PHILIPPINE ENGINEERING JOURNAL
PEJ 2022; Vol. 43, No. 1: 65-88

Received: 17 January 2022
Accepted: 03 May 2022

A Text Mining Framework for the Classification and Prioritization of Disaster-Related Tweets for Disaster Response

Raymond Freth A. Lagria, Eugene Rex L. Jalao and Joanna Z. Resurreccion
Department of Industrial Engineering and Operations Research
College of Engineering, University of the Philippines Diliman
Quezon City, Philippines

Abstract — *Disasters have enormously disrupted the normal way of life of countries around the world and the Philippines is one of these countries. It is one of the most badly hit by disasters every year and due to its lower coping capabilities, it has constantly been ranked in the top 10 of the World Risk Index. This paper proposes a text mining framework that classifies and prioritizes disaster-related social media data, particularly Twitter tweets for the use of disaster managers for disaster response decision making. Validation of the framework during the classification stage resulted in an average of 90.67%, 99.25%, and 72.84% recall, on the test cases pertaining to training data and two different typhoon datasets. The prioritization module also prioritized tweets that were deemed urgent indicating the need for immediate response or attention.*

Keywords — classification, twitter, text mining, disaster management, disaster response

I. INTRODUCTION

For generations, disasters have enormously disrupted the normal functions and ways of life of countries and communities around the world. These brought tangible and intangible damages to properties and the lives and livelihood of those affected. Over the years, the engagement of governments and civic-society organizations in activities to diminish the effects of disasters have given birth to disaster risk reduction and management (DRRM).

The Philippines is one of the most disaster-affected countries, consistently being in the top ten of the World Risk Index [1, 2, 3, 4, 5]. Notable disasters in recent history include Typhoon Ondoy (2009), Bohol Earthquake (2013), Typhoon Yolanda (2013), Typhoon Rolly (2020) and the most recent is the Taal Volcano eruption in January 2020. The World Risk Report cites that the lack of technological capabilities and insufficient disaster management policies contribute to the Philippines' high world risk index. On a yearly average, disasters caused by natural hazards lead to 1,370 deaths and 1.2 billion dollars-worth of damages to the country [6, 7] with the highest recorded in 2013.

The Philippine government restructured and institutionalized its DRRM law – the Republic Act No.10121 [8] in 2010. It provides the legal basis for the installation of early warning systems; collaboration efforts among agencies; easier access to disaster funds (both emergency and preparedness); hazard/risk community mapping, and implementation of technology-based disaster prevention and mitigation systems and research.

In times of emergencies, systems in place fail in achieving "on-time" delivery of intended response due to different scenarios. For instance, the Philippine National Disaster Risk Reduction and Management Council (NDRRMC) had delays in its response to an emergency in August of 2018 because it declared a red alert warning, one day late. This warning should have directed all Philippine disaster management personnel to be on 24/7 duty to respond immediately after a calamity [9]. Another operational delay happened in an earlier typhoon in July of 2018 where the NDRRMC was not able to disseminate an orange rainfall warning among those affected [10]. This again led to delayed disaster preparation for response. These issues initiated a call for better disaster management coordination and government officials suggested tapping the latest technology in managing disasters. This specific call also came after NDRRMC disseminated information on an earthquake one day late. Systems in place will fail at times making disaster response in the Philippines very challenging [11].

When it comes to disaster events, social media has become experimental and users are observed to be more engaging during disasters [12]. These types of platforms are identified as effective means of communication among people during disasters [13]. Social media users can also disseminate information in a public setting, may serve as witnesses through sharing, and contribute efforts to the disaster events currently happening [14].

Twitter is a social media platform established in 2006. It allows its users to post short messages (tweets) in public, now with a 280-character limit on every tweet. A lot of users, government and non-government agencies, have utilized social media to use hashtags (#) during disasters, to spread awareness, and to relay information to the public especially on Twitter [15]. Furthermore, researchers claimed that Twitter may be able to predict natural events using data mining [16] or aid first responders particularly during disasters [17].

Filipinos use social media in various ways – to react, to help, to inform, and to relay important news. It is estimated that 80.7% of the total Filipino population are active social media users[18]. Hootsuite also reported that Filipinos spent time on social media with an average of 4.28 hours in 2016 and 3.95 hours in 2017 which is greater than the rest of the world [19]. In 2021, the daily average internet usage time of Filipino users (aged 16-64) is at 10 hours and 56 minutes [18].

Twitter, in 2012, ranked 10th in terms of the total number of Twitter users at 9.5 million [20]. In 2017, the social media giant in the Philippines had around 68.3 million visits with an average of 12.03 minutes visit time per user – 2nd only to Facebook. In 2018, Twitter comes in as the 3rd social media platform with the highest activity in the Philippines at 30% total activity per user. Facebook (57%) comes 1st and Instagram (36%) comes 2nd in terms of activity [19]. By 2021, Twitter became the top 5 in terms of the percentage of all Filipino internet users aged 16-24 (62.7%) that use it [18]. Southeast Asia Twitter cites that the number of Filipino users grew larger than the global average growth rate of 12% and that Filipinos are very interested in three major things: entertainment, sports, and civil society involvement [21].

Thousands of Filipino internet users flock to Twitter during typhoons to spread critical information and contribute to rescue and relief operations. Likewise, the government utilizes Twitter for public announcements during emergencies [22]. Filipino users usually follow celebrities on Twitter but interestingly, the National Meteorological and Hydrological Services Agency (PAGASA) has 6.3 million followers (ranked 10th most followed Filipino Twitter account in the country) – this shows that Filipinos are also concerned with their geophysical environment. The media outlets ABS-CBN and GMA News do not fall behind with 7.8 million and 6 million followers which are also within the top 15 accounts being followed by Filipino users (based on the researchers' tally). Hence, it is only logical that Twitter data can be used to the advantage of disaster managers aiding in disaster management.

During emergencies, the local government units (LGUs) are the first to coordinate the efforts for disaster response. Furthermore, it is their primary responsibility to prepare and distribute relief supplies and manage evacuation centers for the affected. Apart from that, the LGU also leads in conducting rescue and recovery operations. A lot of LGUs have created their Twitter accounts to facilitate such activities in their localities. Currently, to the best of our knowledge, no system investigates these local tweets and prioritizes them in terms of urgency during disasters.

One of the most successful platforms abroad that uses microblog data for disaster response is the Artificial Intelligence for Disaster Response (AIDR) in Qatar. It is a text processing platform developed to automatically perform classification on crisis-related microblog communications by Imran, et al. [23]. AIDR empowered disaster managers by categorizing microblogs during disasters into classes such as "Damage" and "Needs". However, this system primarily focuses on classifying microblogs. Ultimately, there is a need to develop a non-existent text mining framework that can classify and prioritize Filipino-authored disaster-related tweets for faster disaster response.

The main objective of this research is to aid reduce the disaster response time of disaster responders and other agencies by classifying and prioritizing tweets. Specifically, this research proposes a framework that (1) processes Filipino-authored tweets, (2) determines the relevance of the tweets to the disasters, and (3) prioritizes tweets that might be of interest in disaster response and disaster planning activities by disaster decision-makers.

The rest of the paper is organized as follows: Section 2.0 describes related literature to classification and prioritization frameworks of disaster-related tweets. Section 3.0 discusses the proposed framework that aims to solve the problem and the methodology used to validate the solution. Section 4.0 details the results of the test cases used in the study. Section 5.0 concludes the paper and sets up future work.

II. REVIEW OF RELATED LITERATURE

The problem identified in this paper can be split into two components – classification and prioritization of disaster-related tweets for disaster managers. This section also explores literature on decision support systems that aid in disaster management.

2.1 Classification Methodologies

One of the techniques that has been prominently used in processing tweets is classification. It is a data mining task that aims to assign or predict certain items (e.g., tweets) with “class labels” or simply “classes”. Specific to tweets, the goal is to assign a class label to a certain tweet based on a selected set of labels (e.g., important or not important). Classification engages machine learning algorithms and learns based on given data (training data) to develop a resulting classification model with its attributes as criteria for predicting classes [24]. The usual practice is for the given data to be coded with class labels which serve as the training set data from which a classifier (classification algorithm) will learn rules. The resulting model becomes the classification model that can be used to predict future or unseen records. The assignment of classes can be done manually, crowd-sourced, or automatically by using a customized program. Classification of disaster-related tweets has drawn a lot of motivation to different research but along with it comes differences in class labels used and more importantly, comes disparities in classification algorithms used.

