Utilizing Twitter Data for Identification of Need and Availability in Disaster Relief

KAVITA ASNANI[1]
NIKITA SALKAR[2]
SHRAVAN MANERIKAR[3] MUGDHA KHATAVKAR[4] DELSTAN D'SOUZA[5]

[1]State Higher Education Council, Directorate of Higher Education, Porvorim, Goa, India
[1]Adjunct Faculty, Goa College of Arts, Science and Commerce,Sanquelim, Goa, India. (kavitadhegoa)@gmail.com
[2]Khoury College of Computer Sciences, Northeastern University, Boston, United States. salkar.n@northeastern.edu
[3]Visteon, Goa, India. shravan.manerikar@visteon.com
[4]University of Queensland, Australia. khatavkar.mugdha1@gmail.com
[5]Dehlivery Private Limited, Goa, India.delstan2007@gmail.com

Abstract. Social media in today's world has substantial impact in several domains. Disasters caused by natural hazards witness active communication on social media platforms. In such situations, the popular microblogging forums like Twitter have been resourceful in providing information related to relief operations. Tweets about disaster related to need and availability of emergency resources have been referred to as need-tweets and availability-tweets respectively. Automatic identification of such need-tweets and availability-tweets could be of good aid for timely action in relief operations. Our research exploits posts from Twitter to investigate the feasibility of using machine learning techniques of clustering and deep learning to assist in identification of need and availability of resources during crisis.

We performed experiments using Nepal Earthquake dataset. The results obtained by clustering algorithms yielded clusters that were specious with the occurrence of noise. We also performed experiments using recurrent neural network with long short term memory and compared the results with two baseline techniques using need and availability tweets. The proposed recurrent neural network achieved the precision of 0.772 and F-score of 0.600 for both need as well as availability tweets, which were higher than the baseline techniques.

Keywords: Social media, Twitter, Disaster Relief, Earthquake, Need, Availability, Clustering, Deep learning

(Received July 2nd, 2021 / Accepted November 19th, 2021)

1 Introduction

Disasters are the major destructive factor on the earth. In integration to the natural disasters, un-natural developments by the human species have contributed to man-made and technological disasters. The classification of disasters based on the cause is depicted in Figure 1

In any given disaster situation, several rescue and recovery operations are carried out in order to reduce catastrophic loss of life, physical and property eradications and mass trauma in afore, during and after stages of the disaster. In such circumstances, expeditious communication and necessary action always reduces the losses. In recent times, lots of efforts have been taken for utilizing the technological development in every aspect of human life. In specific, utilization of popular social media forums to get the information in bulk amount and in the speedy manner has been of recent interest in research.

When the disaster strikes, the immediate rescue is the provision of resources that are required by the people

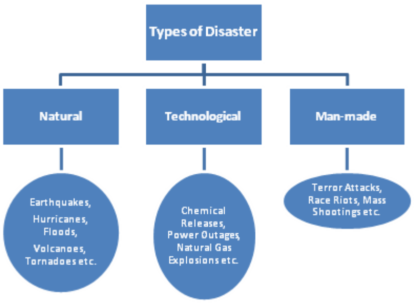

Figure 1: Types of Disaster

stuck in the disaster. Resources such as medical facilities, basic necessities namely food, water, shelter and infrastructural resources are imperative for survival. In such situations, there is a need for the resources to be made available to the victims as early as possible. Consequently, there is control on the number of human casualties and serious after-effects and the affected regions eventually resume to stability. However, the faster dispatch of required resources to the victims is possible only after identifying the resources needed and the resources that are then available for help.

In the 21st century, social media has touched every aspect of human life. Information about disaster and during disaster is promulgated across the globe via social media. Social media is widespread and the vast data from social media forums can be exploited to help expedite the relief efforts when the disaster strikes. News about casualties, the scale of devastation, immediate requirements etc. are shared by the social media users, which enables the government, organizations and people willing to help, gain cognizance of the situation and pool in resources and donations to offer help to the disaster-hit victims.

