International Journal of Civil Engineering and Technology (IJCIET)
Volume 8, Issue 8, August 2017, pp. 20–29, Article ID: IJCIET_08_08_003
Available online at http://http://www.iaeme.com/ijciet/issues.asp?JType=IJCIET&VType=8&IType=8
ISSN Print: 0976-6308 and ISSN Online: 0976-6316

TEXT MINING ON REAL TIME TWITTER DATA FOR DISASTER RESPONSE

Myneni Madhu Bala
Professor, Department of Computer Science,
Institute of Aeronautical Engineering, Hyderabad, India

K. Navya
Assistant Professor, Department of Computer Science and Engineering,
MLR Institute of Technology, Hyderabad, India

P. Shruthilaya
Assistant Professor, Institute of Aeronautical Engineering, Hyderabad, India

ABSTRACT

Social media such as micro blogging services have a significant impact on the day-to-day lives of people. These services are currently being used by government agencies to interact and communicate information to general public. They also bring an effective collaboration of all stakeholders for dissemination of information during an emergency. Social media is capable of providing spontaneous information during emergency/disaster situations unlike news media, therefore, particularly micro blogging services, have the potential to be adopted as an additional tool for emergency services. In the present work the authors by mining real time data from twitter TM tried to predict the impending damage in the following days during flood scenario. The users of twitter provide important information such as warnings, location of an event, first hand experiences. Such information is collected, preprocessed, geo located and filtered. From the collected information, geo-coded data is prioritized to that of text data. Then the data is analyzed to find the course of the disaster through regression analysis. Later, disaster curve is extrapolated for prediction of damage susceptible locations in the following days. The results are validated by analyzing the past events. In this study, 2015 Chennai flood data is used to validate the results. The study has the potential to facilitate disaster managers for better response operations during emergencies.

Key words: Data mining, disaster computing, Chennai floods, sentiment polarity score.

Cite this Article: Myneni Madhu Bala, K. Navya and P. Shruthilaya, Text Mining on Real Time Twitter Data for Disaster Response. International Journal of Civil Engineering and Technology, 8(8), 2017, pp. 20–29.
http://www.iaeme.com/IJCIET/issues.asp?JType=IJCIET&VType=8&IType=8

1. INTRODUCTION

The number of users and messages in blogs and micro blogs has been continuously growing in recent years, fostered by the unfold of always-connected mobile devices. Microposts, the short messages revealed in microblogs, typically report the standing of users in an exceedingly social or physical context, remodeling the users themselves in period of time sensors concerning their native surroundings. Therefore, microblogging services like Twitter have been tested to be AN innumerable supply of data for different tasks, like opinion mining for industrial purposes [8], detection social events like festivals or political unrest [6], and detecting natural disasters in period of time, for instance earthquakes [12, 10]. Event detection is typically performed by discovering unusual activity patterns, focused on a particular geographic area or on a given topic (usually specified by means of keywords). In order to carry out the detection tasks effectively, it is important to mine the right kind of information from the huge flow of posts (277, 000 tweets per minute2 all over the world). Detection results may vary greatly depending on the search terms, according to [12].

The purpose of a tweet could additionally vary de unfinished on the user: [5] identified four totally different user intention varieties, varied from coverage news to oral communication. Therefore, analysing the content of microposts could facilitate to pick solely those who are relevant to a determined task. Within the case of natural disasters, and additional specifically within the case of disaster management, finding posts that indicate a scenario of danger, worrying or generic alarm, could prove crucial. In such cases, posts that report in progress news, opinions, {or even|orcouldbe|or perhaps} ironic comments don't seem to be significantly relevant and so may complicate the task of analysing the flux of knowledge in such a crucial scenario. Recently, there's a growing interest of the tongue process (NLP) analysis community on Sentiment Analysis (SA) or opinion mining, as testified by the new challenges in SA at totally different information processing conferences, like SemEval [11, 7] or Evalita [2]. As a results of this analysis, some nlp ways and tools were tailored to work on microposts, overcoming a long gap attributable to the actual nature of the language utilized in these short messages. The results obtained in these challenges show that it's currently attainable to find some forms of irony and classify posts consistent with their polarity or judgment. during this paper, we have a tendency to formulate the hypothesis that we will use judgment, polarity and irony detection tools to filter effectively microposts associated with natural disasters and so enhance the accuracy of the knowledge available for the disaster management tasks. We suppose that subjective tweets are more important that objective ones because they are more likely to come from a person involved in the event rather than an objective tweet which is just reporting some news. We also suppose that ironic tweets are more likely to appear afterwards, as an example to criticize the response or blame the government. Finally, we have a tendency to suppose that tweets with a negative polarity square measure additional probable to contain info concerning dangers or emergency things within the context of a natural disaster.