A common use of classifying tweets includes mainly content analysis like trend activity, tweets content patterns, user activity, and followers/following statistics. This includes an examination of functions of disaster social media [25]. Assessment of situational awareness may have been the first application that focused on classifying tweets with disaster information labels of natural hazard events like grassfire and river flooding [26]. This study used class labels such as “Warning” and “Flood Level”. This stipulated designating disaster information class labels to tweets even if no classification algorithm was used. A subsequent study discussed the importance of user behavior-based features of datasets on contributing to situational awareness during mass emergencies. It developed a set of features called the “Verma features” that focused on linguistically motivated features like personal style, subjectivity, and tone [27].

A separate study employed classification by classifying tweets into confirming “truths” or baseless “rumors” during an earthquake event [28]. It particularly explored trend activity and noise that the disaster generated on Twitter. In another study, the motivation for building a classification model was from a storm hitting a festival that explored “positive” and “negative” sentiments [29]. In the Philippines, a similar study was conducted after Typhoon Haiyan (2013) hit and used generic disaster information class labels like Disaster Relief and Expressions of Support. These classes were used to extract knowledge from the tweets. [14].

A more sophisticated usage of classifying disaster-related tweets comprises learning a classification model and then predicting the classes of tweets based on different goals. The situational awareness study mentioned before used Naïve Bayes and Maximum Entropy (MaxEnt) as their classification methods to extract information during mass emergencies [26, 27]. A study in 2010 combined different classification methods for event detection and location estimation during earthquakes and typhoons. This study rose to the development of an almost real-time warning system that can warn people around the epicenter of an earthquake and a path of a typhoon [30]. A study in 2011 revolved around event detection of non-disaster-related events which compared Support Vector Machines (SVM) and Naïve Bayes as the classifiers. This led to SVM having the advantage over the Naïve Bayes classifier in terms of event detection classification [13].

The class labels generated by Imran, et al. were used as criteria in another study for the development of a mobile application that classified tweets to be "Relevant" or "Not relevant" to the 2010 Chilean earthquake. The dataset used was then evaluated using four (4) classifiers and found that SVM and Random Forests were the best performers [31].

Another study focused on categorizing Hurricane Sandy (2012) related tweets based on the relevance and the theme of the message. It presented a richer set of features and a comprehensive set of class labels. A binary classification step was employed first to identify "Relevant" & "Irrelevant" tweets and then a second classification step was used to assign the set of classes to the tweets. The study assessed three classifiers namely, SVM, Naïve Bayes, and MaxEnt that resulted in having better performance (in terms of model accuracy) as compared to previous similar studies [32].

Tweets usually contain "sentiments" which are often the reflections of the users that post them. Sentiment analysis of tweets using classification is another focus during disasters. Terpstra and Stronkman [29] conducted an interpretative analysis of tweets related to the 2011 storm in Belgium using manually coded "Positive" and "Negative" sentiments. The datasets were based on citizen tweets for disaster relief to the affected. A similar study used "Positive" and "Negative" sentiments on publicly available tweets datasets of different languages that used SVM [33]. Another paper offered sentiment analysis application on Filipino and English disaster-related tweets which concentrated on eight (8) sentiments based on Plutchik's Wheel of Emotions, however, this resulted in having very low precision and recall scores [34].

2.2 Disaster Management Decision Support Systems using Microblogs

Decision support systems (DSS) are the kind of information systems that aid decision-makers through determinations, judgments, and actions for organizations. There are several types of DSS such as data-driven, knowledge-driven, and model-driven DSS. There have been numerous studies related to DSS for disaster management. However, most of these systems center on developing DSS using spatial data with multi-criteria decision analysis (MCDA) methods for earthquakes [35], geographic information system (GIS) data paired with remote sensing, and hydrologic models for flood disasters [36, 37, 38], hybrid meta-heuristics model for disaster response scheduling problems [39], deep learning and machine learning methods for classifying geotagged images for rescue operations [40] and operations research models model for disaster response and recovery operations [41, 42].

With the vast data contained in social media in general, there are also significant efforts in developing DSS's that utilize microblogs. One of the earliest applications is disaster event detection for earthquakes through the use of social sensors by Sakaki, et al. [30]. Their model used semantic analysis, classification, and Kalman filtering for location estimation of such events. Another system that focused on identifying events using social media is Twitcident which was developed by a group of researchers in the Netherlands [43]. Twitcident used classification for semantic enrichment by classifying messages into reports about casualties and damages. This resulted in improved search and filter on relevant information to incidents.

Microblog data has also been used in a system that assesses risks and damages of an incoming disaster event. Risks and damages are quantified by degrees that result from composite scores before and after a disaster event. The method that identified these scores is estimated using an index model (with the assistance of emergency experts) which was extracted from a Chinese microblog. The index model featured 3 levels of indices which are composed of factors such as type of disaster, the demand for food, activities being held before the disaster, and public preparedness. The model proved to efficiently assist in risk and damage assessment as part of disaster management activities [44].

Sentiments from tweets, combined with economic loss data and geo-location information, have also been used as input to an early warning system for hurricanes. One such study aggregated reverse geocoding, sentiment analysis, hashtag consideration, and frequency-based approaches to be able to answer how people, affected or not, respond to disasters. These responses to the disaster they experience are quantified via sentiment analysis and scoring. The research paper determined that there is a positive correlation between the relationship of the severity of damages in an area and the intensity of the disaster-related event in the same area [45]. A similar study was conducted which used sentiment analysis but was implemented in the Chinese microblog, Weibo. Weibo is similar to Twitter which allows its users to read and post short messages with a predetermined limit. This research primarily took advantage of the negative sentiments of Weibo users on the affected victims. The study proposed a framework that identifies disaster-related messages, filter the negative sentiments using machine learning methods such as SVM, and analyze the negative sentiments. Using an earthquake dataset, the study concluded that the framework is useful in post-disaster incidents like public crises [46].

One of the practical applications of DSS utilizing microblogs data is the system proposed by Chae, et al. which aims to analyze public behavior for disaster response planning. [47]. Their proposed DSS features four main components namely, spatial analysis, spatial decision support, temporal pattern analysis, and spatiotemporal visualization. All components take advantage of Twitter data considering the location and temporal attributes of tweets. Particularly, the spatial decision support merged multiple sources of location data since data Twitter lacked the number. The temporal analysis component then considers how users behave before, during, and after a disaster event which identifies abnormal situations. Their method displayed how spatial data can be combined with microblogs for response planning.

A similar visualization-based DSS was developed which combines semantic annotators, a classifier, and spatial map data for prioritizing emergency response in Italy. However, the DSS mainly featured the generation of impromptu crisis maps which are processed from Twitter data, geolocation, and methods for damage detection. Specifically, data from Twitter were analyzed based on mentions of damages and location information. The test cases provided promising results in terms of damage detection [48].

2.3 Prioritization Methods

Only a handful of literature was found related to the prioritization of disaster-related microblog posts and tweets. Strohmaier (2010) proposed two scales that indirectly measure the strength (Twichter) and the impact of earthquakes (Twicalli). It was hypothesized that the Twichter scale may be able to match with the published earthquake intensity. However, the study further posits that the Twicalli scale made more sense as it used effect on people and is much more relevant to a social scale. Twicalli was found to have consistent findings on Twitter based on visualization methods produced [49]. Twicalli was coined and was used in a study that confirms official earthquake intensities as part of an earthquake detection system in Chile [50]. The research resulted in very good precision quantities even in different locations and different languages. Lastly, a classification-based prioritization system for earthquake-related tweets in Turkey was proposed using SVM, Random Forests, and Naïve Bayes as the core classifiers. The research used a binary classification method with “High Priority” and “Low Priority” as the class labels. The results proved to contribute that the identification of high priority tweets with the assistance of the right location and the right time is indeed helpful [51].

Most of the related research focused on the applicability of classifiers in processing and analyzing microblogs posts and tweets. This study extends this classification into prioritization based on microblog data. Furthermore, the DSS found in the literature concentrated on post-disaster management, GIS-based information systems, flood incidents, and crisis mapping. Most of these DSS focused on combining geolocation data and microblog data. In contrast, this study explores the development of DSS from microblog data prioritization focusing on its application to (near real-time) disaster response. Moreover, this paper contributes to the body of knowledge through the development of a unique classification and database based on or for Filipino related data (e.g., new prioritization lexica for typhoons). There are existing disaster lexica but with the best effort done by the authors, no other disaster prioritization-based lexica have been developed yet.

III. METHODOLOGY

Figure 1 shows the proposed text mining framework that extracts, pre-processes, classifies tweets into relevant tweets, and finally, prioritizes tweets that can be used by disaster managers in disaster management.

