International Journal on Emerging Technologies 10(4): 345-353(2019)

ISSN No. (Print): 0975-8364
ISSN No. (Online): 2249-3255

Assorted Sentiment Analysis Model for Natural Crisis Response and Recovery using Big Data driven Technology

S. Dhyani[1], G.S. Thakur[2] and Y. Sahu[1]
[1]Research Scholar, Department of Computer Application, MANIT Bhopal (Madhya Pradesh), India.
[2]Assistant Professor, Department of Computer Application, MANIT Bhopal (Madhya Pradesh), India.

(Corresponding author: S. Dhyani)
(Received 09 September 2019, Revised 05 November 2019, Accepted 16 November 2019)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Social networking sites are generating big data from which we can fetch valuable data. Social media is increasingly used for communication during emergency situation caused by natural crisis and also used for helping related requests. During natural crisis situation, pool of big data is used to extract emergency request for moving forward to provide timely help. Though emergency responders and government agencies work together with the help of their respective framework of natural crisis response and recovery mechanism, the sentiment of the affected persons during natural crisis and after that determines the success of natural crisis response and recovery mechanism. In this research paper, we have proposed natural crisis response and recovery through analysis of assorted sentiment model with the help of big data driven approach. Assorted sentiment analysis model is a combination of various phases for finding sentiment from any textual content of social media rather than to find sentiment in one phases from textual content of social media. Assorted sentiment model would extract sentiment from covering many phases from textual content of social media. These phases are namely collection phase, preprocessing phase, and filtering phase. The proposed model gathers natural crisis data from Twitter and categorizes them according to the requirement of needy persons. The categorized natural crisis dataset is classified through SVM (Support Vector Machine) algorithm for finding analysis of sentiment of affected people. We have chosen SVM algorithm for classification because SVM performs better than other machine learning algorithm after applying feature generation method. Various feature like, Bag of Word, part of speech and lexicon are analyzed to identify best classification strategy for natural crisis data. Result shows that Bag of Word feature combined with SVM is suitable for analyzing the needs of people during natural crisis. This model helps rescue team and emergency responders to improve good approaches for making effective MIS (management information system) of frequently changing natural crisis situation.

Keywords: Big Data, Natural crisis Management, Natural language Processing, Sentiment analysis, Social media analysis, Text mining.

I. INTRODUCTION

Twitter has generated Big Data which has made significant position in the industries all over the world. Various applications are involved in the big data analytics just like traffic control, sports management and telecommunications industries manufacturing industries, crime analysis and prediction, smarter healthcare [1]. Now-a-days people are using social media for promoting critiques and evaluation about services related to products. Social networking sites generate huge volume of data every day for a specific query, analysis and filtration for those big data is a major challenge for industries analytics of big data is being used in several sectors but in their application in the natural crisis data recovery is still at in beginning phase [2]. Social networking sites are rarely used for helping emergency related query in the natural crisis prone area. As crisis conditions during natural crisis are more disorganized and chaotic, the big data which are generated through social networking site analysis of that big data is the good suitable to capable of handling the jumbled and the fully unorganized situation which is chaotic [3]. It is significant to generate good decision to the required time for helping needy persons with their requirements during the period of natural crisis. Direct communication between affected people and natural crisis management is the need of the hour. Most of the times due to shortage of direct communication between them lead natural crisis management relies on incorrect or incomplete message. Computational intelligence and analysis of big data play a vital role in such a situation. Big data analytics can help rescue team using computational intelligence to obtain the right information after analyzing big data and take after that best course of action. There are three phases of natural crisis management [4]. First is an early warning and preparedness, second is the impact and risk and third is a risk which depends on the modeling and susceptibility. In all the stages two types of input data are generated. First is data which is generated by sensors such as drones, an image related to satellite and the second is content which is generated by user of social media such as Flicker, Facebook and Twitter. The Natural crisis situation can be handled efficiently when these data are analyzed in an efficient way. Big data analytics provides solution to grip big data in such a technique that all the three stages of a natural crisis situation are managed efficiently.