2. RELATED WORK

Current SA systems utilized in polarity classification and irony detection ar largely supported machine learning approaches that exploit each surface options like emoti- 1185 A Study on the 2014 urban center Floodings cons, exclamation marks and capital magnitude relation, and lexicon based options. Lexicons may be thought-about as have an effect on dictionaries that map a word into a polarity score (positive, negative). for example, SentiWordNet3 [1] maps word senses (WordNetsynsets) into a polarity score: “good” (first sense) features a positive score of zero.75, and “worst” features a negative score of zero.75. we tend to participated within the Evalita2014 Italian SENTIPOLC (SENTIment and POLarity Classification) task with such a system, named IRADABE [3], getting one in all the most effective ends up in the sound judgement (3rd with zero.6706 Fmeasure), polarity classification (2nd with zero.6347) and irony detection (2nd with zero.5415) tasks [2]. IRADABE depends on a Support Vector Machine (SVM) exploitation surface and lexiconbased options. The lexicons are custom-made from English to Italian exploitation artificial intelligence. For the experiments given during this paper, IRADABE was trained on the entire SENTIPOLC training set and test set, consisting in 6448 tweets in Italian on numerous random topics, from politics to soccer. The dataset was POS-tagged exploitation TreeTagger4 . Here we tend to describe the options employed by IRADABE:

Bag-of-Words.

The foremost frequent words from the coaching corpus;

Emoticons frequency.The frequency of emoticons expressing sound judgement, positivity or negativeness; Negative Words frequency. Frequency of words that triggers negation, non/no ,avversative conjunction or adverbs; computer address info frequency. the amount of hyperlinks during a tweet; sound judgement options. we tend to took under consideration the presence of verbs conjugated at the primary and second persons (those endings in “-o”, “-i”, “-amo”, “-ate/ete”) and private pronouns (“io”, “tu”, “noi”, “voi”, and their direct and object versions); Tweet Length and capital magnitude relation. The length in words of every tweet.we tend to took under consideration additionally the magnitude relation between the capital words and length of the tweet; SentiWordNet. we tend to used the positive and negative scores to derive six features: positive/negative words count, the add of the positive scores within the tweet, the add of negative scores within the tweet, the balance (positive-negative) score of the tweet, and therefore the variance of SentiWN scores within the tweet; Hu-Liu Lexicon5 . we tend to derived 3 options from this lexicon: positive and negative words count, balance (# of positive words- # of negative words); AFINN Lexicon6 . This lexicon contains 2 word lists labelled with polarity valences from -5 (negative) to +5 (positive). we tend to derived five options from this lexicon: three http://sentiwordnet.isti.cnr.it four http://www.cis.uni-muenchen.de/~schmid/tools/ TreeTagger/ five http://www.cs.uic.edu/~liub/FBS/ sentiment-analysis.html vi https://github.com/abromberg/sentiment_analysis/ blob/master/AFINN/AFINN-111.txt positive/negative word count, add of positive and negative scores; overall balance of scores within the tweet; Whissel wordbook [13]. This lexicon contains eight, 700 Italian words with values of Activation, representational process and Pleasantness associated with each. vary of scores go from one (most passive) to three (most active). We tend to derived six features: average activation, representational process and pleasantness, and therefore the variance of the various scores. We tend to thought that Associate in Nursing elevate score in one in all these options might indicate Associate in Nursing out-of-context word, so indicating a probably ironic comment; Italian “Taboo Words”. Knowing the operate of taboo words to trigger humour, catharsis, or to spice up opinions, we tend to set to use an inventory of taboo Italian words that we tend to extracted from Wiktionary7 ; Counter-Factuality and Temporal Compression [9]. Frequency of terms that indicate Associate in Nursing abrupt amendment during a narrative.