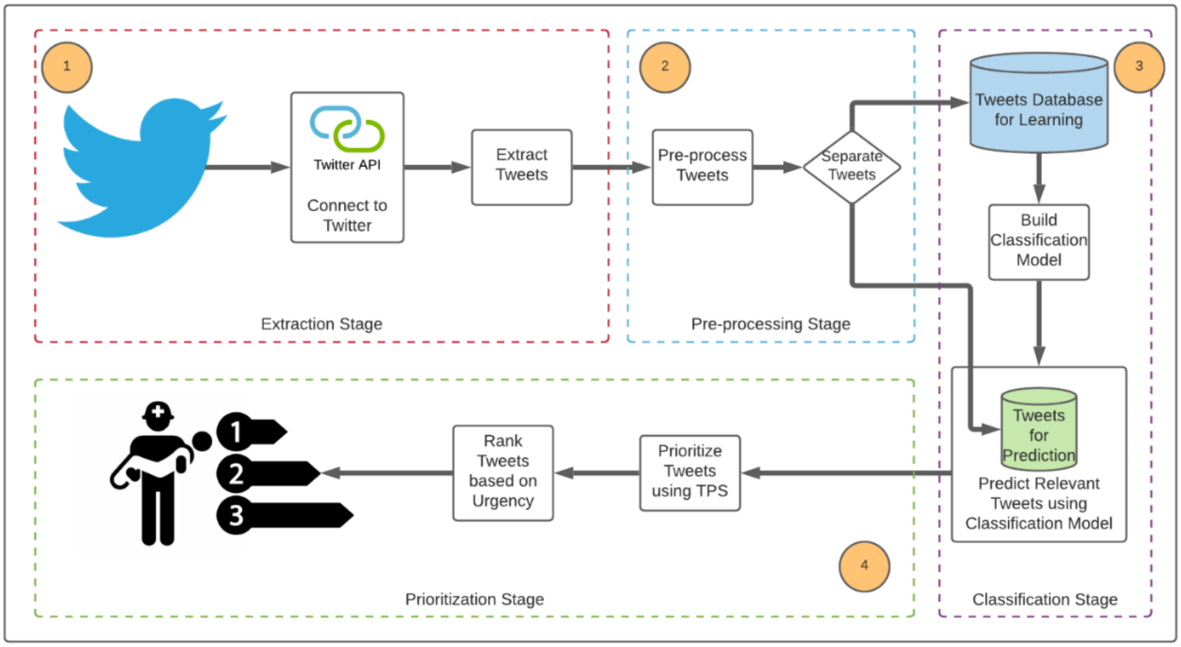


Figure 1. Text Mining Conceptual Framework

3.1 Extraction Stage

The first component of the proposed framework is the extraction stage which captures tweets based on search keywords and hashtags specified by the user. Keywords should be related to the natural hazard or the disaster event happening (e.g., typhoon name, #haiyan). The extraction procedure is recommended to be conducted during the following events: before, during, and post-disaster timeline. This component is aligned with the incident managers' goal to monitor and respond to affected people efficiently in times of disasters.

The extracted data will be a collection of tweets in semi-structured table format with attributes/features as its columns. The extracted data will be further enriched to include additional attributes such as "UserType" and if the tweet "ContainsURL". The selection of these features will be based on features used by related research as well as the authors' selection. The complete feature set, along with its descriptions, is detailed in Table 1.

Table 1. List of Attributes/Features

Feature/Attribute	Type	Sample	Notes
Unigrams*	Binary	0, 1	Single terms part
isRetweet[1]	Categorical	Yes, No	If the tweet itself is a retweet
ReTweetCount[1]	Numeric	0,1,2, …	Retweet counts of the original tweet
ifMention[2]	Categorical	Yes, No	If there is a mention
ifReply**	Binary	0, 1	If the tweet is a reply
ContainsURL*	Categorical	Yes, No	If the tweet contains a URL
UserType[3]	Categorical	Layman	If the user is Government, Celebrity, etc.
Favorited**	Binary	0, 1	If the tweet is favorited at least once
FavoritedCount**	Numeric	0,1,2, …	The number of favorites on the tweet
ReTweeted**	Categorical	Yes, No	If the actual tweet is retweeted

*1-Becker, 2-Pekar, 3-Takahashi, 4-Vieweg, 5-Zhou, 6-Verma, *-All, **-default in TwitteR*

The categorical attributes may need to be transformed to binary digits (e.g., 0, 1) since some classifiers only accept numerical values. The attribute UserType adopts Takahashi's (2017) codes namely, NGO, Government, News Organization, Journalist, Lay People (Ordinary People), Celebrity and Others (e.g., personal blog, feature profile, meme profile, quotes profile, TV show). Takahashi explored the typology tweets relating to disaster information dissemination and relief efforts coordination [52].

3.2 Extraction and Pre-processing

The second stage of the framework is the pre-processing of tweets. Pre-processing helps in the removal of noise like unwanted characters or words. For instance, in text mining, the human language would usually not include the words "blah blah blah" or "Meh" as critical words. The following are the steps included in the pre-processing:

a. Convert non-alphanumeric characters to alphanumeric
b. Replacement of emoticons with spaces
c. Removal of the number sign "#"
d. Removal of the mentioned user (mentioned user starts with the character "@")
e. Removal of HTML links
f. Removal of numbers (will be focusing on letters only)
g. Converting everything to lower case
h. Removal of extra whitespaces
i. Removal of stopwords
j. Perform stemming algorithm – stems are root words with derivational affixes

For the stemming algorithm, this paper proposes a stemming algorithm that features lookup tables. Lookup tables are static tables that are used to look for matching values. Using lookup tables in this technological age has become exceedingly efficient and fast. The proposed framework will look for the stem of a term from a lookup table and replaces these terms with the corresponding stems in the concerned tweet. If there is no stem in the lookup table, then it uses the original term as the stem. Hence, the tweets are transformed to contain stems.

The term-stem lookup tables describe a repository of all terms collected by the authors and their equivalent stems. Different tables are maintained according to the 2 main languages used in the Philippines (e.g., English, Tagalog), thus catering to the need of handling multilingual tweets. Additional lookup tables are also maintained such as the Slang Words lookup table. This will contain slang words like “resq” for “rescue”. Table 2 lists down maintained tables:

Table 2. Maintained Lookup Tables

Table Name	**Description**
Stopwords - English	Master list of English stopwords
Stopwords - Tagalog	Master list of Tagalog stopwords
Language Tagalog Stem Lookup	Stems for Tagalog Words
Language English Stem Lookup	Stems for English Words
Slang Tagalog Stem Lookup	Stems for Slang Words

Text mining involves the creation of a document-term-matrix (DTM) as part of the text mining and modeling process. A DTM is a matrix that lists the frequency of terms that appear in a collection of documents or this paper – the tweets. The framework utilizes the DTM as a step to transforming its original form into the desired format ready for the classification stage. Table 3 displays a DTM example. The framework sets a threshold on the words that would appear in the columns (e.g., only terms that appear in 3% of all tweets are included). These words become part of the feature set in addition to the default set extracted from Twitter.

Table 3. Sample Document-Term-Matrix

Tweets	**Help**	**Us**	**Need**	**Water**	**We**	**Are**	**Flooded**	**Food**
Help us, need water	1	1	1	1	0	0	0	0
We are flooded	0	0	0	0	1	1	1	0
We need water food	0	0	1	1	1	0	0	1

The framework then separates the collection of tweets to be used in building the classification model and then the other set to be used in prediction accuracy validation.

3.1 Classification Modeling

Twitter does not filter tweets in the public space, hence, even if there is a disaster event in a certain area, the public Twitter stream still allows all users to post tweets. Moreover, even if a hashtag is used, there could be still be tweets that show sarcasm, memes, or punchlines during disaster events. This leads to the third stage of the framework which is classification modeling. A classification model is proposed to filter out only the relevant tweets to the disaster event. This component utilizes binary class labels “Relevant” and “Irrelevant”. The classification stage proposes three different classifiers namely, Naïve Bayes, Support Vector Machines, and Random Forests. As identified by related literature, these three classifiers were found to be commonly used when it comes to disaster-related text data.