Though natural crises are huge, devastating, and chaotic they carry people under one roof where persons help one another and contest against the devastation. It is the tendency of persons to express the emotions and opinion surrounding them. Sentiment analysis of big data has been used for several years in the types of opinionated material such as news articles or reviews of online content. However Social Networking material has been imposed many unique challenges to the processing of natural language in particularly and to find Sentiment analysis in effective way [5]. It is important to understand the emotionally text of message and analyze it to extract real meaning of content. Such kind of analysis was performed on the content belonging to the Haiti Earthquake [6]. It was the first incident where affected people were survived through sentiment analysis with the help of big data technologies. Crowd Sourcing technique was used to produce crisis maps during Haiti earthquake [7]. Digital technology related to crisis response has come into existence after the incident of Haiti earthquake. Through various studies have been done related to analyze opinion of the people during disaster [8]. They are not much more effective to find sentiment according the need of affected people during any devastation. To acquire the requirements of development, people are irrationally utilizing and exploiting resources which are natural and exacerbating the happening of natural crisis such as floods, earthquake and hurricane by changing the climate directly and indirectly [9]. In real life, there are many applications which are using social media for finding sentiment analysis. These application include brand monitoring using social media monitoring, Market Research, product analytics, voice of customer analysis, feedback analysis, political election analysis, natural crisis management, etc. Among these applications we have focused on the natural crisis management related application in our research. In this research Paper we have implemented an assorted sentiment analysis model to recognize the sentiment towards help received by persons during and after a natural crisis. Though rescue team and government try to aid the people during natural crisis. In this research paper we have classified the tweets during natural crisis and helps in constructing assorted sentiment analysis techniques on several requirements of the persons. The proposed sentiment model will provide help to rescue team to understand the natural crisis situation and work according to that situation. Firstly, we have categorized natural crisis data and after that we have analyzed the various requirements of the citizens throughout natural crisis situation. Secondly, different kinds of feature like, POS (part of speech) tagging based feature, 2-gram (Bigram) and 3-gram (Trigram) based feature and various kinds of lexicon based features are analyzed. Each category of needs is identified and after that best performing machine learning algorithm is applied for extracting sentiment analysis. Advantage of proposed model is that it would help rescue team and emergency responders to develop better strategies for effective information management system of the rapidly changing natural crisis environment.

We have organized paper in VI sections. Section II describes the literature review related to tasks in the area of sentiment analysis. The proposed model is presented in section III. Result of experimental setup is presented in section IV. Section V provides the conclusion of the research work. Main objective of our research is to categorize natural crisis related tweets from twitter according the need of the people during natural crisis period and after that find the sentiment of affected people on the categorized data so that rescue team can provide timely help according to the need of affected people.

II. LITERATURE REVIEW

Social media can be used extensively during devastation situation but for monitoring crisis situation social media relatively using a smaller number of tools [3]. Impact of ICT on the 2011, flooding of Thailand was studied [10]. Learning was absorbed on three dimensions, namely structural, psychological way and empowerment of resources these dimensions empowered the community after analyzing the way through which dimensions are extracted from social media. It also has studied that how to improve communication channel during crisis response. During Tohoku earthquake the role of social media was studied [11]. In their work, two groups were formed for twitter users. The first group was formed from users who are not affected by the disaster and a second group was formed from users who are affected by disaster. The effect of manipulators affected by earthquake was analyzed. The main goal of sentiment analysis is to deal with classifying text into positive and negative. First work related to sentiment analysis was to classify review data which is belonging to movie using machine learning approaches into positive and negative [12]. Study was carried out analysis of challenges related to sentiment. SVM (Support Vector Machine) algorithm was implemented to reduce the dimensionality [13]. Sentiment analysis carried out on reviews of product over the data fetched from site of Amazon [14]. The experiment setup fetches result from both levels. First level is depending on the review and the second level is depending on the sentence. For performing the classification of sentence these two levels are applied. In their experiment social networking mining technique was applied in sentiment analysis and modeling of topic for finding changing needs of the customer. Sentiment analysis was performed on various regions like hotel reviews [15] and pizza industry [16] for analyzing satisfaction of the customer. Chinese stock market was predicted through twitter messages [17]. Authors worked for both machine learning technique and technique based on lexicon for dataset from financial website of sina. Authors revealed that semantic technique had accuracy of classification less than machine learning technique. Support Vector Machine (SVM) algorithm was implemented to categorize the tweets pertaining to influenza disease [18]. The experiment concluded that classification accuracy was significantly improved using Natural Language Processing (NLP). Data extracted from the web is used to improve the services related to airport [19]. Sentiment analysis is mainly used to find out the hidden meaning of subjective expression from the textual data in the research. Various features of social media content are analyzed by different researchers in order to fetch the subjective content from social media content.

Adjective Verb Adverb (AVA) framework used to classify subjective sentences [20]. It calculates sentiment of any type of documents through the use of adverbs and adjectives. Mapping of the sentiment from English language to Dutch language multilingual sentiment analysis based on lexical approach was performed [21]. Results revealed that the language specific dimension associated accordingly with the meaning of the word for finding sentiment of the text. Sentiment analysis of Urdu blog framing for several areas is done by two widely used approaches, namely support vector machine approach and lexicon-based approach [23]. They compared two approaches with accuracy, precision, recall, f-measure and found that lexical based approach outperformed. In current days, researchers of areas of opinion mining have engrossed on applying sentiment analysis technique in social media content during a disaster. Features like number of words and keywords were used to detect event which was targeted [22]. To estimate the center of trajectory particle filtering and Kaman filtering were used along with known feature. During the Japan earthquake a mixture of emotions like sadness, anxiety, fear, relief, calm and unpleasant was studied [24]. A model of sentiment analysis was built which routinely fetched tweets related to natural crisis and classified them into various kinds of categories like impersonal or personal style, informal or formal linguistic text, the subjectivity of linguistic text. Machine learning algorithm was used to classify tweets related to crisis [25]. In Linguistic pattern blended with textual content with the help combining various semantic rules based on the syntactic structure of the sentences authors use linguistic sentiment flow algorithm for finding result [34].