Social media sites such as twitter witness a deluge of tweets and has become the active reservoir of information [9, 15] when a calamity strikes. During the Kerala floods in 2018, more than 2.62 million tweets were shared [1]. Hurricane Sandy that made landfall on 29th October, 2012 saw more than 20 million tweets within the first week itself. According to the Pew Research Center, photos and videos contributed to more than half the twitter conversations [2]. During the Great East Japan Earthquake in 2011, around 55 million tweets were collected pertaining to the disaster. The information pertinent to the disaster contains useful facts that can be used to accelerate the process of supplying resources. Such information is classified in two groups, namely, need and availability. Need messages indicate the requirement of resources with more information like the need for blood and type of causalities [11]. Availability messages indicate the availability of resources, giving information about people, volunteers, and organizations willing to help with provision of required resources [11]. The tweets related to need are referred to as need-tweets and the tweets related to availability are referred to as availability tweets. During the disaster, we can expedite the efforts of sending out the required resources to the locations that need the respective resources by matching the identified need-tweets with the availability-tweets. In such situations, automatic identification of need and availability could be of great aid. Our proposed work aims at using machine learning techniques for identification and automatic classification of the need and availability from social media data gathered during crisis.

2 Related Work

Tweets are typically expressed in the text form and have a restriction on the word limit and occurrence of slang makes it difficult for the retrieval systems to fetch relevant information. They are mostly informal and may have the use of multiple languages in the same tweet. The text preprocessing operations on tweets include the tasks related to cleanup and normalization which includes removal of non-ASCII characters, ASCII transliterations of Unicode text [14], tokenization [14], stemming [14],[20], normalizing the words into lowercase [14], translation of words using Google Translate [16][20], pruning words with hashtag [5][12], removing URLs [8][20][12], removing duplicate tweets and retweets [5][8], removal of punctuation [12] and stopword removal of words occurring in multiple languages [5][14][20][12].

Prior works include pattern matching methods [17][3] and also word embedding based retrieval techniques [4] to capture the semantics of need and availability tweets to retrieve the tweets. For this purpose a dataset of 50068 tweets pertaining to the Nepal earthquake was used. The study revealed that pattern matching methods could not identify the required need and availability tweets because the tweets seldom contained intuitively complementary terms such as ‘need‘, ‘availability‘ or ‘distribute‘. On the other hand, the contextual

[1] https://www.bgr.in/news/kerala-floods-twitter-2-62-million-tweets-during-august-2018-deluge-686121/

[2] http://www.pewresearch.org/fact-tank/2013/10/28/twitter-served-as-a-lifeline-of-information-during-hurricane-sandy

word2vec based methods successfully identified need tweets and availability tweets. Thus, the performance of word2vec based methods is much superior compared to pattern matching method establishing the efficiency of contextual matching. The accuracy was found to be 18% and 45% for need and availability respectively [4]. Classification methods have been used for need and availability classification. The existing research shows the classification of data using two main classes of learning algorithms namely offline and online learning algorithms. The offline classification has been done by using semantic word-embedding by Logistic Regression(LR) [20]. This method consisted of three phases a) Preprocessing of tweets b) Adding semantic knowledge using neural network based on tf*idf scores c) Training of the classifier using the logistic regression. This approach resulted in 74% accuracy enabling the feature selection [20]. Also, the offline classification using Support Vector Machine(SVM) have been tried for classification and further for performing matching of the need and availability tweets [18]. The use of three classifiers was made to classify the data and then the best out of them were selected depending on the results and the feature selection. The classifiers used were AdaBOOST, SVM of Linear Kernel and SVM of non-linear kernel. The classified data was further used to match the need with availability [18]. It was observed that SVM performed the best for classification task based on the average MAP value; whereas AdaBOOST worked better for matching task [18]. A hybrid approach with Logistic Regression(LR) model along with the Support Vector Machine(LibSVM) was used for information retrieval from the microblogs [12]. The experimental results showed that depending on the MAP values, the LR model could more efficiently classify the tweets as compared to LibSVM. The need and availability matching resulted in almost same results w.r.t precision, recall and F-score. Due to the informal social media platform there was comprise done with sorting of text and also filtering and and feature selection [12].
Online learning algorithms comprises of machine learning in situations where it is necessary for the algorithm to dynamically adapt to the new patterns in data, or when the data itself is generated as a function of time. As the offline methods used aspects like feature selection, text matching and tweet processing; Deep Neural Network (DNN) uses information from posts and type of the information. It has not only been able to benefit from the aspects but is also used to predict about the future events. Convolution Neural Network (CNN) was used under DNN to effectively classify the key features at different levels of abstraction since it is effective at sentence level classification [6]. DNN was trained using the Stochastic Gradient Descent (SGD) and this could fare the classification part with more efficiency [6].
CNN being an online training method, learns in small batches which suits any dynamic situation perfectly unlike the previously used offline methods like SVM, LR, etc. Also auto feature extraction by NN did not require any manual feature engineering [6].