Current Storm Troops systems employed in polarity classification and irony detection area unit largely supported machine learning approaches that exploit each surface options like emoti- 1185 A Study on the 2014 metropolis Floodings cons, exclamation marks and majuscule magnitude relation, and lexiconbased options. Lexicons is thought-about as have an effect on dictionaries that map a word into a polarity score (positive, negative). as an example, SentiWordNet3 [1] maps word senses (WordNetsynsets) into a polarity score: “good” (first sense) includes a positive score of zero.75, and “worst” includes a negative score of zero.75.

we have a tendency to participated within the Evalita2014 Italian SENTIPOLC (SENTIment and POLarity Classification) task with such a system, named IRADABE [3], getting one in all the most effective leads to the subjectiveness (3rd with zero.6706 Fmeasure), polarity classification (2nd with zero.6347) and irony detection (2nd with zero.5415) tasks [2]. IRADABE depends on a Support Vector Machine (SVM) victimisation surface and lexiconbased options. The lexicons are custom-made from English to Italian victimisation MT. For the experiments conferred during this paper, IRADABE was trained on the entire SENTIPOLC training+test set, consisting in 6448 tweets in Italian on numerous random topics, from politics to soccer. The dataset was POS-tagged victimisationTreeTagger4 . Here we have a tendency to describe the options employed by IRADABE:We hand-picked manually a collection of hashtags, toponyms and matter fragments (we might decision them topics) that we have a tendency to judged to be vital from a disaster management perspective. as an example, we have a tendency to enclosed the toponyms associated with the affected zones (Montoggio, Bisagno, Sturla, Fereggiano,...) as they're listed on the Italian Wikipedia page on the disaster; we have a tendency to enclosed hashtags like #allertameteo (meteo alert), #protezionecivile (civil protection agency), #alluvionege (Genoa flooding); we have a tendency to enclosed fragments like “ondata di piena” (“surge wave”), “mancatoallarme” (“missed alarm”), “invasodekalitrefango” (“flooded by mud”). Then we have a tendency to extracted for every time-frame the list of trending hashtags, toponyms and topics, consistent with (1). we have a tendency to calculated accuracy because the range of trending vital things (hashtags, toponyms or topics) divided by the whole range of detected trending things, and coverage because the range of detected trending vital things all told time frames divided by the whole range of vital things. The results of this analysis area unit shown in Table one. The results show that if we have a tendency to limit the analysis to subjective and negative tweets, we will acquire higher average accuracy (in different words, the next share of the rumored things for every fundamental measure area unit relevant) however at the expenses of coverage. Note that fifteen and sixteen of the things detected victimisation the negative and subjective tweets, severally, weren't detected victimisation the entire dataset. In Figure three we have a tendency to show the typical accuracy, hour by hour, on all things, victimisation the total dataset or solely the subjective and negative tweets. we have a tendency to manually associatealysed a number of the tweets with an assigned polarity or irony label. we have a tendency to discovered that a lot of positive tweets (58) were originated by one account that was posting video links, adding “buonavisione!” (“enjoy the show!”) with associate happy emoticon, yet their content (probably an automatic posting). The second most frequent poster of positive tweets totalled solely eight tweets, all thanking the volunteers that helped in removing mud from the streets. Most positive tweets looked as if it would be encouragement messages to the population like “ForzaGenova!” (“Come on Genoa!”). we will conclude that the majority of positive tweets don't seem to be pertinent from the disaster management perspective, corroborating our initial hypothesis on this category. we have a tendency to found a high rate of false positives among the tweets tagged as ironic, indicating that this task is however a tough one. However, among the properly tagged ones, we have a tendency to were able to notice several tweets that criticized the native authorities, like “Ragazzitranquilli #Renzi hour angle dettoche non ci lascer`asoli” (“Don’t worry guys, #Renzi - the Italian prime minister - same that he won’t leave U.S. alone”). These tweets were in all probability detected owing to the high range of politically themed ironic tweets within the SENTIPOLC coaching set. within the case of negative tweets, we have a tendency to were able to establish totally different styles of negative feelings, from worry and worry to rage and frustration.The author done the extensive work on cloud applications data and social network data on multiple themes [18,19].