Dataset Description: We have used twitter to collect disaster related data. Disaster related data that are considered in this research include Kedarnath flood another flood named Chennai and Fani Cyclone. For text analysis we have collected 23,500 tweets using Streaming of Twitter API which allows extracting only past seven-day data, rest of disaster related data was extracted from Historical data recovery tool for twitter 2017.

Table 1: **Description of Dataset.**

Disaster	No of Tweets	Type	From	To
Kedarnath	10,000	Flood	17/06/2013	25/12/2013
Chennai	6,000	Flood	08/11/2015	04/01/2016
Fani	7500	Cyclone	26/04/2019	01/06/2019

III. PROPOSED MODEL

Fig. 1 shows the broad view of the proposed model. Model work for 3 stages namely, data collection stage, data filtering stage and data classification stage. In data collection stage, we have extracted natural crisis dataset from twitter using API provided by the twitter. After extracting dataset preprocessing step has to be taken. Next preprocessed data would be used for the filtering stage. In filtering stage, we have filtered category based on the keyword. We have considered keyword which comes in the dataset more times. Mainly keywords are the words which are used mostly by user of twitter during any kinds of natural crisis situation in the last stage we have implemented support vector machine algorithm on the categorized data for finding sentiment related to each category during natural crisis.

Data Preprocessing: Natural crisis dataset is preprocessed in such a way that SVM algorithm can understand data in the next phase. Data Preprocessing is used to discard unnecessarily contents from the input data. In case of Twitter message, we have removed numbers, URL, language of foreign words, symbols and emoticons, abbreviations, Hash tags. Natural crisis data is characterized according to several requirements of persons after data preprocessing is performed

Data Categorization: Keyword filtering technique was used to categorize natural crisis related data [26]. For this technique data was disaster related tweets of twitter. Keywords are selected for each category based on words that are found more than five times [27]. Categories contain water, medical emergency, shelter, transportation, electricity, and food is most demanding needs of individual affected during disaster. Table 2 shows Keywords related to each class of identified needs. Keywords are used to fetched natural crisis dataset which contains 6832 tweets Fig. 2 shows distribution of tweets under various categories.

Subjective Sentence Categorization: Distributing the text into objective and subjective sentence accordingly to usage of words, is major work which involved in sentiment analysis is to filter Objective and Subjective sentence from tweet. Objective kinds of sentence do not contribute to recognize sentiment of the persons affected by natural crisis while subjective kind of sentence do contribute to identify the sentiment of people who are affected by natural crisis. Table 3 shows a sample of categorization of sentences which are objective and subjective Fig. 1 shows proposed assorted

Table 2: Keywords associated to Category.

Keywords	Category
Hungry, Starves, Food, Bread, Stuff, Eat, Restaurant, Packet	Food
Living Place, Rest, Accommodation, House, Sleep, Hotel, Cave	Shelter
Hospital, Medicine, Ambulance, Doctor, Nurse, Clinic, Tonic	Medical Emergency
Electricity Power, Fan, Light, Current, Charge, News,	Electricity
Drinking Water, Thirsty, Thirst, Dehydration, Water Tank	Water
Road, Bus Spot, Car, Ambulance, Weather, Airport, railway Station	Transportation

Table 3: Categorization of Sentences.

Category	Tweet
Subjective	No Proper facility is provided by government to flood affected area.
Subjective	Rescue team provides help to cyclone hit area.
Objective	Chief Minister visited to cyclone hit area.

