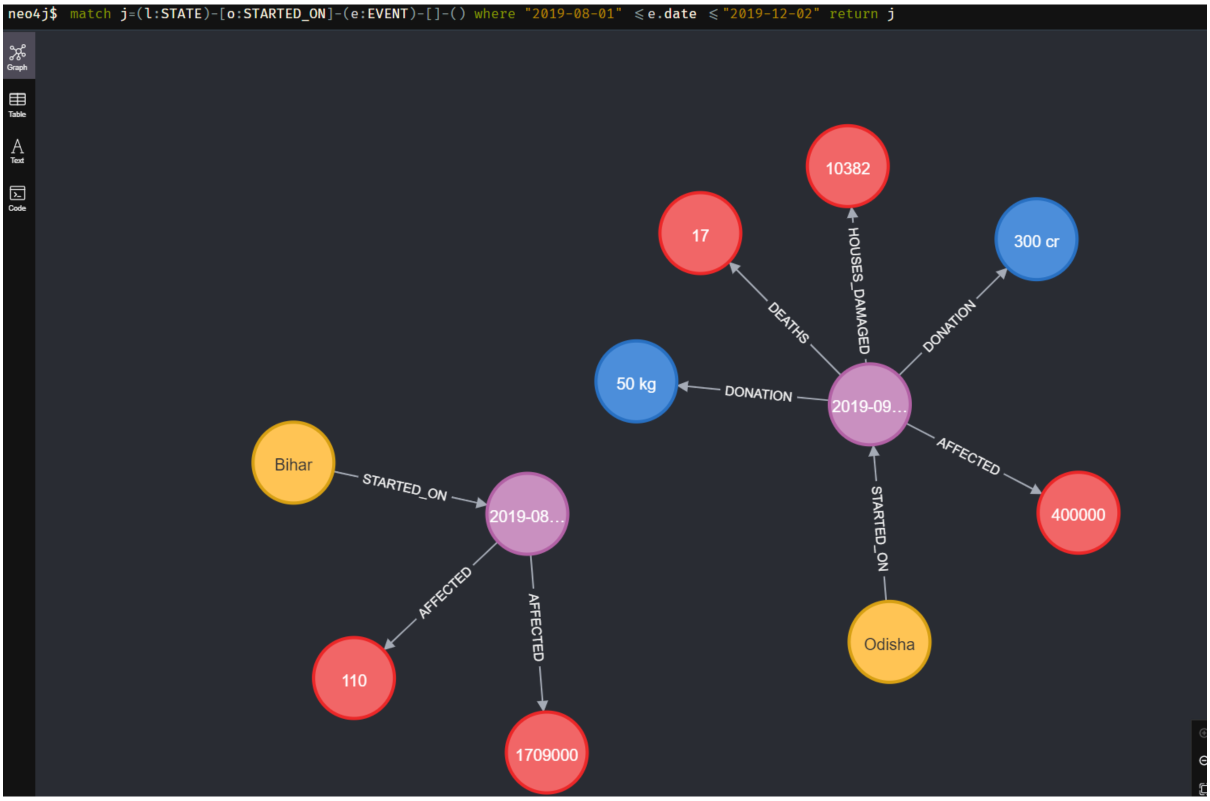

Figure 12: Graph which shows the floods that occurred in between August and December of 2019