3. SENTIMENT ANALYSIS

The sentiment analysis is the classification of texts into four categories as positive, negative, neutral, and objective texts. The neutral texts are subjective, but they include almost equal amounts of positive and negative charges. These categories are not satisfactory for our disaster management, because negative sentiments have important role. Namely, fear, panic, anxiety, resentment are negative sentiments that must be identified as much as accurate as possible. Each of these sentiments dictates a specific line of action in disaster relief. These sentiment and emotion analysis in very effective in the assessment of disaster impact and people's needs or help in the time of emergency conditions.

The advancements in automated tools based on analytics applied to social media are helpful for disaster relief and mitigation. The tools are capable of doing the following functionalities for the dynamic and effective disaster management system:

- Data is collected in almost real-time (Dynamic)
- The collected data is cleaned from noise (Trustiness)
- Effective processing to produce timely and relevant information (Management and Assessments)

Such systems can be integrated in larger, multi-layered emergency warning, monitoring and management of disasters. Current reserves systems utilized in polarity classification and irony detection are largely supported machine learning approaches that exploit each surface options like emoti-1185 A Study on the 2014 urban center Floodings cons, exclamation marks and majuscule magnitude relation, and lexicon based options. Lexicons is thought-about as have an effect on dictionaries that map a word into a polarity score (positive, negative). as an example, SentiWordNet3 [1] maps word senses (WordNetsynsets) into a polarity score: "good" (first sense) contains a positive score of zero.75, and "worst" contains a negative score of zero.75. we have a tendency to participated within the Evalita2014 Italian SENTIPOLC (SENTIment and POLarity Classification) task with such a system, named IRADABE [3], getting one in every of the most effective leads to the sound judgement (3rd with zero.6706 Fmeasure), polarity classification (2nd with zero.6347) and irony detection (2nd with zero.5415) tasks [2]. IRADABE depends on a Support Vector Machine (SVM) exploitation surface and lexicon based options. The lexicons are tailored from English to Italian exploitation computational linguistics. For the experiments conferred during this paper, IRADABE was trained on the whole SENTIPOLC training and test set, consisting in 6448 tweets in Italian on numerous random topics, from politics to soccer. The dataset was POS-tagged exploitation TreeTagger4 . Here we have a tendency to describe the options employed by IRADABE:

Sentiment analysis may be a text classification downside that deals with extracting data gift at intervals the text [6, 8]. This extracted data is then more classified per its polarity as positive, negative or neutral. It is outlined as a procedure task of extracting sentiments from the opinion. Some opinions represent sentiments and a few opinions don't represent any sentiment. Sentiments: Opinions or in alternative sense is recognized as someone's linguistic expressions of emotions, beliefs, evaluations etc. Analysis: To capture the opinions from a pool of users whether or not the opinion is positive, negative or neutral [9]. Benefit: give economical data in deciding.

Twitter messages have several distinctive attributes as shown below [10, 11]:

Length: absolutely the most length of a Twitter communication is 100 and forty personas.

Data access: yet one more distinction may be the specifications of data without delay accessible. Where as exploitation Twitter API, it's quite all to simple to amass infinite twitting referring to teaching.

Dialect style: Twitter customers submit mail messages from many alternative media, furthermore as their specific cellular phones. Your frequency of misspellings in conjunction with slang throughout twitting is way higher than throughout alternative names.

4. MACHINE LEARNING

Many ways of Machine learning algorithms are designed to live the likelihood of uncertainty supported past expertise. The MLalgorithms learn the patterns on information then provide the output. High the educational observations smart is that the mil rule. Basic options of the machine learning algorithms are as Learning speed, Memory utilization, Accuracy, comprehensibility.

KNN methodology K nearest neighbors may be an easy rule that stores all accessible cases and classifies new cases supported a similarity live (e.g., distance functions). KNN has been utilized in applied math estimation and pattern recognition already within the starting of 1970's as a non-parametric technique [13].