3.1.1 Naïve Bayes (NB)

The Naïve Bayes classifier is family of algorithms that are based on Bayes Theorem [53]. It is a classifier that allows to predict a class primarily based on probabilities. Naïve Bayes was found to have a practical applicability in text mining problems because it tends to simplify its features through the usage of frequency of terms. A simple explanation using the Naïve Bayes classifier starts with the Naïve Bayes theorem:

$$\boldsymbol{P(A|B) = \frac{P(B|A)x\,P(A)}{P(B)}} \qquad \text{eq. (1)}$$

The NB classifier finds the probability of a class of a tweet similar to the following:

$$\boldsymbol{P(Relevant|Help\ us\ here\ in\ Quezon\ City) = \frac{P(Help\ us\ here\ in\ Quezon\ City|Rescue)x\,P(Relevant)}{P(Help\ us\ here\ in\ Quezon\ City)}} \qquad \text{eq. (2)}$$

Eq. (2) looks to calculate the probability of a tweet "Help us here in Quezon City" to be classified as a "Relevant" tweet. NB assumes that every term in the tweet are independent (Naïve) and the probability of the event P(B) or P(Relevant) becomes:

$$\boldsymbol{P(Help\ us\ here\ in\ Quezon\ City|Rescue) = P(Help)x\,P(us)x\,P(here)x\,P(in)x\,P(Quezon)x\,P(City)} \qquad \text{eq. (3)}$$

Eq. (3) is a transformation that is also applicable in the conditional probabilities. This produces a network of conditional probabilities that classify using terms found in tweets.

3.1.2 Support Vector Machines (SVM)

SVM is a non-probabilistic classifier that classifies unseen objects (response variables) into two separate groups based on the attributes of each data point by creating a linear partition or a hyperplane between the two classes [54]. The separated data points are called support vectors. A practical example would be dividing a set of tweets into two separate groups – Relevant and Irrelevant. Essentially, SVM places a tweet either above or below the linear partition based on the tweets' attributes. Figure 2 shows an example of data points separated by a line and a hyperplane. The goal of SVM is to find the greatest margin (represented by the hyperplane) between the two classes. The figure also shows separation in 3D but the underlying concept is to find higher dimensions that would produce a hyperplane that can segregate classes.

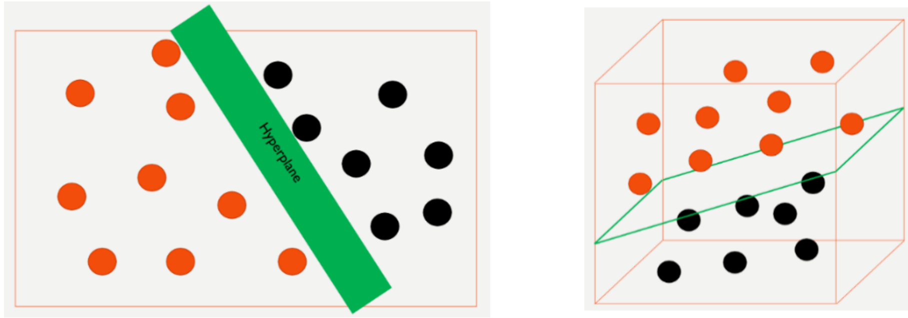


Figure 2. SVM Hyperplane in 2D and 3D

3.1.3 Random Forests (RF)

RF is a classification model that operates by building an ensemble of decision trees. A decision tree (DT) is another classification model in the form of a tree structure. DTs break down datasets into smaller subsets until a leaf node (rule for class prediction) is reached. RF then builds multiple decision trees to provide a more accurate classification. The final predicted class of an object is chosen by averaging or by voting. The diagram below shows a simple explanation of how RF induction works.

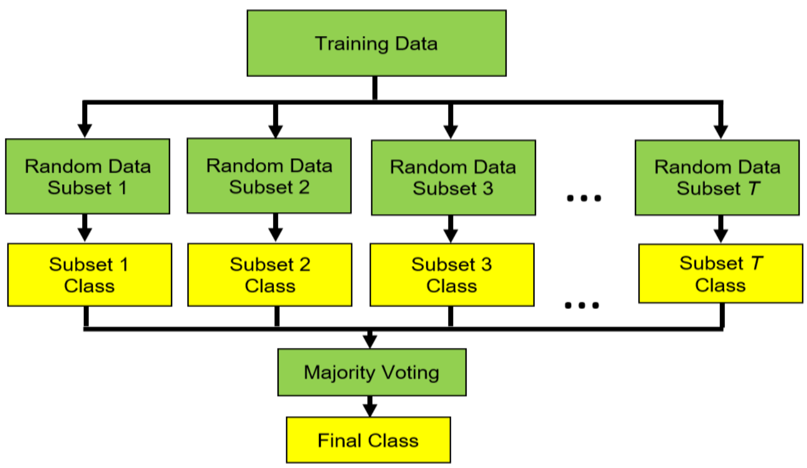


Figure 3. Illustrating Random Forests Classifier

Given training data, RF induces its classification model based on the following:

- learn a number of subset DTs, T, based on random subsets of N cases or records
- subsets, T, will have a number of attributes, m, where m = ½√p, √p, or 2√p

Each random subset will produce a different decision tree and will provide different predicted classes. The two points above ensures that randomness is embedded in each sub-classifier. The final step in the algorithm is a majority-voting scheme that provides the final predicted class for each record.

3.2 Prioritization Module

After the classification model classifies tweets as “Relevant” or “Irrelevant”, the tweets are subjected to the prioritization process. Figure 4 shows the general steps:

Figure 4. Prioritization Module Steps

The classified relevant tweets will be prioritized based on the degree of urgency. The proposed method provides five (5) different scores for each prioritization level. To score each relevant tweet, a lexicon of DRRM word lists with pre-assigned prioritization scores according to urgency was developed and will be used as the scoring reference. The basic algorithm is to sum up the scores of each word (Tweet Prioritization Score or TPS) in a tweet to have a total urgency scoreEach tweet will then have its total TPS. The scored tweets will then be sorted and ranked according to the highest TPS. The result of the prioritization module is a list of ranked tweets to be used by the intended users.

To the best of our knowledge, current literature does not discuss DRRM words with corresponding “urgency” scores. However, there are existing lexica that keep DRRM related word lists which were used as bases in developing the **TPS lexicon**. Crisis Lexicon preserves two lists - CrisisLex Lexicon and EMTerms (Terminological Resource for Crisis Tweets). These two lexica are used as part of the TPS lexicon in addition to the authors’ DRRM word lists. To add more validity, the TPS description also adapts the severity rating component statements of a risk scoring matrix tool which depicts risks as impact to a group of people [55]. Table 4 describes each TPS score while the lexica are available upon request from the authors.

Table 4. Tweets Prioritization Score (TPS) Descriptions

TPS	Tweets Contain Words having the Subject of	Description
-1	Others, Humor, Unrelated, Irrelevant Commentary	Words that generate negative effect to urgency and pose no significance to the urgency
1	Prayers, Sentiments, Relevant Commentary, Announcements	Words that produce sentiments and impart small impact to urgent attention
2	Weather Information Updates, Damage, and Injury Reports	Words that inform disaster-related status and may potentially be called for attention
3	Evacuation Information, Relief Goods	Words that attract substantial attention from disaster managers
6	Urgency, Rescue, Help, First Aid, First Response	Words that contribute to extreme need and urgency. Calls for immediate attention from responders

The TPS is based on the presence of words depicting the subject matter as described in Table 4. Currently, the TPS Lexicon contains both English and Filipino words. Urgency is expressed by ultimately considering how the immediate response is depicted from a tweet. A score of negative one (-1) will be assigned to those words that would de-prioritize a tweet, enabling words that contribute to humor, and irrelevance to disasters. On the other side, a score of positive six (6) indicates words that are disaster-urgent and needs immediate attention. The exclusion of numeric values 4 and 5 for the TPS is supported by using a simulation process done before the TPS scheme was finalized. The simulation was initiated using different scales of points (e.g., 1 to 5 or -1 to 4). By using the proposed TPS (i.e., -1, 0, 1, 2, 3, 6), the most urgent terms are given the highest emphasis and exaggerate the urgency of tweets that the TPS aims to prioritize.

New words can be inserted into the TPS Lexicon as the words relating to disaster urgency evolve over time. It is recommended that the addition of new words and phrases into the TPS Lexicon should be done every noteworthy disaster that the country experiences. However, a weekly check and update of words based on tweeting dynamics by Filipinos are also recommended. The authors take note that in every disaster event, there could be different dynamics and behavior in tweeting. Also, it is highly suggested to develop different TPS Lexicon for different types of disasters (e.g., typhoon, earthquake) in the future.