Table 1: Tweet Statistics for Nepal Earthquake Disaster

Tweet Retrieval for Nepal Earthquake Disaster	**Training Set**	**Test Set**
	20000 tweets posted on Twitter during the disaster	40000 tweets posted on Twitter during the disaster

Table 2: Annotated Need and Availability Tweets

Training Tweets 16834 of 20000 tweets retrieved	Training Need Tweets	194 valid tweets out of 211
	Training Availability Tweets	624 valid tweets out of 719
Testing Tweets 40772 of 46000 tweets retrieved	Testing Need Tweets	368 valid tweets out of 427
	Testing Availability Tweets	813 tweets out of 980

3 Experiments Performed

In this section we first present the implementation details essentially presenting the dataset used and its representation, followed by the approaches used to evaluate usefulness of the system.

3.1 Dataset Used

We used the dataset of the 2015 Nepal earthquake [10]; which took place from 25th to 27th April, 2015. On 25th April, 2015, Nepal was struck by a devastating earthquake of magnitude 7.8Mw [3]. Twitter API was used to retrieve the tweets. The tweets were collected and stored as text files. The dataset contained 66000 tweets posted during the course of Nepal Earthquake, April 2015.

The statistics of the initial data retrieved from twitter is presented in Table 1. In order to make the information from the dataset useful, preprocessing was done.

[3]https://en.wikipedia.org/wiki/April-2015-Nepal-earthquake

Table 3: Word Count for Monolingual and Mixed Tweets

	Training	Testing
Monolingual	Tweets: 11001 Words: 218000	Tweets: 34190 Words: 668100
Mixed	Tweets: 9000 Words: 170000	Tweets: 5023 Words: 99200

Table 4: Word Count for Language Specific Monolingual and Mixed Tweets

	English	Hindi	Nepali
Training Availability	12400	693	57
Training Need	3690	82	68
Testing Availability	13000	1106	694
Testing Need	6700	374	236

Preprocessing techniques included cleaning of data to make the tweets written in social media jargons, appropriate for use. We used Natural Language Processing (NLP) techniques using Natural Language ToolKit (NLTK) library to remove hashtags, URLs, usernames, stopwords, punctuations and duplicates. The preprocessed tweets were manually annotated in two classes namely ‘need tweets‘ and ‘availability tweets‘. This was done for preparing data for training. The dataset provided data related to ‘need availability‘ retrieval. The details of the same are provided in Table 2. The 3166 training and 5228 testing tweets could not be retrieved since they were deleted or users that tweeted had deleted their accounts and were no longer twitter users.

Tweets have a stringent word limit, and users often make use of abbreviations which are difficult to interpret for retrieval systems. They are mostly informal and may involve the use of multiple languages in the same tweet, called code mixing, or even multiple scripts in a single tweet. We scanned the initial tweet dataset manually to identify the monolingual and mixed tweets. The details of the same are provided in Table 3.

The tweets in the dataset were multilingual in English, Hindi and Nepalese language. The language specific word count of the tweets occurring in the dataset is given in Table 4.

3.2 Data Representation

In view of further processing, we convert the tweets into vector representation. The word2vec was used to get local word vectors which are expected to capture the local context. The doc2vec model was used for converting the textual tweets into vector representations. The tweets were translated into feature vectors for processing and analysis. A feature vector is a n-dimensional vector. For the ease of processing, the dimensionality of the vectors has been reduced using the t- distributed Stochastic Neighbor Embedding (t-SNE). It reduces multi-dimensional vectors into two dimensional vectors, while preserving the semantic relations between the vectors. The doc2vec model is trained on 44921 tweets and the output is the two dimensional vector of the tweets.