Classification KNN predicts the classification of some extent Xnew employing a procedure love this: one. Notice the NumNeighbors points within the coaching set X that ar nearest to Xnew. 2. Notice the NumNeighbors response values Y to those nearest points. Assign the classification label Ynew that has smallest expected misclassification value among the values in Y two.2. Naïve Bayesian methodology Naïve Baye's ways ar supervised learning ways supported baye's theorem.

5. ISSUES AND CHALLENGES

Information Overload Issue:

Information overload or the complete volume of the data stream, which exceeds human processing capacity, has been identified in past research as one of the major issue to the use of social media for gathering information during a disaster [12]. Miller [14] has discussed about levels in human cognitive and perceptual tasks. This suggests that technological enhancements to SM that "chunk" social media data into groupings that are identifiable by the user prior to actual examination of the data may be useful to combat information overload. That "chunking" is an effective way to mitigate the natural human limitations on the amount of data that can be retained in short term memory, processed, etc. has been suggested by other researchers as well. For example, Tegarden [14] reports that visualized techniques can help in chunking information in a way that increases the amount of information people can retain. The further improvements in chunking[13] is to be done automatically.

In other words, a third dimension of information overload is the extent to which the information is organized and categorized. Users can deal with more information if it is "chunked" and organized in some way by the system. A necessary step in making sense of SM data is to organize or classify the information into useful categories. SM are to some extent selforganizing. For example, Twitter users invent and use hashtags to describe topics, and Facebook users establish and name new groups to discuss topics. However, the basic organizing tools and conventions for public SM use are inadequate for the purpose of disaster management.

The Information Overload is a problem that presents a barrier to use of social media during management of emergencies. However, it also points the way forward to solutions to this problem. The research contributions are needed in this direction for creating the effective bridge between developers of technology to enhance SM data use and the potential users of those technologies in crisis situations. This can provide guidance to both technologists and researchers as they work to find ways to improve the presentation of SM data so that the wealth of information on it can be effectively and efficiently used.

6. USE OF ANALYTICS FOR MONITORING SOCIAL MEDIA DATA

Disaster management systems require tools that detect dynamically on time the occurrence of disasters and accurate assessment on detailed picture of the situation. The dynamic data source is social media data. The monitoring of people responds in SM is a critical role in crowd sources. Hence the need of such tools may help under harsh conditions related to natural disasters by automatic identification of the type, scope, location, force, and implications of the disaster. The research focuses on the use of analytics to identify emergencies and recent disasters, based on social media search and direct to the relief organizations according to the needs.

Method and Tools

The main focus is finding relevant keywords to search, ways to improve the search algorithms, and techniques to interpret the findings for assessment.

Multilevel Search Method

Single keyword search produces a vast amount of ‘noise’, that is, of messages that are not related to emergency situations. For example, some terms as ‘disaster’ and ‘catastrophe’ are too general and may have figurate meanings. Their use as primary search keywords produces irrelevant information. For avoiding these drawbacks, the solutions are:

1. Collection of relevant hashtags/keywords from web crawlers.
2. Hashtags with highest frequency of tweets.
3. Geographic location with the combination of tagwords.
4. Instead of single word search hierarchical search.

In hierarchical search, start with a few set of keywords and then search the messages with first level keywords. The second level keywords are selected from a set of words that have a large chance of occurrence together with the first one identified. The high frequency of joint occurrence in a message classifies the keywords in families of words, related to specific scenarios. The search optimization is very essential step due to the importance of timely and accurate results for disaster warnings, classified into different services. The terms in the controlled vocabulary are considered with significance to a type of emergency and the most appropriate terms are searched first on the second level.

7. ANALYSIS ON PAST EVENT ON CHENNAIFLOODS

The data Extraction include the hashtag as #Chennaifloods. The sample twitter data after preprocessing is shown in figure 1.