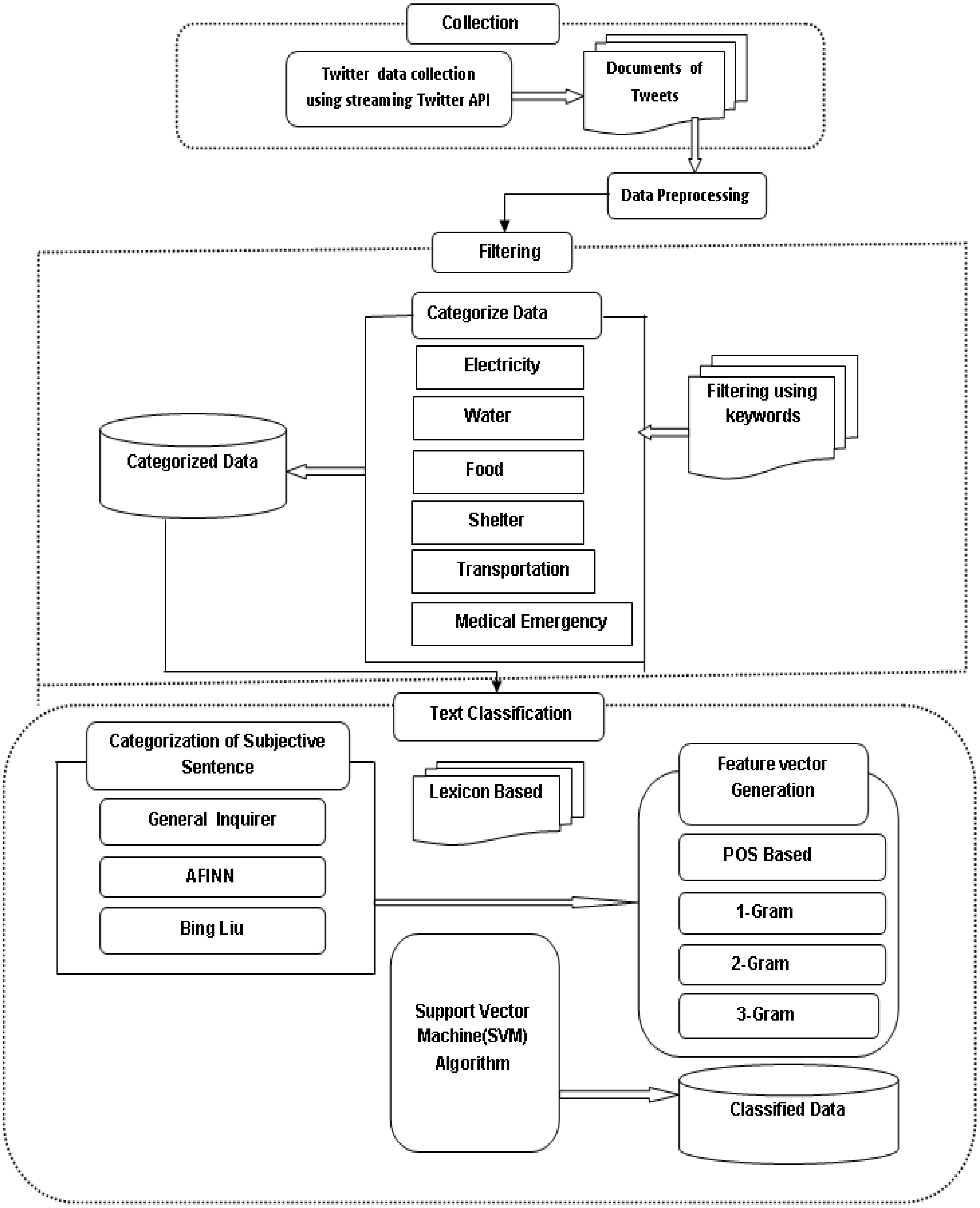


Fig. 1. Proposed Model.

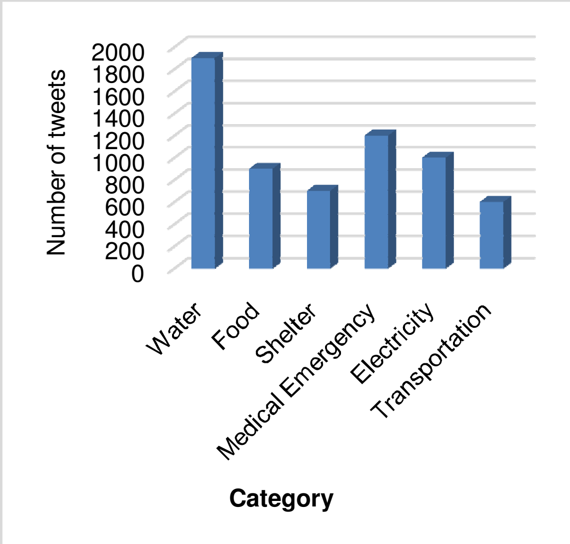


Fig. 2. Number of tweets related to each category.

Bing Liu lexicon: Bing Liu lexicon has all the combination of words which include slang words, morphological of any word and misspelled word [28]. This is polarity-based lexicon which contains around 4683 words which are negative and 2006 words which are positive.

AFINN lexicon: ANEW (Affective norms for English words) is a lexicon that includes words associated with its emotional rating. Psychological view of person indicates the rating of any kind of word. ANEW was come into existence before to rise of micro blogging services. After rising of micro blogging sites there was need of extension of ANEW with slang words and any other type of services belonging to micro blogging. AFINN lexicon is the extension of Affective Norms for English Words which concentrate on the words extracted from social networking sites [29]. AFINN includes jargons of website, obscene of the content of social media, slang words and short form words which are used to determine strengthens of text. Lexicon related to AFINN has around 2476 words. These words belonging to lexicon are categorized based on the negative score and positive score negative score of negative word contains from -1 to -5 and score of positive word contains from 1 to 5.

General Inquirer: GI lexicon used for analysis of content of English written text using Lass well and Harvard dictionaries [30]. It contains 26 parameters of sentiment to adjust the different type of text analysis depending on the requirements. In proposed model we have used first category named “large valence of two categories” of General Inquirer. This category has around 1914 positive words and 2292 negative words. For the purpose of training to the machine learning system, categorize data need to be classified. In order of classifying text of the tweet into negative and positive, we have used 3 lexicons namely General Inquirer, AFINN and Bing Liu lexicon. The natural crisis data is filtered into positive and negative word using list of good word and bad word of these lexicons. The content from a tweet that is not found as either negative or positive is discarded and classified as neutral Table 4 shows the samples of negative and positive tweets from natural crisis data. Positive tweets contribute to find sentiment towards positivity and negative tweets contribute to find sentiment towards negativity. In Table 4 we have taken only sample of large dataset. Neutral sentences indicate sentences which do not contribute to find either positive sentiment or negative sentiment. Neutral kind of sentence also has been included for research propose. It means neutral kind of sentences did not discard from tweets dataset.