The proposed framework shows that the classification module acts as preliminary screening of relevant tweets and the prioritization component assigns the quantitative degree of urgency to tweets enough to call the attention of the disaster managers or disaster monitoring personnel for disaster response and disaster preparedness.

3.5 End-users

The end-users, which may comprise disaster managers, monitoring personnel, responders, or DRRM officers, will consider the results of the prioritization module during critical disaster time decision making. In the Philippine setting, the NDRRMC, along with its national partner agencies, carries out its Pre-disaster Assessment-Action, Plans, and Protocols to prepare and assess potential disasters of its risks which happens before a disaster event. After this, the

NDRRMC activates its Response Cluster and Incident Management Teams (IMT) that deploys response teams. The response is characterized by critical response and relief operations provided to the affected LGUs who requested assistance. In this sense, the end-users can make use of the system before, during, and after a disaster event. As long as people are tweeting, disaster managers can take advantage of social media. The end-users have to decide and act upon seeing the list of prioritized tweets. To summarize, the following use case diagram was developed to demonstrate the different use cases of the proposed framework.

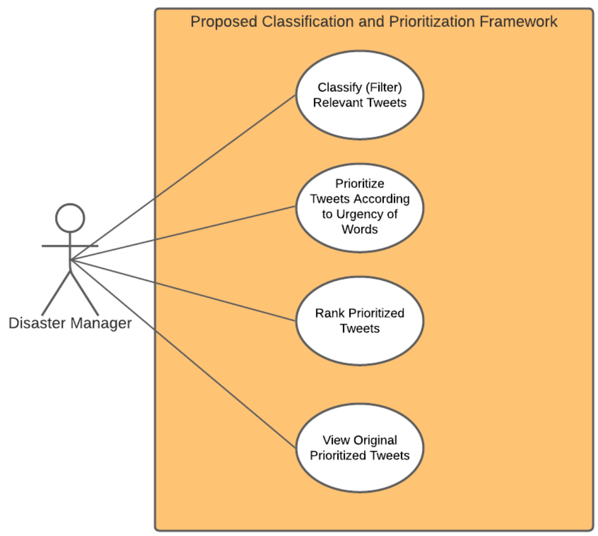


Figure 5. Proposed Framework Use Case Diagram

Currently, Twitter has a location-based tagging of its tweets. However, in the Philippines, it was observed by the authors that most of the Filipino users do not turn on their location/GPS services while tweeting, hence, the majority of tweets lack location tags. However, while the proposed framework lacks location-tagged tweets, the disaster responders should know ahead of time where the disaster event might happen or had happened. Furthermore, upon viewing original prioritized tweets, the disaster managers will have the chance to look into the original tweets which may contain locations and contact numbers.

IV. RESULTS AND DISCUSSION

In this paper, the framework is tested and validated by implementing the framework on several test cases. The framework and test procedures were programmed using the R language. The following diagram depicts the methodology used in testing and validating the framework:

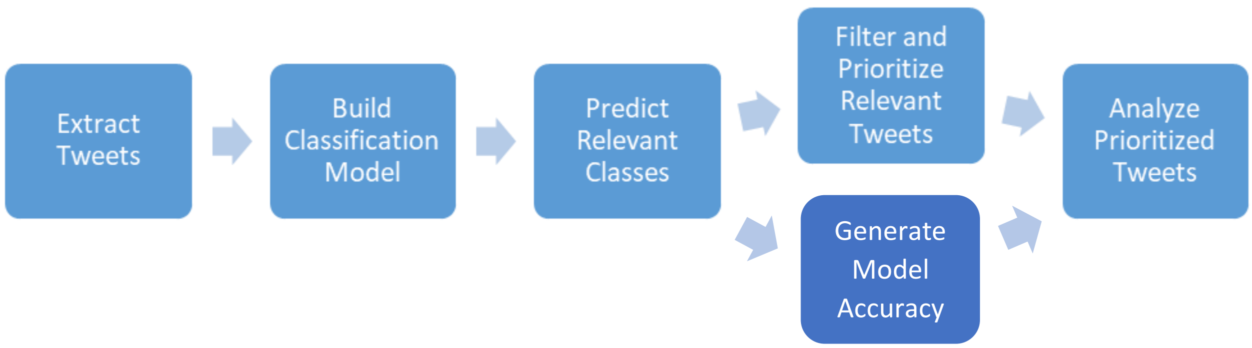
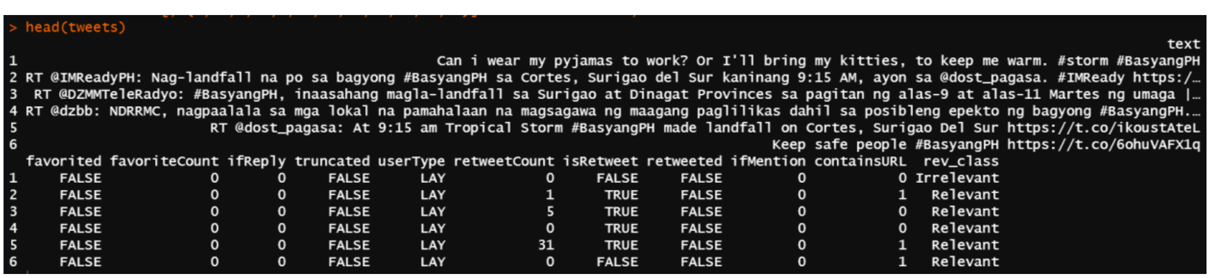


Figure 6. Methodology for Evaluating the Proposed Framework

The package twitteR in R was used to facilitate the extraction of tweets using the free version. To be able to build the classification model, a total of 3,906 tweets were collected from 2017-2020 in different typhoon events with non-disaster event tweets. twitteR contains default attributes such as "ReTweetCount" and "isRetweet". As mentioned previously, new attributes like "ContainsURL" and "UserType" are added to enrich the dataset. The tweets were then coded manually with the class labels "Relevant" or "Irrelevant" by the authors and are used as the main classification model of the framework. On top of default and added attributes, DTM attributes are also augmented. Examples of DTM attributes are "rescue" and "typhoon". These DTM attributes are integer-only values counting of the number term occurrences on the corresponding tweet. The derivation of DTM attributes was done through a combination of custom R code and the package tm. A snippet of the training data is shown in the figure below.



```
> head(tweets)
                                                                                                                        text
1                                                       Can i wear my pyjamas to work? Or I'll bring my kitties, to keep me warm. #storm #BasyangPH
2 RT @IMReadyPH: Nag-landfall na po sa bagyong #BasyangPH sa Cortes, Surigao del Sur kaninang 9:15 AM, ayon sa @dost_pagasa. #IMReady https:/…
3  RT @DZMMTeleRadyo: #BasyangPH, inaasahang magla-landfall sa Surigao at Dinagat Provinces sa pagitan ng alas-9 at alas-11 Martes ng umaga |…
4 RT @dzbb: NDRRMC, nagpaalala sa mga lokal na pamahalaan na magsagawa ng maagang paglilikas dahil sa posibleng epekto ng bagyong #BasyangPH.…
5                        RT @dost_pagasa: At 9:15 am Tropical Storm #BasyangPH made landfall on Cortes, Surigao Del Sur https://t.co/ikoustAteL
6                                                                                    Keep safe people #BasyangPH https://t.co/6ohuVAFX1q
  favorited favoriteCount ifReply truncated userType retweetCount isRetweet retweeted ifMention containsURL  rev_class
1     FALSE             0       0     FALSE      LAY            0     FALSE     FALSE         0           0 Irrelevant
2     FALSE             0       0     FALSE      LAY            1      TRUE     FALSE         0           1   Relevant
3     FALSE             0       0     FALSE      LAY            5      TRUE     FALSE         0           0   Relevant
4     FALSE             0       0     FALSE      LAY            0      TRUE     FALSE         0           0   Relevant
5     FALSE             0       0     FALSE      LAY           31      TRUE     FALSE         0           1   Relevant
6     FALSE             0       0     FALSE      LAY            0     FALSE     FALSE         0           1   Relevant
```

Figure 7. Sample Snippet of the Relevant and Irrelevant Tweets

For the classification stage, the following table shows the RWeka functions used to build the classification models per classifier. Each model uses the same set of attributes and training data (Juan) for consistency. The reader will find the parameters for each classifier in Appendix A.