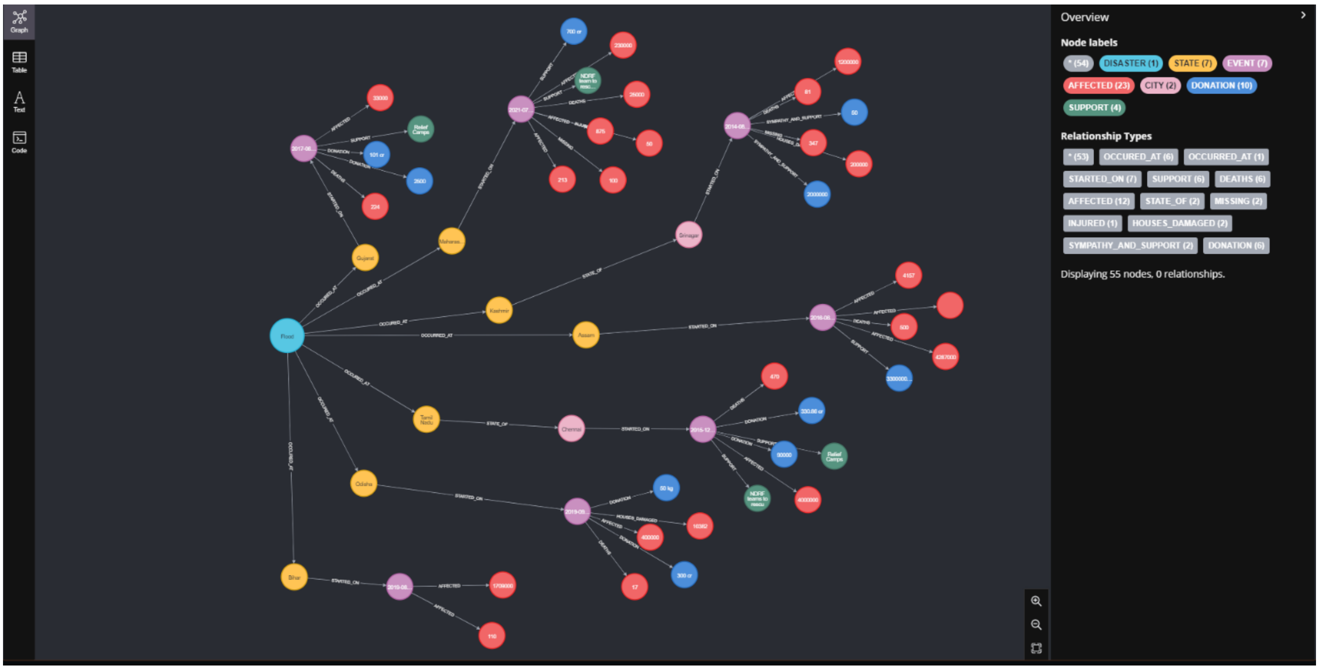

Figure 13: Global Storyline/India

6.1.3.1 Wikidata Comparison

We compared our knowledge graph with wikidata and observed that our knowledge graph provides much more data and is much easier to read than wikidata