Table 4: Sample of Polarity of tweets.

Tweet	Polarity
The army provides better service to victims of flood	Positive
Govt. provide no relief to victim of the cyclone	Negative
CM visits to flood hit area	Neutral

Data analysis and preparation: Bing Liu lexicon and general inquirer lexicon which contains slang words and AFINN lexicon which contains a domain of twitter that are chosen to carry out better classification of natural crisis dataset. In the text of tweets words that are not chosen with these lexicons, they are discarded from tweet. Large numbers of words which are not found for classification of sentiment in each category of natural crisis data are filtered from the negative and positive list of words.

Feature vector Generation: For sentiment categorization feature vector play important role. To train classifier natural crisis dataset needs to be changed into a feature vector most important features for sentiment classification are adjective, adverb, bigram, trigram, bag of words.

Part of Speech Feature: Verbs, adjectives and adverbs play vital role to find the sentiment of the content of the tweets. It is important to carry out part of speech tagging feature to extract adverbs and adjectives from the flow of tweets. In English language, there are eight POS (Part of Speech). Among those part of speech adverb and adjective mainly contribute to identify sentiment of content.

The Pen Banktree Tagger has 46 tags [31]. It put from NLTK (Natural Language Tool Kit) for tagging each sentence of text file. Table 5 shows different forms of adjective and adverbs in proposed assorted sentiment analysis model. From part of speech we can find sentiment using words which are adjectives in English sentence and after adding adverb to adjective it shows the more precise meaning of English sentence example of adjectives are good, bad, fine etc. and examples of adjective with adverb like very good, not bad etc. these adjective and adverbs are defined within predefined dictionaries within NLTK. NLTK Library is built with python language to implement dictionaries.

Table 5: POS Tagging for Adjectives and Adverbs.

Set	Definition
RB	Adverb
RBR	Adverb, comparative
RBS	Adverb, superlative
JJ	Adjective
JJR	Adjective, comparative
JJS	Adjective, superlative

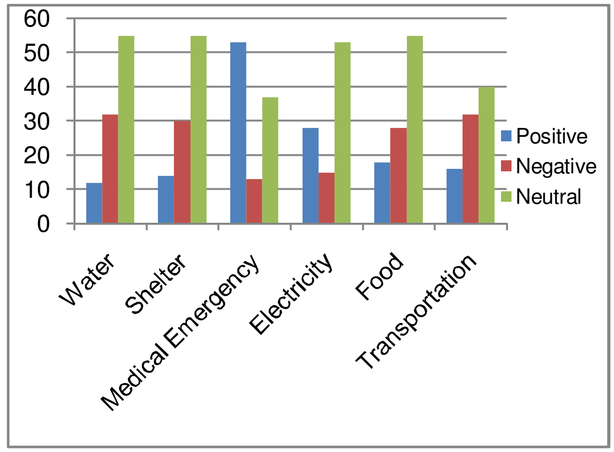
Lexicon Feature: Phrases or opinion words are important for any sentence to identify the sentiment of the text. In different occasion, same phrase or same opinion word may convey different meaning. To understand the pattern in which words come together in the text it is necessary to create size of windows three which contains the word, its successor and its predecessor. Subjective phrase is framed well with the help of window size three. Linguistic feature like 2-gram and 3-gram are applied to subjective phrase to extract meaningful pattern from text. Entire set of 2-Gram and 3-gram is scanned from text when word like not or no is finding in the polarity of positivity it is moved forward to the combination of features with negativity.

Support Vector Machine (SVM): Though there exist several machine learning algorithm Support vector machine algorithms perform better for text classification other than any algorithm [32, 33]. SVM is a machine learning algorithm which works on the boundaries of decision. Boundary of decision splits input into different classes depending on the membership of class. Though there are several machine learning algorithms available SVM is verified to perform better classification than any other machine algorithm in the case of problems of text mining. Support Vector Machine is supervised technique that performs classification and regression tasks by constructing nonlinear decision boundaries. For a given category C = {P, N} where P is the collection of positive samples and N is the collection of negative samples. P is defined as P = $\sum_{n}^{i}(d_i, +1)$ and N is defined as N = $\sum_{n}^{i}(d_i, -1)$.

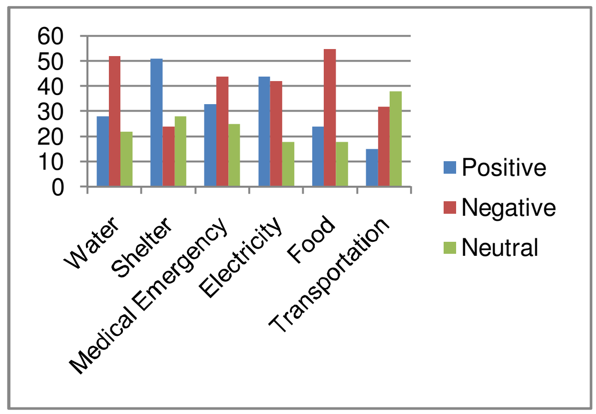
In SVM, dimensional space is used to construct set of hyper planes or hyper plane that divides data. We have converted natural crisis related data into feature vector V which consists set of features. SVM divides disaster related text based on category into positive text and negative text with maximum margin through the calculating of hyper plane in SVM algorithm.