1 brought people together despite challenges fightnatures fury demonetisation story repeats now blackmoney
2 jugaad that helped peoples during chennaifloods
3 un loved her app that saved many in chennaifloods but this yr old techie wants to do more for people
4 missing chennaifloods
5 just got info supply trucks frmbnglr n places getting diverted chnai border loc

6	better news chennai start week chennaifloods showers return midweek darren

The major discussions are shown in word cloud of figure 2.

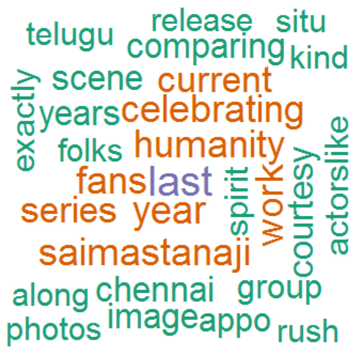

Figure 1 Word cloud of sample data

Location wise data is shown in Figure 2.

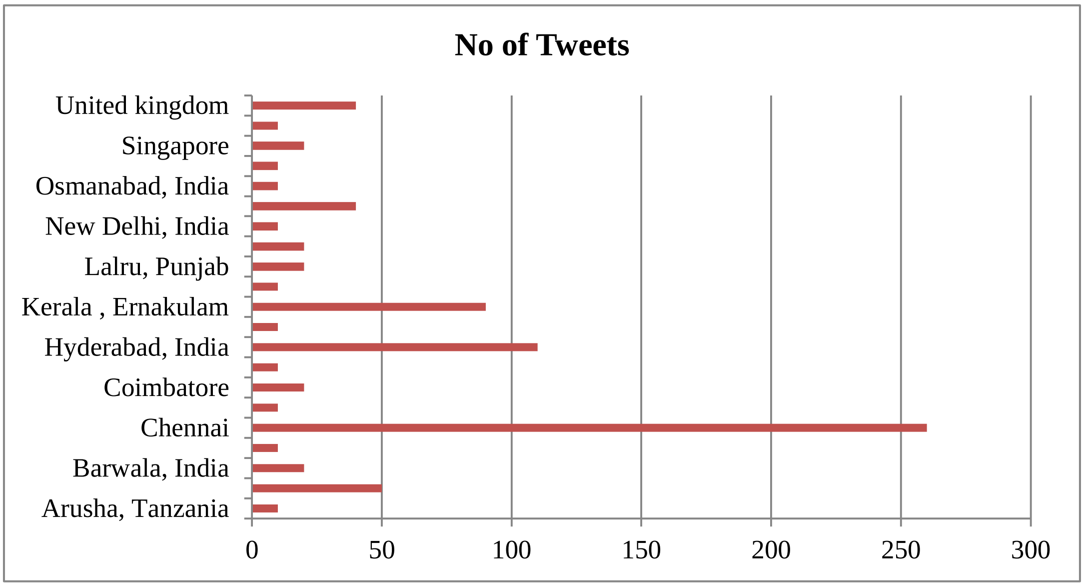

Figure 2 Bar graph of location wise tweets

The trained data description and sentiment score of sample twitter data is shown in table 1.

Table 1 Training data statistics

total.sentences	113
total.words	1169
ave.polarity	0.138
sd.polarity	0.35
stan.mean.polarity	0.395
Uniq word	frequency

The sentiment score of floods is shown in figure 3.

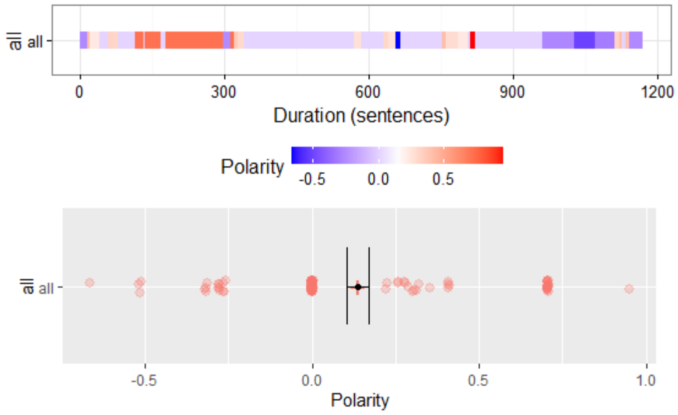

Figure 3 Polarity score of sentiment on chennaifloods disaster response

The positive opinion is expressed about chennai floods on disaster response.The average polarity score of .3 is found on sample data about chennai floods.