It can be observed in the following two examples-

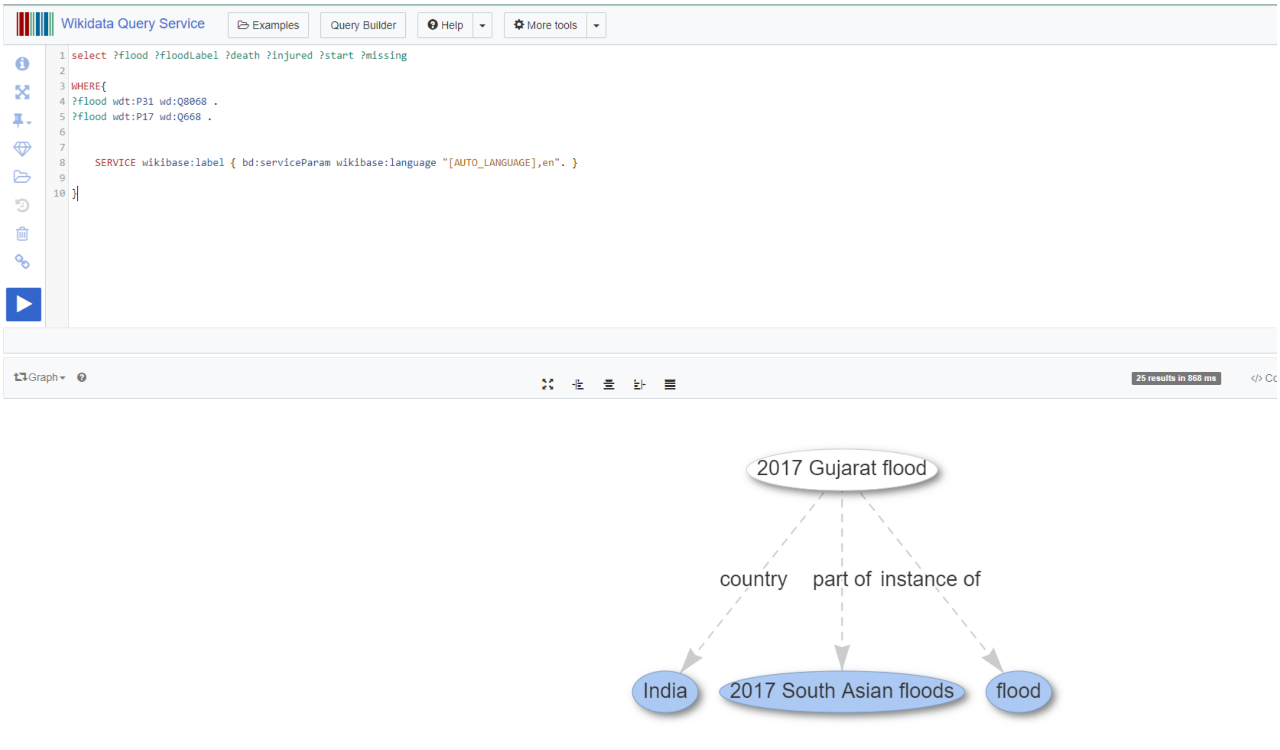

Figure 14: Wikidata of 2017 Gujarat Flood

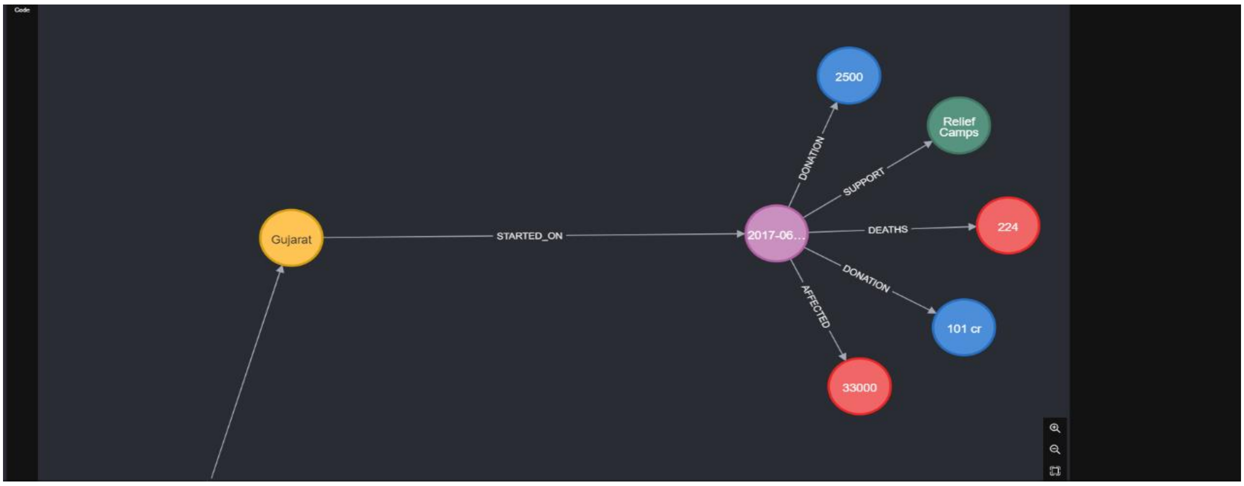