IV. EXPERIMENTAL SETUP

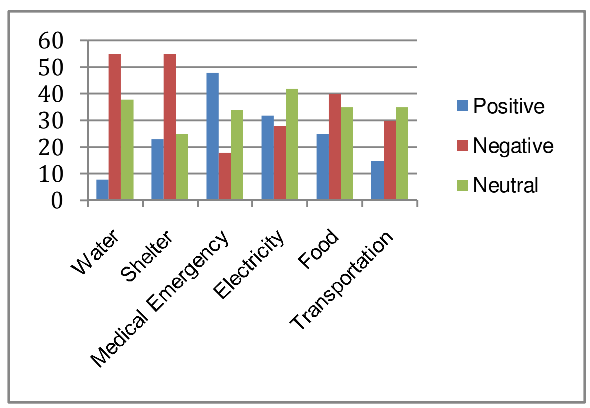
We have collected twitter data from natural crisis events for experimental purpose. We have extracted tweets from using Twitter API. Natural crisis related tweets was extracted from Kedarnath flood, Chennai flood and fani cyclone. In June 2013, cloudburst centered on the north India state of Uttarakhand caused devastating floods in the Kedarnath Shrine and becoming the country's worst natural crisis since the 2004 tsunami. We have collected 10,000 kedarnath flood related tweets from 17/06/2013 to 15/12/2013. In November 2015, Chennai flood was generated by heavy rain. We have collected 6000 Chennai flood related tweets from 08/11/2015 to 04/01/2016. In April 2019, Cyclone fani was the strongest tropical cyclone to strike the Indian state of Odisha. We have collected 7500 fani cyclone related tweets from 26/04/2019 to 01/06/2019. After combining tweets of all the three natural crisis events, we have extracted 23500 tweets for experimentation proposes. For text analysis we have collected 23,500 tweets using Streaming of Twitter API which allows extracting only past seven-day data, rest of disaster related data was extracted from Historical data recovery tool for twitter 2017. The proposed methodology is implemented using Python Programming language and Apache Spark Big Data Framework.



(a)



(b)



(c)

Fig. 3. Percentage of positive and negative tweets of categorized data by (a) GI (b) AFINN (c) Bing Liu.

Fig. 3 shows result of data scanning strategy with lexicon to finding sentiment of each category related to tweet dataset during natural crisis period. Each category of natural crisis data is scanned with lexicon to provide the result of sentiment. We have extracted result in the three form of polarity such as positive, negative and neutral. Support Vector machine-based classification has two sets namely classification set and training set. In SVM, the training set builds a model of training which is used to predict whether the input text is negative or positive. We have used two methods to evaluate the classification and learning phase of proposed model. First method which is used in proposed model is based on lexicons. In this method, entire text is scanned with GI Lexicon, AFINN lexicon and BL Lexicon. Once the positive and negative word are found in the text by scanning with lexicons then subjective sentences are changed into vector of feature by applying 2-gram (Bigram) and 3-gram (trigram) feature. Second method separates subjective sentence from text related to natural crisis data. Each category of content related to natural crisis dataset filters from subjective sentences. Neutral (Objective) sentences are discarded and POS (Part of Speech) tagging is performed using sentences which are subjective. Result of support vector machine is evaluated by calculating recall, precision and F-Score. These are the three measures for evaluating performance of classification system. Precision is percentage of classified text which is relevant. Recall is Percentage of classified text which is retrieved. F-score is the ratio of combination of recall and precision. Table 6 shows result of disaster text classification using 2-gram feature vector and Table 7 shows result of natural crisis text classification using 3-gram feature vector 2-gram feature and 3-gram feature we have used for first method for classifying natural crisis related text. Table 8 shows result of classifying of natural crisis related text using second method. For second method, we have used POS tagging feature before applying classification to natural crisis related text. In the natural crisis text classification using phrase of subjective, the phrase which formed using 2-gram (bigram) performed better than 3-gram (trigram). Phrase of subjective sentence combined with SVM using 2-gram or bigram of BOW (Bag of word) feature yield better accuracy of classification as shown in Table 6. Table 7 shows the performance of phrase of subjective combined with SVM (support vector machine) using 3-gram or trigram of BOW (bag of word) feature as shown in Table 6. It not performed as well as shown in Table 5. It means bigram (2-gram) feature with support vector machine beats to the trigram (3-gram) feature with SVM if we applied to natural crisis data. Table 8 shows the accuracy of POS tagging feature combined with SVM (support of vector machine is) it not performed as well as Bag of words feature combined with SVM. It In the comparison of first method using BOW combined with SVM as shown in Table 6 and 7 and second method using POS tagging combined with SVM as shown in Table 8, it is found that first method performs better for classification of natural crisis data based on categorization.