The spatial layout of the two dimensional vectors in the vector space, after implementing the doc2vec model on the dataset, is shown in the Figure 2. The dense regions in Figure 2 indicate that points forming those areas are similar and hence the tweets they represent may be co-related in the meaning.

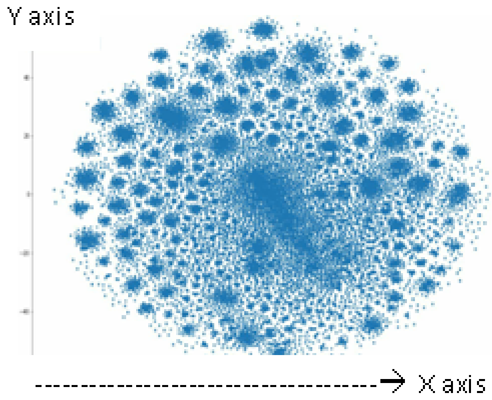

Figure 2: Doc2Vec Representation

3.3 Experimental Approaches and Results

In order to evaluate the quality of need and availability, we performed experiments using clustering and deep learning approach.

3.3.1 Clustering Approach

Clustering is an unsupervised approach which clusters unlabeled data in groups such that each group contains items that are similar to each other while items in different groups are dissimilar. We performed experiments using k-means and the DBSCAN algorithms [1]. The partitioning of tweets was based on the notion of need and availability. We implemented the k-means algorithm with k =2. Instead of dividing the entire dataset of 44,921 vectors into two distinct groups, the k- means

algorithm divided the 44,921 vectors into two mixed groups. Therefore, we concluded that k-means algorithm did not work for segregation of need and availability. The clusters obtained were not discrete due to the varying density in the vector space.
We then performed experiment using Density based Spatial Clustering of Applications with Noise (DBSCAN) algorithm which is a density based clustering algorithm. This algorithm clusters regions of high density. Since the output of the doc2vec model places similar tweets close to each other in the vector space, the DBSCAN algorithm was expected to detect the regions of high density formed by similar clusters. We assigned the epsilon to 0.06 and the minpts to 26. The Figure 3 depicts clusters formed after implementing DBSCAN.

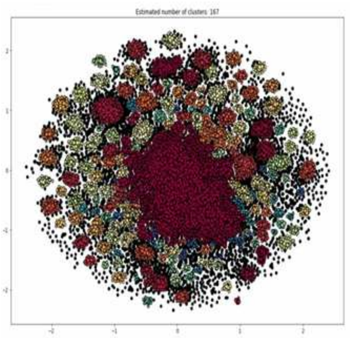

Figure 3: Clusters after implementing DBSCAN

The DBSCAN algorithm groups all the points that form dense regions, which can be seen in the Figure 3. On implementing the DBSCAN algorithm on 44,921 vectors we obtained 13,091 outliers. After removing 13,091 outliers, the remaining 31,830 vectors were clustered by the DBSCAN algorithm subsequently generating 167 groups. We observed that the obtained 167 groups had no specific bonding characteristics of need or availability. Rather they were simply a motley cluster of need tweets, availability tweets and also a large amount of noise with no particular outcome. We could associate this behavior with the characteristic of the DBSCAN clustering algorithm which is generally useful for identifying multiple clusters in vector spaces having arbitrary shapes. It is seen to work optimally when the data is homogenous. However, the dataset we used was found to be heterogeneous in nature. We attribute this pattern in the output to the fact that since it is the social media data, the dataset contained a large amount of noise in the form of condolences messages, rumors and other irrelevant data.