Figure 15: Knowledge Graph of 2017 Gujarat Flood

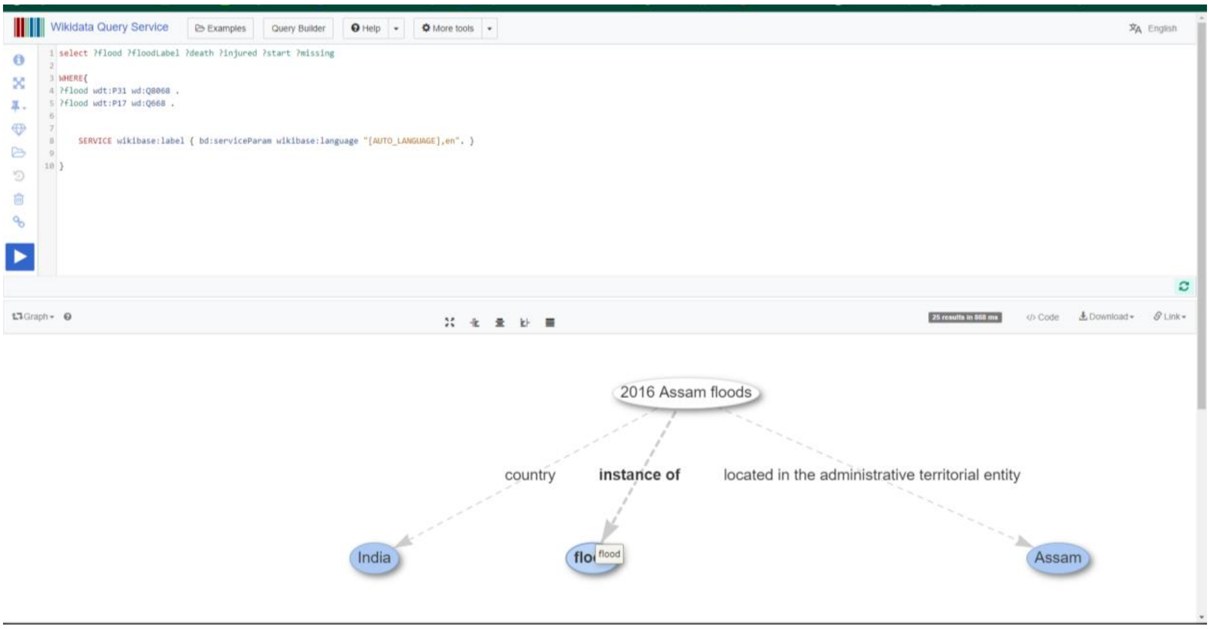

Figure 16: Wikidata of 2016 Assam Flood

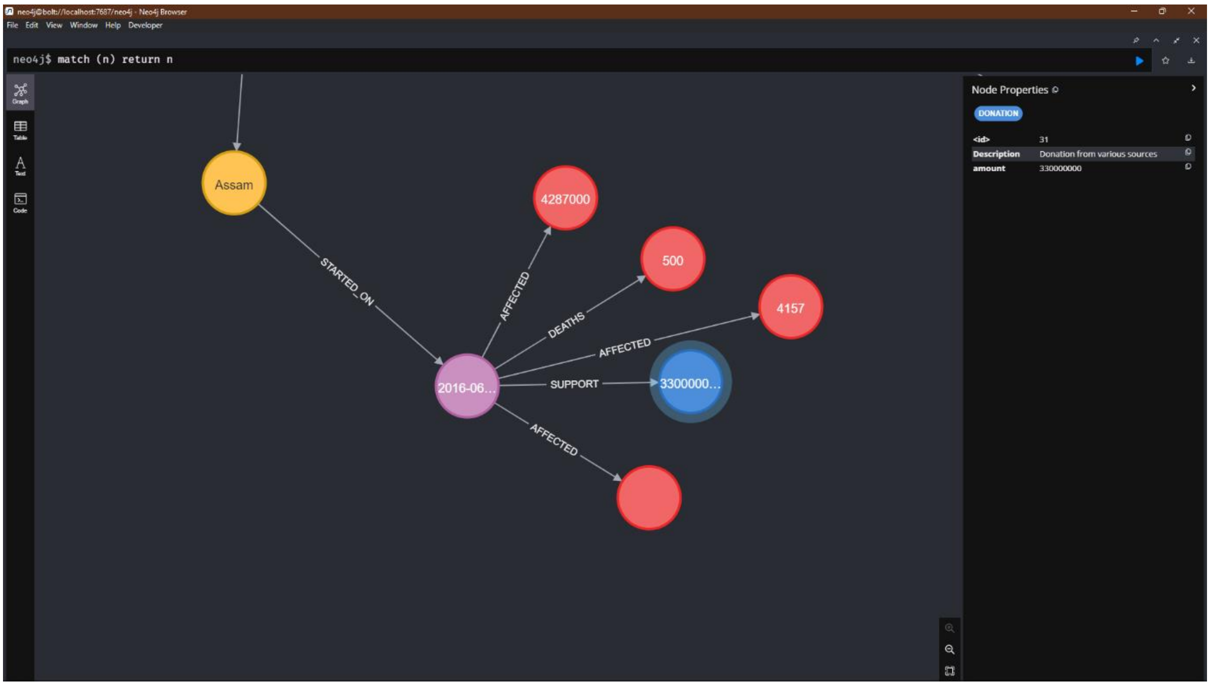

Figure 17: Knowledge Graph of 2016 Assam Flood

Let’s take a question to compare the graphs further What is the effect of 2021 Maharashtra Floods?