Table 6: Text Classification using 2-gram Feature.

Lexicon	Measure	Water	Shelter	Medical emergency	Electricity	Food	Transportation
General Inquirer	Precision	96%	89%	91%	91%	94%	94%
	Recall	96%	87%	89%	87%	91%	93%
	F-Score	96%	86%	89%	88%	92%	92%
Bing Liu	Precision	94%	88%	93%	91%	92%	90%
	Recall	93%	86%	92%	87%	91%	89%
	F-Score	93%	83%	92%	88%	90%	88%
AFINN	Precision	94%	85%	91%	89%	96%	92%
	Recall	94%	81%	92%	87%	95%	96%
	F-Score	93%	79%	91%	86%	95%	93%

Table 7: Text Classification using 3-gram Feature.

Lexicon	Measure	Water	Shelter	Medical emergency	Electricity	Food	Transportat ion
General Inquirer	Precision	94%	88%	89%	80%	88%	90%
	Recall	94%	86%	86%	79%	89%	92%
	F-Score	93%	85%	88%	78%	87%	91%
Bing Liu	Precision	93%	82%	88%	89%	80%	89%
	Recall	92%	83%	86%	76%	79%	87%
	F-Score	91%	81%	85%	77%	78%	87%
AFINN	Precision	91%	83%	83%	84%	83%	90%
	Recall	92%	81%	74%	80%	81%	92%
	F-Score	91%	81%	77%	81%	81%	90%

Table 8: Text Classification using POS tagging.

Lexicon	Measure	Water	Shelter	Medical emergency	Electricity	Food	Transportation
General Inquirer	Precision	50%	70%	62%	62%	63%	62%
	Recall	34%	68%	63%	54%	44%	54%
	F-Score	36%	69%	63%	57%	50%	58%
Bing Liu	Precision	79%	62%	67%	61%	53%	60%
	Recall	61%	54%	60%	61%	58%	58%
	F-Score	67%	57%	62%	60%	63%	58%
AFINN	Precision	57%	62%	50%	50%	44%	57%
	Recall	51%	57%	57%	49%	58%	58%
	F-Score	52%	58%	58%	44%	53%	58%

V. CONCLUSION

In this research paper, various phases of sentiment categorization are discussed and assorted sentiment analysis model for natural crisis using big data driven approach is proposed. The main role of this research paper is to study classification methodology for any kind of disaster situation and another contribution of this paper is to do detailed study of learning during any natural crisis situation. Classification methodology is used to classify the requirements of persons during the time of natural crisis. For our natural crisis dataset Combination of SVM (Support Vector Machine) and subjective phrase with bag of word feature yield better classification accuracy for natural crisis dataset.

VI. FUTURE SCOPE

Despite of various opportunities in utilizing social networking site for natural crisis response and recovery there are few concerns related to its usage. There are challenges involved in using social networking site for natural crisis response and recovery include, difficulties in gathering natural crisis related data to generate better model of sentiment for natural crisis analysis, shortage of standard natural crisis dataset or natural crisis related lexicon for accurate evaluation of need persons, in future, to solve these challenges ontology can be build according to the needs of the persons, and also generate a lexicon with natural crisis related keywords.

ACKNOWLEDGEMENT

We are very thankful to all the faculty members, research scholars and other supporting staffs of MANIT, Bhopal, India, for the continued, encourage and support to complete this research work.

Conflict of Interest. There is no conflict of research in this research article. The manuscript has not been submitted to, nor is under review in another journal or other publishing venue.

REFERENCES

[1]. Gerber, M. S. (2014). Predicting crime using Twitter and kernel density estimation. *Decision Support Systems*, *61*, 115-125.

[2]. Gensler, S., Völckner, F., Liu-Thompkins, Y., & Wiertz, C. (2013). Managing brands in the social media environment. *Journal of interactive marketing*, *27*(4), 242-256.

[3]. Graham, M. W., Avery, E. J., & Park, S. (2015). The role of social media in local government crisis communications. *Public Relations Review*, *41*(3), 386-394.

[16]. He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, *33*(3), 464-472.

[17]. Wu, D. D., Zheng, L., & Olson, D. L. (2014). A decision support approach for online stock forum sentiment analysis. *IEEE transactions on systems, man, and cybernetics: systems*, *44*(8), 1077-1087.

[18]. Aramaki, E., Maskawa, S., & Morita, M. (2011). Twitter catches the flu: detecting influenza epidemics

[4]. Qadir, J., Ali, A., ur Rasool, R., Zwitter, A., Sathiaseelan, A., & Crowcroft, J. (2016). Crisis analytics: big data-driven crisis response. *Journal of International Humanitarian Action*, *1*(1), 10-16.

[5]. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1-135.

[6]. Gurman, T. A., & Ellenberger, N. (2015). Reaching the global community during disasters: findings from a content analysis of the organizational use of Twitter after the 2010 Haiti earthquake. *Journal of Health Communication*, *20*(6), 687-696.