Table 5: Performance of DBSCAN Clustering

	Precision	**Recall**	**F-score**
Need	0.2217	0.0965	0.1781
Availability	0.1977	0.0796	0.1476

We tested with the set of experiments, by the manual annotation process. We found 207 need tweets and 374 availability tweets. The number of need and availability tweets was limited for the clustering algorithm to segregate the total dataset into need tweets and availability tweets. Hence, we conclude that most of the twitter data comprised of data prayers and concerns without specific focus on need and availability. This is reflected from the performance of the clustering techniques implemented with the purpose of segregating the need tweets and availability tweets. It showed the performance which was less than the optimum and yielded results, lower than expected. The performance of the DBSCAN algorithm is shown in Table 5. Low F-score of DBSCAN clustering algorithm indicated the need to explore the deep learning approach for the purpose of identification of need tweets and availability tweets.

3.3.2 Deep Learning Approach

Considering the preclassified tweets for creating the model, we implemented the Recurrent Neural Network (RNN) in order to classify the tweets into the classes need and availability. We performed experiments on Long Short Term Memory (LSTM) which is an artificial RNN architecture used for deep learning [13]. It does not only process single data points, but also entire sequences of data. We trained the dataset using LSTM algorithm having six layers and validated the model using the validation dataset. LSTM deals with the exploding and vanishing gradient problems that can be encountered when training traditional RNNs.

The model was tested for different optimizers:

- Stochastic Gradient Descent (SGD): The SGD algorithm oscillates along the path of steepest descent towards the optimum [7].

- Root Mean Square Propagation (RMSprop): The RMSprop keeps a moving average of the elementwise squares of the parameter gradients. Using RMSProp effectively decreases the learning rates of parameters with large gradients and increases the learning rates of parameters with small gradients [2].

Table 6: Performance of Deep Learning * N.D. : Not Defined

Optimizer	Precision	Recall	F-score
Need -Tweets			
Adam	**0.77234**	**0.34715**	**0.600304**
SGD	0	N.D.	N.D.
RMSprop	0.149856	0.253659	0.30677
Availability-tweets			
Adam	**0.33727**	**0.764881**	**0.568648**
SGD	1	N.D.	N.D.
RMSprop	0.799213	0.674419	0.1588586

- Adaptive Moment Estimation (Adam): The Adam uses a parameter update that is similar to RMSprop, but with an added momentum term. It keeps an element-wise moving average of both the parameter gradients and their squared values. The full Adam update also includes a mechanism to correct a bias the appears in the beginning of training [19].

The performance of the optimizers for need and availability tweets is presented in Table 6. The highest precision and F-score was obtained for the Adam optimizer, for both need and availability tweets. The results of the SGD optimizer shows N.D. values, which is indicative of fact that the training data contained more of availability tweets as compared to need tweets.

From the performance shown in the Table 6, we conclude that RNN model is capable of identifying need and availability tweets. The better availability of both need and availability tweets would aid in desirable classification of the tweets. Also, the precision in the Table 5and the Table 6 indicate that the neural network models perform relatively better than clustering for pattern finding from twitter data.

4 Conclusion

Our work proposes to exploit the social web of twitter to assist in offering rescue to the affected in the disaster event. We sought to use the results of machine learning techniques to coordinate the rescue and relief efforts towards provision of resources to the people adversely affected in the disaster stricken areas.

Our work entails the processing of the dataset of tweets pertaining to a 2015 Nepal earthquake disaster event obtained using the twitter API, and then identification of need-tweets and availability tweets. The task was carried out by experiments using clustering techniques which included the k-means algorithm and the DBSCAN algorithm. The task of identification was also done using deep learning techniques. RNN was used to classify the tweets to serve the output of need-availability identification model. From the results of the experiments conducted, we conclude that the tweets collected during the disaster event showed the presence of need-availability pattern in the deep learning approach. However, the skewness was due to the noise from rumors, prayers and other irrelevant data not specific to the need/availability requirement. It is also concluded that the sufficient and relevant training data, would increase the accuracy of the model drastically.

As the part of the future work, we intend to map the need tweets to that of availability tweets, so as to find the exact resource for the victims. For the given disaster like situation-based domain we shall employ the identification of the need and availability module. Further, for the given need tweet, we will map availability tweets to find the most matching one. Such a service to be offered in real time mode during the disaster will be a fast mode of providing the essential help to the needy people stuck in the disaster from those offering the help.