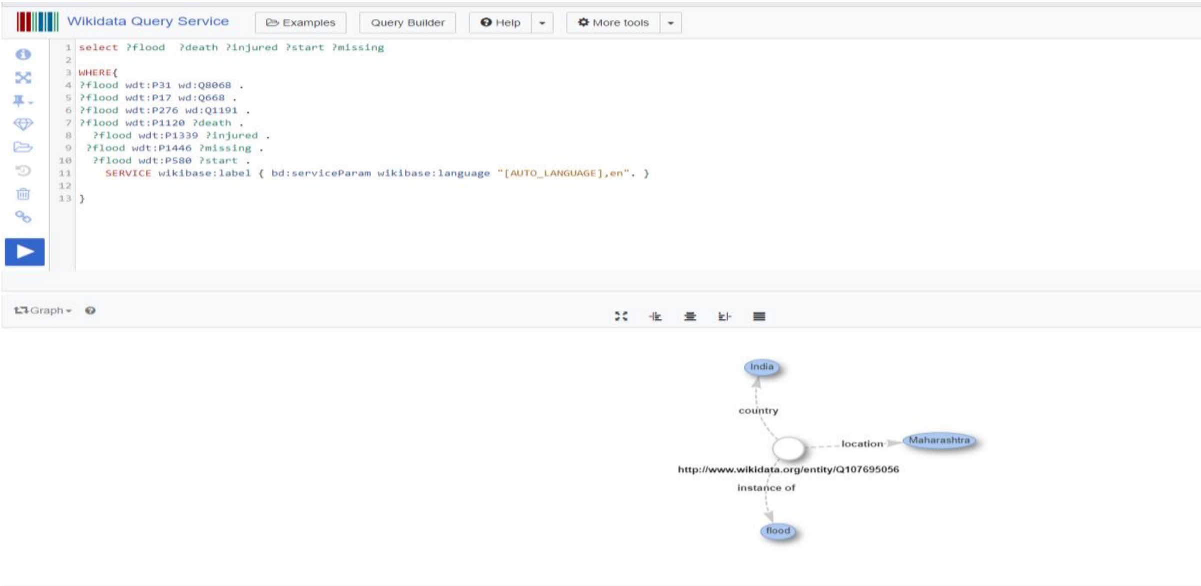

Figure 18: Wikidata of 2021 Maharashtra Flood

Figure 19: Wikidata of 2021 Maharashtra Flood in Table Format

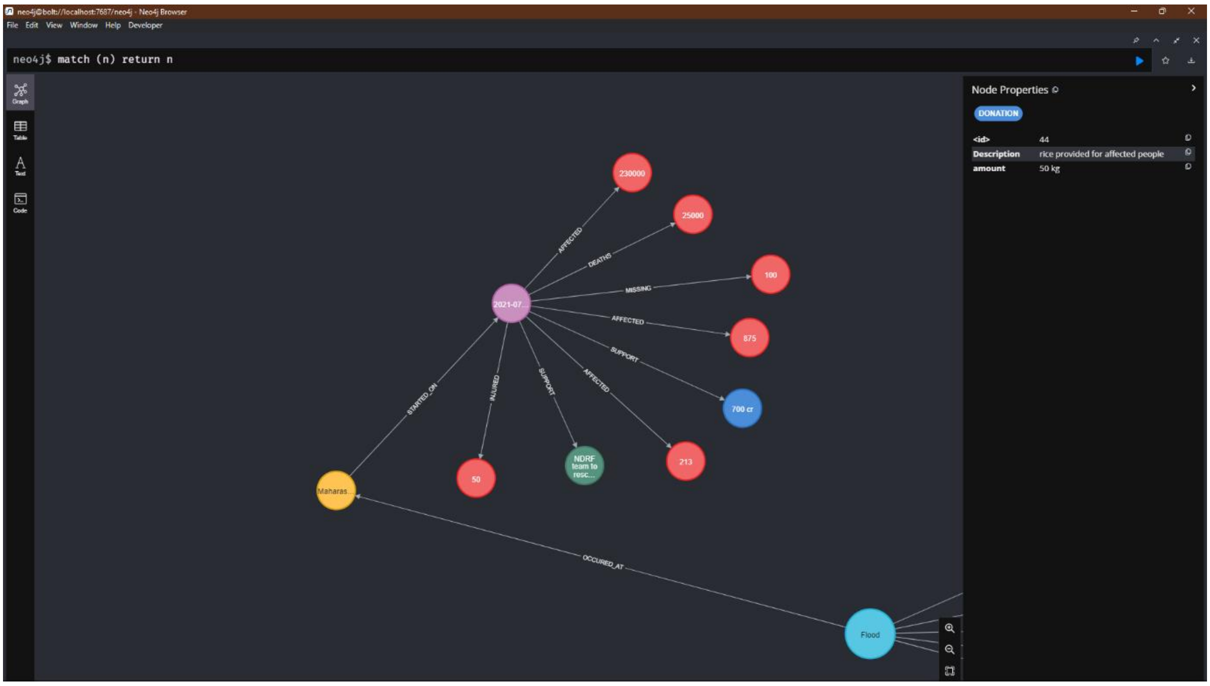

Figure 20: Knowledge Graph of 2021 Maharashtra Flood

From the above example, it can be observed that both wikidata and our knowledge graph provides relevant information regarding the effect of the event that occurred. But our knowledge graph provides much more information and that too in separate nodes which improves readability whereas we have to look at the table to get more information in wikidata.

6.2 Conclusion and Future work

Disaster related casualties have an extensive impact in our society. Disaster management focuses on preventing and minimizing the risks these scenarios face humanity with. Social media, although raw, presents a way of using humans as sensors to detect such hazards with the utmost brevity. Our model builds upon existing tools, contributing with a dynamic way of extracting and representing spatial and temporal relationships, as well as providing this knowledge to decision-makers, in real-time. As we've already completed the model design for informative and non-informativeness, humanitarian classes using BERT and CNN along with that we are done with NER Bi LSTM model. So, this project can be further enhanced by including other disasters as well so that more people will benefit from this program and can be made accessible to more people by creating a web application where we can display this information present in Neo4j about all the disasters to several organizations like NGOs and government which will help them in providing support to potentially lakhs of people.