[7]. Tapia, A. H., Moore, K. A., & Johnson, N. J. (2013). Beyond the trustworthy tweet: A deeper understanding of microblogged data use by disaster response and humanitarian relief organizations. *Proceedings of the 10th international ISCRAM conference*, 770-779.

[8]. Zielinski, A., Middleton, S. E., Tokarchuk, L. N., & Wang, X. (2013). Social media text mining and network analysis for decision support in natural crisis management. *Proceedings of the 10th international ISCRAM conference,* 840–845.

[9]. Haarsma, R. J., Hazeleger, W., Severijns, C., De Vries, H., Sterl, A., Bintanja, R., ... & van den Brink, H. W. (2013). More hurricanes to hit western Europe due to global warming. *Geophysical Research Letters*, *40*(9), 1783-1788.

[10]. Leong, C. M. L., Pan, S. L., Ractham, P., & Kaewkitipong, L. (2015). ICT-enabled community empowerment in crisis response: Social media in Thailand flooding 2011. *Journal of the Association for Information Systems*, *16*(3), 174-212.

[11]. Umihara, J., & Nishikitani, M. (2013). Emergent use of Twitter in the 2011 Tohoku Earthquake. *Prehospital and Disaster Medicine*, *28*(5), 434-440.

[12]. Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up?: Sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing, 10,* 79-86.

[13]. Kim, H., Howland, P., & Park, H. (2005). Dimension reduction in text classification with support vector machines. *Journal of Machine Learning Research*, *6*, 37–53.

[14]. Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, *2*(5), 1-14.

[15]. Hu, Y. H., & Chen, K. (2016). Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings. *International Journal of Information Management*, *36*(6), 929-944.

using Twitter. In *Proceedings of the 2011 conference on empirical methods in natural language processing* (pp. 1568-1576).

[19]. Gitto, S., & Mancuso, P. (2017). Improving airport services using sentiment analysis of the websites. *Tourism management perspectives*, *22*, 132-136.

[20]. Subrahmanian, V. S., & Reforgiato, D. (2008). AVA: Adjective-verb-adverb combinations for sentiment analysis. *IEEE Intelligent Systems*, *23*(4), 43-50.

[21]. Hogenboom, A., Heerschop, B., Frasincar, F., Kaymak, U., & de Jong, F. (2014). Multi-lingual support for lexicon-based sentiment analysis guided by semantics. *Decision support systems*, *62*, 43-53.
[22]. Sakaki, T., Okazaki, M., & Matsuo, Y. (2012). Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Transactions on Knowledge and Data Engineering*, *25*(4), 919-931.
[23]. Mukhtar, N., Khan, M. A., & Chiragh, N. (2018). Lexicon-based approach outperforms Supervised Machine Learning approach for Urdu Sentiment Analysis in multiple domains. *Telematics and Informatics*, *35*(8), 2173-2183.
[24]. Vo, B. K. H., & Collier, N. (2013). Twitter emotion analysis in earthquake situations. *International Journal of Computational Linguistics and Applications*, *4*(1), 159-173.
[25]. Ragini, J. R., & Anand, P. R. (2016). An empirical analysis and classification of crisis related tweets. In *2016 IEEE International Conference on Computational Intelligence and Computing Research,* pp. 1-4
[26]. Vieweg, S., Hughes, A. L. and Palen, K. (2010). Microblogging during two natural hazards events: What twitter may contribute to situational awareness. *Conference on human factors in computing systems*, CHI'10, 1079-1088.
Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010). Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1079-1088).
[27]. De Albuquerque, J. P., Herfort, B., Brenning, A., & Zipf, A. (2015). A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International Journal of Geographical Information Science*, *29*(4), 667-689.
[28]. Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, *5*(1), 1-168.
[29]. Denge, P. & Wiebe, Y. (2015). An entity/event-l evel Sentiment Corpus. *Proceedings of the 2015 conference of North American chapter of the association for computational linguistics human language technologies*, 1323-1328.
[30]. Stone, P. J., Bales, R. F., Namenwirth, J. Z., & Ogilvie, D. M. (1962). The general inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information. *Behavioral Science*, *7*(4), 484-498.
[31]. Marcus, M., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational linguistic*, *19*(2), 313-330.
[32]. Kim, H., Howland, P., & Park, H. (2005). Dimension reduction in text classification with support vector machines. *Journal of Machine Learning Research*, *6*, 37-53.
[33]. Aramaki, E., Maskawa, S., & Morita, M. (2011). Twitter catches the flu: detecting influenza epidemics using Twitter. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 1568-1576).
[34]. Sahu, Y., Thakur, G. S., & Dhyani, S. (2019). Semantic Rules based linguistic Computational Model of Sentiment Analysis, *International Journal of Emerging Technologies*, *10*(3), 238-243.

How to cite this article: Dhyani, S., Thakur, G.S. and Sahu, Y. (2019). Assorted Sentiment Analysis Model for Natural Crisis Response and Recovery using Big Data driven Technology. *International Journal on Emerging Technologies*, *10*(4): 345–353.