Table 5. Classification Model Parameters

Classifier	Model Building and Attributes
NB	`library(RWeka)` `tweets_df_data_NBmodel <- NaiveBayes(rev_class ~ favorited + favoriteCount + ifReply + userType + truncated + retweetCount + isRetweet + retweeted + ifMention + containsURL + bagyong + hundred + tropical + safe + people + walangpasok + typhoon + haha + city + level + tomorrow + wala + help + lang + rescue + stay + lakas + hangin + pls + need + rescueph + jessica + jessicasoho + spirit + awareness, data = tweets_df_data)`
SVM	`library(RWeka)` `tweets_df_data_SVMmodel <- SMO(rev_class ~ favorited + favoriteCount + ifReply + userType + truncated + retweetCount + isRetweet + retweeted + ifMention + containsURL + bagyong + hundred + tropical + safe + people + walangpasok + typhoon + haha + city + level + tomorrow + wala + help + lang + rescue + stay + lakas + hangin + pls + need + rescueph + jessica + jessicasoho + spirit + awareness, data = tweets_df_data)`
RF	`RF = make_Weka_classifier("weka/classifiers/trees/RandomForest")` `tweets_df_data_RFmodel <- RF(rev_class ~ favorited + favoriteCount + ifReply + userType + truncated + retweetCount + isRetweet + retweeted + ifMention + containsURL + bagyong + hundred + tropical + safe + people + walangpasok + typhoon + haha + city + level + tomorrow + wala + help + lang + rescue + stay + lakas + hangin + pls + need + rescueph + jessica + jessicasoho + spirit + awareness, data = tweets_df_data)`

To evaluate the performance of the classification stage, three models are built using the three classifiers on the main training dataset named Juan. Juan is the dataset from which the main classification model is built and will be used to predict the relevant classes on the other two datasets, Ulysses and Odette. These two are used for measuring model performance only. Each model produces a table called confusion matrix which can be used to calculate model accuracy. Classification accuracy is defined as the performance metric that summarizes how a classification model performs in terms of predicting the number of correctly predicted class labels, may they be relevant or irrelevant classes. Table 6 describes the datasets used.

Table 6. Test Dataset Details

Dataset (n)	Relevant Classes	Validation Method	Years Collected
Juan (3906)	1,453	10-fold Cross-Validation	2017-2020
Ulysses (1289)	893	Test scenario	2020
Odette (1847)	1847	Test scenario	2021

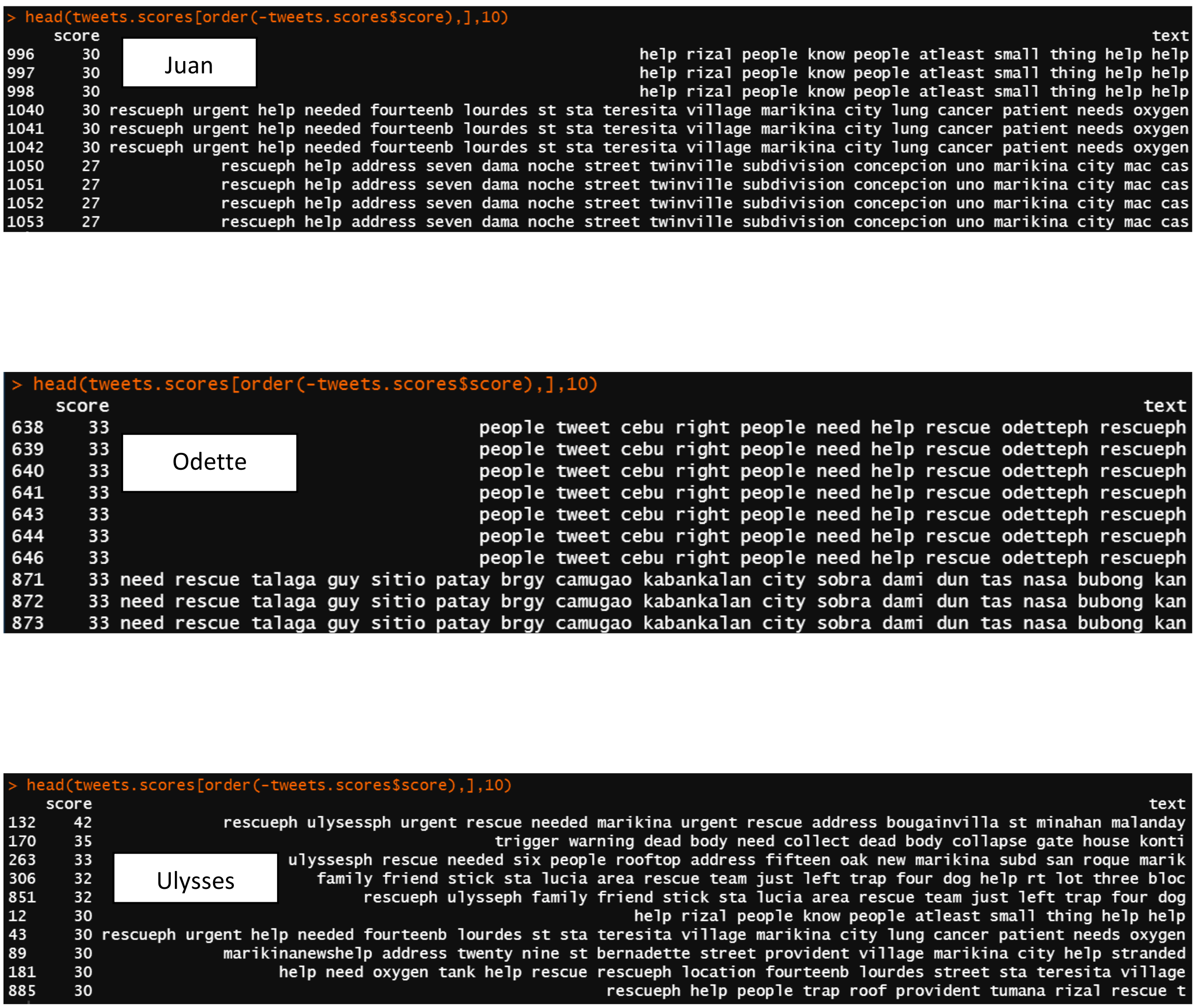
Datasets Ulysses and Odette were also manually coded taking note that these are imbalanced datasets. Ulysses contains several irrelevant classes while Odette contains only relevant ones. This setup tests a scenario where during typhoons where all tweets are relevant. However, not all tweets in disaster events would always be relevant, hence, the classification stage. The resulting classification model was run against the three datasets. The results of the models generated an average of 76.52%, 69.02%, and 90.18% accuracy (see Table 6) respectively which are already acceptable and better than a random model that predicts 50-50.

Table 7. Classification Modeling Stage Performance

Dataset	**Classifier**	**Accuracy**		**Precision**	**Recall**	
Juan	NB	62.54%	Average 79.76%	57%	97.10%	Average 90.67%
	SVM	86.05%		86.90%	84.10%	
	RF	90.68%		90.20%	90.80%	
Ulysses	NB	71.37%	Average 69.74%	71.13%	98.77%	Average 99.25%
	SVM	69.20%		69.25%	99.89%	
	RF	68.66%		69.09%	99.10%	
Odette	NB	97.51%	Average 72.84%	100%	97.51%	Average 72.84%
	SVM	57.50%		100%	57.50%	
	RF	63.51%		100%	63.51%	

To further validate the effectiveness of the classification model against the datasets, additional performance measures were also calculated. Precision is the percentage of the tweets which are "relevant" while recall is the percentage of total "relevant" tweets correctly classified. Since the proposed framework is highly dependent on relevant tweets to be subjected to the prioritization module, it is important to evaluate the performance in terms of recall. In this regard, the focus of the framework is minimizing the false negatives. This translates to minimizing the prediction of "irrelevant" tweets which are actually relevant. The two test scenarios generated average recall values of 99.25% and 72.84% for Ulysses and Odette, respectively, as seen in Table 7. These indicate that the classification stage of the proposed framework is performing very well (above 70% can be considered a good recall) in predicting relevant tweets while minimizing error in predicting irrelevant tweets which are actually relevant.