References

1. Dhanya, V.G., Jacob, M.S., Dhanalakshmi, R. (2021). Twitter-Based Disaster Management System Using Data Mining. In: Pandian, A., Fernando, X., Islam, S.M.S. (eds) Computer Networks, Big Data and IoT. Lecture Notes on Data Engineering and Communications Technologies, vol 66. Springer, Singapore.
2. Chanda, Ashis Kumar. "Efficacy of BERT embeddings on predicting disaster from Twitter data." *arXiv preprint arXiv:2108.10698* (2021).
3. Singh, Jyoti Prakash, et al. "Event classification and location prediction from tweets during disasters." *Annals of Operations Research* 283.1 (2019): 737-757.
4. Boné, João, et al. "DisKnow: a social-driven disaster support knowledge extraction system." *Applied Sciences* 10.17 (2020): 6083.
5. Hu, Yingjie, and Jimin Wang. "How do people describe locations during a natural disaster: an analysis of tweets from Hurricane Harvey." *arXiv preprint arXiv:2009.12914* (2020).
6. Lee K, Ganti R, Srivatsa M, & Mohapatra P (2013) Spatio-temporal provenance: identifying location information from unstructured text. In 2013 I.E. International Conference on Pervasive Computing and Communications Workshops (PERCOM workshops) (pp. 499– 504). IEEE. https://doi.org/10.1109/PerComW.2013.6529548.
7. Kumar, Abhinav, and Jyoti Prakash Singh. "Location reference identification from tweets during emergencies: A deep learning approach." *International journal of disaster risk reduction* 33 (2019): 365-375.
8. Hecht, B., Hong, L., Suh, B., & Chi, E. H. (2011). Tweets from justin bieber's heart: the dynamics of the location field in user profiles. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 237–246). ACM.
9. Kumar, A., Singh, J. P., & Rana, N. P. (2017). Authenticity of geo-location and place name in tweets. In *23rd Americas Conference on Information Systems (AMCIS)*.
10. Huang, Q., Cao, G., & Wang, C. (2014). From where do tweets originate?: a gis approach for user location inference. In *Proceedings of the 7th ACM SIGSPATIAL International Workshop on Location-Based Social Networks* (pp. 1–8). ACM.
11. Nakaji, Y., & Yanai, K. (2012). Visualization of real-world events with geotagged tweet photos. In *Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference on* (pp. 272–277). IEEE.

12. Yuan, Q., Cong, G., Ma, Z., Sun, A., & Thalmann, N. M. (2013). Who, where, when and what: discover spatio-temporal topics for twitter users. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 605–613). ACM.
13. Cheng, Z., Caverlee, J., & Lee, K. (2010). You are where you tweet: a content-based approach to geo-locating twitter users. In *Proceedings of the 19th ACM international conference on Information and knowledge management* (pp. 759–768). ACM.
14. Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. In *ICWSM*.
15. Itoh, M., Yoshinaga, N., & Toyoda, M. (2016). Spatio-temporal event visualization from a geo-parsed microblog stream. In *Companion Publication of the 21st International Conference on Intelligent User Interfaces* (pp. 58–61). ACM.
16. Malmasi, S., & Dras, M. (2015). Location mention detection in tweets and microblogs. In *International Conference of the Pacific Association for Computational Linguistics* (pp. 123–134). Springer.
17. Middleton, S. E.,Middleton, L., &Modafferi, S. (2014). Real-time crisis mapping of natural disasters using social media. *IEEE Intelligent Systems, 29*, 9–17.
18. Zhang, W., & Gelernter, J. (2014). Geocoding location expressions in twitter messages: A preference learning method. *Journal of Spatial Information Science, 2014*, 37–70.
19. Gelernter, J., & Mushegian, N. (2011). Geo-parsing messages from microtext. *Transactions in GIS, 15*, 753–773.
20. Giridhar, P., Abdelzaher, T., George, J., & Kaplan, L. (2015). On quality of event localization from social network feeds. In *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on* (pp. 75–80). IEEE.
21. Unankard, S., Li, X., & Sharaf, M. A. (2015). Emerging event detection in social networks with location sensitivity. *World Wide Web, 18*, 1393–1417.
22. Xu., & Li, S. Knowledge graph visualization of intelligent emergency research. International Journal of Innovative Computing, Information and Control ICIC International ISSN 1349-4198 Volume 17, Number 4 (2021).
23. Son, J., Lim, C. S., Shim, H. S., & Kang, J. S. Development of Knowledge Graph for Data Management Related to Flooding Disasters Using Open Data. Future Internet, 13(5), 124 (2021).