References

[1] Aggarwal, C. C. An introduction to cluster analysis. In *Data Clustering*, pages 1–28. Chapman and Hall/CRC, 2018.

[2] Alam, M., Abid, F., Guangpei, C., and Yunrong, L. Social media sentiment analysis through parallel dilated convolutional neural network for smart city applications. *Computer Communications*, 154:129–137, 2020.

[3] Basu, M., Ghosh, K., Das, S., Dey, R., Bandyopadhyay, S., and Ghosh, S. Identifying post-disaster resource needs and availabilities from microblogs. In *2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 427–430. IEEE, 2017.

[4] Basu, M., Ghosh, K., and Ghosh, S. Information retrieval from microblogs during disasters: In the light of irmidis task. *SN Computer Science*, 1(1):1–10, 2020.

[5] Baweja, S., Aggarwal, A., Goyal, V., and Mehta, S. Automatic retrieval of actionable information from disaster-related microblogs. In *FIRE (Working Notes)*, pages 43–45, 2017.

[6] Cleuziou, G., Martin, L., and Vrain, C. Poboc: an overlapping clustering algorithm. *Application to rule-based classification and textual data*, 440:444, 2004.

[7] Diab, S. Optimizing stochastic gradient descent in text classification based on fine-tuning hyper-parameters approach. a case study on automatic classification of global terrorist attacks. *arXiv preprint arXiv:1902.06542*, 2019.

[8] Gautam, B. and Basava, A. Automatic identification and ranking of emergency aids in social media macro community. *arXiv preprint arXiv:1810.11498*, 2018.

[9] Guimarães, R., Rodríguez, D. Z., Rosa, R. L., and Bressan, G. Recommendation system using sentiment analysis considering the polarity of the adverb. In *2016 IEEE International Symposium on Consumer Electronics (ISCE)*, pages 71–72. IEEE, 2016.

[10] Khosla, P., Basu, M., Ghosh, K., and Ghosh, S. Microblog retrieval for post-disaster relief: Applying and comparing neural ir models. *arXiv preprint arXiv:1707.06112*, 2017.

[11] Kumar, N. and Dubey, M. Databros@ information retrieval from microblogs during disasters (irmidis). In *FIRE (Working Notes)*, pages 41–42, 2017.

[12] Mehrotra, H., Soni, R., and Pal, S. Iit bhu at fire 2017 irmidis track-fully automatic approaches to information retrieval. In *FIRE (Working Notes)*, pages 59–60, 2017.

[13] Nistor, S. C., Moca, M., Moldovan, D., Oprean, D. B., and Nistor, R. L. Building a twitter sentiment analysis system with recurrent neural networks. *Sensors*, 21(7):2266, 2021.

[14] Raf, M., Ahmed, S., Ahmed, F., and Ahmed, F. Tweet classification using semantic word-embedding with logistic regression. *Technometrics*, 49(3):291–304, 2007.

[15] Rosa, R. L., Rodriguez, D. Z., and Bressan, G. Sentimeter-br: A social web analysis tool to discover consumers' sentiment. In *2013 IEEE 14th international conference on mobile data management*, volume 2, pages 122–124. IEEE, 2013.

[16] Talukdar, A., Bhargava, R., and Sharma, Y. Bits_pilani@ imridis-fire 2017: Information retrieval from microblog during disasters. In *FIRE (Working Notes)*, pages 38–40, 2017.

[17] Temnikova, I. P., Castillo, C., and Vieweg, S. Emterms 1.0: A terminological resource for crisis tweets. In *ISCRAM*, 2015.

[18] Xin, D., Xiaoyu, W., and Ziyao, Z. A hybrid model for information retrieval from microblogs during disaster.

[19] Zhang, J., Hu, F., Li, L., Xu, X., Yang, Z., and Chen, Y. An adaptive mechanism to achieve learning rate dynamically. *Neural Computing and Applications*, 31(10):6685–6698, 2019.

[20] Zicheng, Z., Hui, N., Ziyao, Z., Jinmei, Z., and Jun, L. Hljit2017-irmidis@ irmidis-fire2017: Information retrieval from microblogs during disasters.