After the framework generated relevant tweets, these tweets were subjected to the prioritization stage. To demonstrate the prioritization scheme, this paper used the NB predictions to generate the relevant tweets because NB generated the highest average in terms of accuracy. In a more realistic scenario, the disaster manager can be presented with different predictions used by these three classification models. After the relevant tweets are filtered, the assignment of TPS was then conducted using another custom R code. Figure 8 shows the results of the first 10 prioritized tweets ranked from the highest scored to the lowest scored from each of the datasets.



```
> head(tweets.scores[order(-tweets.scores$score),],10)
     score                                                                                         text
996     30                                              help rizal people know people atleast small thing help help
997     30                                              help rizal people know people atleast small thing help help
998     30                                              help rizal people know people atleast small thing help help
1040    30 rescueph urgent help needed fourteenb lourdes st sta teresita village marikina city lung cancer patient needs oxygen
1041    30 rescueph urgent help needed fourteenb lourdes st sta teresita village marikina city lung cancer patient needs oxygen
1042    30 rescueph urgent help needed fourteenb lourdes st sta teresita village marikina city lung cancer patient needs oxygen
1050    27          rescueph help address seven dama noche street twinville subdivision concepcion uno marikina city mac cas
1051    27          rescueph help address seven dama noche street twinville subdivision concepcion uno marikina city mac cas
1052    27          rescueph help address seven dama noche street twinville subdivision concepcion uno marikina city mac cas
1053    27          rescueph help address seven dama noche street twinville subdivision concepcion uno marikina city mac cas
```



```
> head(tweets.scores[order(-tweets.scores$score),],10)
    score                                                                          text
638    33                              people tweet cebu right people need help rescue odetteph rescueph
639    33                              people tweet cebu right people need help rescue odetteph rescueph
640    33                              people tweet cebu right people need help rescue odetteph rescueph
641    33                              people tweet cebu right people need help rescue odetteph rescueph
643    33                              people tweet cebu right people need help rescue odetteph rescueph
644    33                              people tweet cebu right people need help rescue odetteph rescueph
646    33                              people tweet cebu right people need help rescue odetteph rescueph
871    33 need rescue talaga guy sitio patay brgy camugao kabankalan city sobra dami dun tas nasa bubong kan
872    33 need rescue talaga guy sitio patay brgy camugao kabankalan city sobra dami dun tas nasa bubong kan
873    33 need rescue talaga guy sitio patay brgy camugao kabankalan city sobra dami dun tas nasa bubong kan
```



```
> head(tweets.scores[order(-tweets.scores$score),],10)
    score                                                                                                  text
132    42              rescueph ulysessph urgent rescue needed marikina urgent rescue address bougainvilla st minahan malanday
170    35                                         trigger warning dead body need collect dead body collapse gate house konti
263    33                   ulyssesph rescue needed six people rooftop address fifteen oak new marikina subd san roque marik
306    32                      family friend stick sta lucia area rescue team just left trap four dog help rt lot three bloc
851    32                          rescueph ulysseph family friend stick sta lucia area rescue team just left trap four dog
12     30                                                     help rizal people know people atleast small thing help help
43     30 rescueph urgent help needed fourteenb lourdes st sta teresita village marikina city lung cancer patient needs oxygen
89     30              marikinanewshelp address twenty nine st bernadette street provident village marikina city help stranded
181    30                   help need oxygen tank help rescue rescueph location fourteenb lourdes street sta teresita village
885    30                                                rescueph help people trap roof provident tumana rizal rescue t
```


Figure 8. Prioritized Tweets

Inspecting the sample prioritized tweets, we see that most tweets seek immediate response (e.g., rescue, needs oxygen, trap). Furthermore, these tweets contain most of the words from the TPS 6 category indicating that the framework is consisted and is verified. Reading further into the tweets and content, several prioritized tweets generated addresses of people needing the actual disaster response. These prioritized tweets can then be traced back to the original form of the tweets to find more information about the user (e.g., location details and photos). Through visual inspection, it can be said that the proposed framework indeed generated prioritized tweets that can be used by disaster managers in their response actions.

Lastly, as part of the implementation plan with the inclusion of the proposed framework, a set of steps that a disaster manager can take from the results can be the following:

1. Check the prioritized tweets if there are specific locations or nearest landmarks mentioned; match the tweets to where the typhoon has struck (it is noted that the current model is limited to what the content contains, location estimation is not part of the scope of the model)
2. Check if there are contact numbers mentioned
3. Determine the type of response needed
4. Identify the next plan of action and which agency or responder is needed
5. Communicate with and organize the disaster responders for the response/prevention
6. Execute disaster response/prevention

V. CONCLUSION AND FUTURE WORK

This paper proposed a text mining framework that classifies tweets as relevant or irrelevant and then prioritizes them in terms of urgency. The proposed framework was tested and validated using three different classifiers on three different datasets. The classification stage provided an average of 90.67%, 99.25%, and 72.84% recall respectively, on the test cases. Furthermore, the prioritization stage provided a ranked list of urgent tweets which were verified through the presence of TPS 6 terms. The results of the test scenarios showed that the proposed framework is indeed capable of classifying and prioritizing disaster-related tweets for aiding disaster managers in disaster response with significant predictive power.

5.1 Future Work

This paper only focused on typhoon-related tweets and words. It is highly suggested to develop different TPS Lexica for different types of disasters (e.g., earthquake, volcano eruption). By doing so, the proposed framework can be extended to include other types of disaster events. Another future scope can be the development of a user-friendly application that incorporates the proposed framework ready for use by disaster managers and responders. Lastly, a component that caters the lack of location tags of tweets can be explored and added to the proposed model. The authors initially thought of considering locations contained in the actual tweets to be used as estimates for the locations. However, this still needs even further research on location estimation methods.

NOMENCLATURE

Symbol	**Description**
AIDR	Artificial Intelligence for Disaster Response
DRRM	Disaster risk reduction and management
DSS	Decision support system
DTM	Document-term-matrix
EMTerms	Terminological Resource for Crisis Tweets
GIS	Geographic Information System
IMT	Incident Management Teams
LGU	Local government unit
MaxEnt	Maximum entropy
MCDA	Multi-criteria decision analysis
NB	Naïve Bayes
NDRRMC	National Disaster Risk Reduction and Management Council
PAGASA	National Meteorological and Hydrological Services Agency
RF	Random Forests
SVM	Support vector machines
TPS	Tweet prioritization score

REFERENCES

[1] United Nations University. 2017. World risk report 2017. Alliance Development Works. Berlin.

[2] United Nations University. 2017. World risk report 2016. Alliance Development Works. Berlin. 2016.

[3] United Nations University. 2015. World risk report 2015. Alliance Development Works. Berlin.

[4] Mucke P, Kirch L, Walter J, Heintze HJ, Kuppers B, Mann H, Mischo F, Pazdzierny T, Radtke K, Strube F, Weller D. 2018. World risk report 2018. Bündnis Entwicklung Hilft and Ruhr University Bochum – Institute for International Law of Peace and Armed Conflict (IFHV).

[5] Mucke P, Kirch L, Walter J, Forster T, Himmelsbach J, Korte L, Radtke K, Thielborger P, Weller D. 2019. World risk report 2019. Bündnis Entwicklung Hilft and Ruhr University Bochum – Institute for International Law of Peace and Armed Conflict (IFHV).

[6] Prevention Web. 2014. Retrieved from https://www.preventionweb.net/publication/country-risk-climate-profile-philippines on 1 Dec 2021.

[7] The World Bank Group and the Asian Development Bank. 2021. Climate risk country profile: Philippines. World Bank Group: Washington.

[8] Relief Web. 2010. Philippine disaster risk reduction and management act of 2010. Retrieved from https://reliefweb.int/report/philippines/philippine-disaster-risk-reduction-and-management-act-2010 on 1 Dec 2021.

[9] Rappler (S Tomacruz). 2018. Explainer: Who's supposed to be in-charge during disasters?. Retrieved from https://www.rappler.com/newsbreak/iq/209482-explanation-philippine-government-disaster-emergency-response/ on 1 Dec 2021.

[10] ABS-CBN News (A Nieva-Nishimori). 2018. A work in progress: Delay sa text alerts inamin ng NDRRMC. Retrieved from https://news.abs-cbn.com/news/07/20/18/a-work-in-progress-delay-sa-text-alerts-inamin-ng-ndrrmc on 1 Dec 2021.

[11] UNTV. 2017. Retrieved from https://www.youtube.com/watch?v=lhR_bQ77yEk on 1 Dec 2021.

[12] Government Technology. 2013. Retrieved from http://www.govtech.com/em/disaster/Sandy-Social-Media-Use-in-Disasters.html.

[13] Becker H, Naaman M, Gravano L. 2011. Beyond trending topics: Real-world event identification on Twitter. Association for the Advancement of Artificial. p. 1-4.