24. Boné, J., Dias M., Ferreira J. C., & Ribeiro R. DisKnow: a social-driven disaster support knowledge extraction system. Applied Sciences, 10(17), 6083. (2020).

25. Maulana and W. Maharani, "Disaster Tweet Classification Based On Geospatial Data Using the BERT-MLP Method," 9th International Conference on Information and Communication Technology (ICoICT), 2021, pp. 76-81, doi: 10.1109/ICoICT52021.2021.9527513 (2021).
26. Madichetty, S., Muthukumarasamy, S., & Jayadev, P. Multi-modal classification of Twitter data during disasters for humanitarian response. *Journal of Ambient Intelligence and Humanized Computing*, 1-15 (2021).
27. Kersten, J., & Klan, F. What happens where during disasters? A Workflow for the multifaceted characterization of crisis events based on Twitter data. Journal of Contingencies and Crisis Management, 28(3), 262-280 (2020).
28. Madichetty S et al Re-ranking feature selection algorithm for detecting the availability and requirement of resources tweets during disaster. Int J Comput Intell IoT 1(2):207–211 (2018).
29. Yang, J., Yu, M., Qin, H., Lu, M., & Yang, C. A twitter data credibility framework—Hurricane Harvey as a use case. ISPRS International Journal of Geo-Information, 8(3), 111 (2019).
30. K. Utsu, A. Manaka, K. Nakafuri and O. Uchida, "Web application prototype for collecting disaster-related information focusing on tweets immediately after retweeting news posts," 4th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), Münster, pp. 1-6, 2017.10.1109/ict-dm.2017.8275696 2017.
31. Kwon, H. Y., & Kang, Y. O. Risk analysis and visualization for detecting signs of flood disaster in twitter. Spatial Information Research, 24(2), 127–139 (2016).
32. Ni, J., Liu, X., Zhou, Q., & Cao, L. A knowledge graph based disaster storyline generation framework. In 2019 Chinese Control And Decision Conference (CCDC) (pp. 4432-4437). IEEE (2019).
33. Alam, Firoj, et al. "CrisisBench: Benchmarking crisis-related social media datasets for humanitarian information processing." *arXiv preprint arXiv:2004.06774* (2020).
34. Olteanu, A.; Castillo, C.; Diaz, F.; and Vieweg, S. 2014. CrisisLex: A Lexicon for Collecting and Filtering icroblogged Communications in Crises. In ICWSM.

35. Imran, M.; Mitra, P.; and Castillo, C. 2016. Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages. In LREC.
36. Nguyen, D. T.; Al-Mannai, K.; Joty, S. R.; Sajjad, H.; Imran, M.; and Mitra, P. 2017. Robust Classification of Crisis-Related Data on Social Networks Using Convolutional Neural Networks. InICWSM, 632–635.
37. He, H.; Balakrishnan, A.; Eric, M.; Liang, P. Learning Symmetric Collaborative Dialogue Agents with Dynamic Knowledge Graph Embeddings. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, BC, Canada, 30 July–4 August 2017.