[14] David CC, Ong JC, Legara EFT. 2016. Tweeting supertyphoon Haiyan: Evolving functions of Twitter during and after a disaster event. PLOS ONE. p. 1-19.
[15] Murzintcev N, Cheng C. 2017. Disaster hashtags in social media, International Journal of Geo-Information. p. 1-17.
[16] Chouhan SS, Khatri R. 2016. Data mining based technique for natural event prediction and disaster management. International Journal of Computer Applications. p. 34-39.
[17] Science Daily (Pennsylvania State University). 2017. Retrieved from https://www.sciencedaily.com/releases/2017/08/170818123540.htm
[18] Data Reportal. 2021. Retrieved from https://datareportal.com/reports/digital-2021-philippines on 1 Dec 2021.
[19] Digital Information World (Hootsuite and We Are Social). 2018. Retrieved from https://www.digitalinformationworld.com/2018/07/global-internet-stats-infographic.html on 1 Dec 2021.
[20] Inquirer (PG Montecillo). 2012. Retrieved from https://technology.inquirer.net/15189/philippines-has-9-5m-twitter-users-ranks-10th on 1 Dec 2021.
[21] Inquirer (MR Camus). 2018. Retrieved from https://business.inquirer.net/247267/healthier-twitter-looks-philippines-boost on 1 Dec 2021.
[22] Marketing Internactive (H Bulos). 2018. Retrieved from https://www.marketing-interactive.com/the-philippines-is-driving-twitter%E2%80%99s-growth-in-apac,-says-the-social-media%E2%80%99s-md-for-sea on 1 Dec 2021.
[23] Imran M, Castillo C, Lucas J, Meier P, Vieweg S. 2014. AIDR: Artificial intelligence for disaster response. International World Wide Web Conference 2014. Seoul, Korea.
[24] Brownlee J. 2017. Difference between classification and regression in machine learning. Retrieved from https://machinelearningmastery.com/classification-versus-regression-in-machine-learning/ on 1 Dec 2021.
[25] Houston JB, Hawthorne J, Perrreault MF, Park EH, Hode MG, Halliwell MR, Turner McGowen SE, Davis R, Vaid S, McElderry JA, Griffith SA. 2014. Social media and disasters: a functional framework for social media use in disaster planning, response, and research. Disasters. p. 1-22.
[26] Vieweg S, Hughes AL, Starbird K, Palen L. 2010. Microblogging during two natural hazards events: What Twitter may contribute to situational awareness. CHI: Conference on Human-Computer Interaction. p. 1-10.
[27] Verma S, Vieweg S, Corvey WJ, Palen L, Martin JH, Palmer M, Schram A, Anderson KM. 2011. Natural language processing to the rescue? Extracting situational awareness tweets during mass emergency, Association for the Advancement of Artificial Intelligence. p. 385-392.
[28] Mendoza M, Poblete B, Castillo C. 2010. Twitter under crisis: can we trust what we RT?. 1st Workshop on Social Media Analytics. p. 1-9.
[29] Terpstra T, Stronkman R, de Vries S, Paradies G. 2012. Towards a realtime Twitter analysis during crises for operational crisis management. Proceedings of the 9th International ISCRAM Conference: Vancouver, Canada. p. 1-9.
[30] Sakaki T, Okazaki M, Matsuo Y. 2010. Earthquake shakes Twitter users: Real-time event detection by social sensors. WWW2010. Raleigh, North Carolina. p. 851-860.
[31] Cobo A, Parra D, Navon J. 2015. Identifying relevant messages in a Twitter-based citizen channel for natural disaster situations. International World Wide Web Conference Committee; Florence, Italy. p. 1189-1194.
[32] Stowe K, Paul M, Palmer M, Palen L, Anderson K. 2016. Identifying and categorizing disaster-related tweets. Proceedings of the Fourth International Workshop on Natural Language Processing for Social Media; Austin, Texas. p. 1-6.
[33] Agarwal A, Xie B, Vovsha I, Rambow O, Passoneau R. 2011. Sentiment analysis of Twitter data. Proceedings of the Workshop on Languages in Social Media. p. 30-38.
[34] Lapitan RG, Batista-Navarro R, Albacea EA. 2016. Crowdsourcing-based annotation of emotions in Filipino and English tweets. Proceedings of the 6th Workshop on South and Southeast Asian Natural Language Processing; Osaka, Japan. p. 74-82.
[35] Nyimbili PH, Erden T. 2018. Spatial decision support systems (SDSS) and software applications for earthquake disaster management with special reference to Turkey. Natural Hazards. 90:1485-1507.
[36] Levy JK, Gopalakrishnan C, Lin Z. 2007. Advances in decision support systems for flood disaster management: Challenges and opportunities. Water Resources Development. 21(4): 593-612.

[37] Shrestha MS, Grabs WE, Khadgi VR. 2015. Establishment of a regional flood information system in the Hindu Kush Himalayas: Challenges and opportunities. International Journal of Water Resources Development. 31(2):238-252.

[38] Zhang Z, Hu H,Yin D, Kashem S, Li R, Cai H, Perkins D, Wang S. 2019. A cyberGIS-enabled multi-criteria spatial decision support system: A case study on flood emergnecy management. International Journal of Digital Earth. 12(11):1364-1381.

[39] Rolland E, Patterson RA, Ward K, Dodin B. 2010. Decision support for disaster management. Operations Management Research. 3:68-79.

[40] Chaudhuri N, Bose I. 2020. Exploring the role of deep neural networks for post-disaster decision support. Decision Support Systems. 130:1-10.

[41] Cavdur F, Sebatli-Saglam A, Kose-Kucuk M. 2020. A spreadsheet-based decision support tool for temporary-disaster-response facilities allocation. Safety Science. 124:1-14.

[42] Liu B, Sheu JB, Zhao X, Chen Y, Zhang W. 2020. Decision making on post-disaster rescue routing problems from rescue efficiency perspective. European Journal of Operational Research. 286:321-335.

[43] Abel F, Hauff C, Houben GJ, Tao K, Stronkman R. 2012. Twitcident: Fighting fire with information from social web streams. International World Wide Web Conference Committee; Lyon, France.

[44] Deng Q, Liu Y, Zhang H, Deng X, Ma Y. 2016. A new crowdsourcing model to assess disaster using microblog data in typhoon Haiyan. Natural Hazards. 84:1241-1256.

[45] Wu D, Cui Y. 2018. Disaster early warning and damage assessment analysis using social media data and geo-location information. Decision Support Systems. 111:48-59.

[46] Bai H, Yu G. 2016. A Weibo-based approach to disaster informatics: incidents monitor in post-disaster situation via Weibo text negative sentiment analysis. Natural Hazards. 83:1177-1196.

[47] Chae J, Thom D, Jang Y, Kim S, Ertl T, Ebert DS. 2014. Public behavior response analysis in disaster events utilizing visual analytics of microblog data. Computers & Graphics. 38:51-60.

[48] Avvenuti M, Cresci S, Del Vigna F, Tesconi M. 2016. Impromptu crisis mapping to prioritize emergency response. Computer. 16:28-37.

[49] Strohmaier M. 2010. Retrieved from https://mstrohm.wordpress.com/2010/01/15/measuring-earthquakes-on-twitter-the-twicalli-scale/ on 1 Dec 2021.

[50] Poblete B. 2017. Twicalli: An earthquake detection system based on citizen sensors used for emergency response in Chile. The 40th International ACM SIGIR Conference: Shinjuku, Tokyo, Japan.

[51] Kebabci K, Karsligil E. 2015. High priority tweet detection and summarization in natural disasters. IEEE. p. 1-4.

[52] Takahashi B, Carmichael C, Tandoc E. 2015. Communicating on Twitter during a disaster: An snalysis of tweets during Typhoon Haiyan in the Philippines. Computers in Human Behavior. p. 392-398.

[53] Yildirim S. 2020. Retrieved from https://towardsdatascience.com/naive-bayes-classifier-explained-50f9723571ed on 1 Dec 2021.

[54] QuantStart Team. 2018. Retrieved from https://www.quantstart.com/articles/Support-Vector-Machines-A-Guide-for-Beginners on 1 Dec 2021.

[55] Smartsheet. 2018. Retrieved from https://www.smartsheet.com/all-risk-assessment-matrix-templates-you-need?fbclid=IwAR1hrYQgSCtqGq9wFsC72ecIAVZgm2H9vEvqPB4JexCMoI0fODLrri00ai8 on 1 Dec 2021.

[56] Shaw F, Burgess J, Crawford K, Bruns A. 2013. Sharing news, making sense, saying thanks: Patterns of talk on Twitter during the Queensland floods. Australian Journal of Communication. p. 23-40.

[57] Pekar V, Binner J, Najafi H, Hale C. 2016. Detecting reports of mass emergency events on Twitter. International Conference Artificial Intelligence; Athens. p. 137-143.