8. CONCLUSIONS

Natural disaster response of people is analysed by using machine learning and text mining algorithms on twitter data is used for this analysis. From this location wise tweets and response of people are identified from twitter data. This is useful for automatic assessment of disasters response.

REFERENCES

[1] Baccianella S., Esuli A., and Sebastiani F, “SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining,” In LREC, vol 10, pp 2200–2204, 2010.

[2] Basile V., Bolioli A., Nissim M., Patti V., and Rosso P., “Overview of the Evalita 2014 SENTIment POLarity Classification Task,” In Proceedings of the 4th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA’14), pp 50–57, Italy, 2014.

[3] Hern´andez-Farias I, Buscaldi D, and Priego-Sanchez B, “IRADABE: dapting English Lexiconsto the Italian Sentiment Polarity Classification task,” In Proceedings of the 4th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA’14), pp 75–80, Italy, 2014.

[4] Ihler A, Hutchins J, and Smyth P, “Adaptive event detection with time-varying poisson processes,” In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, pp 207–216. ACM, 2006.

[5] Java A., Song X., Finin T., and Tseng B, “ Why we twitter: understanding microblogging usage and communities,” In Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis, pp 56–65. ACM, 2007.

[6] Lee R, Wakamiya S, and Sumiya K, “ Discovery of unusual regional social activities using geo-tagged microblogs” World Wide Web, 14(4), pp 321–349, 2011.

[7] Nakov P., Kozareva Z., Ritter A., Rosenthal S., Stoyanov V., and Wilson T, “ Semeval-2013 task 2: Sentiment analysis in twitter” 2013.

[8] Pak A. and Paroubek P, “ Twitter as a corpus for sentiment analysis and opinion mining,” In LREC, 2010.

[9] Reyes, P. Rosso, and T. Veale. A multidimensional approach for detecting irony in twitter. Language Resources and Evaluation, 47(1):239–268, 2013.

[10] B. Robinson, R. Power, and M. Cameron. A Sensitive Twitter Earthquake Detector. In Proceedings of the 22Nd International Conference on World Wide Web Companion, WWW ’13 Companion, pages 999–1002, 2013.

[11] S. Rosenthal, P. Nakov, A. Ritter, and V. Stoyanov. Semeval-2014 task 9: Sentiment analysis in twitter. Proc. SemEval, 2014.

[12] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes Twitter users: real-time event detection by social sensors. In Proceedings of the 19th international conference on World wide web, pp 851–860. ACM, 2010.

[13] C. Whissell. Using the revised dictionary of affect in language to quantify the emotional undertones of samples of natural language 1, 2. Psychological reports, 105(2), pp 509–521, 2009.

[14] Hiltz, S.R. and Plotnick, L. (2013) Dealing with information overload when using social media for emergency management: Emerging solutions. Proceedings of the 10th International ISCRAM Conference, Baden Baden, Germany.

[15] Koroleva, K., Krasnova, H., & Günther, O. (2010) ‘STOP SPAMMING ME!’ – Exploring Information Overload on Facebook. in Proceedings of AMCIS, August 12- 15; Lima, Peru.

[16] Miller, George,1956. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. Psychological Review 63 (2): 81–97.

[17] Tegarden, D.P. (1999) Business Information Visualization. Communications of the Association for Information Systems, 1.

[18] Madhubala Myneni, Narasimha Prasad L V and G Geetha Reddy, Automatic Assessment of Floods Impact Using Twitter Data. *International Journal of Civil Engineering and Technology*, 8(5), 2017, pp. 1228–1238.

[19] Madhu Bala Myneni, Narasimha Prasad LV, Intelligent Hybrid Cloud Data Hosting Services with Effective Cost and High Availability, International Journal of Electrical and Computer Engineering, 7(4), 2017.

[20] Google Analytics, https://developers.google.com/analytics/solutions/integration-upload

[21] Inje Bhushan V. and Prof. Ujwalapatil, A Comparative Study on Different Types of Effective Methods in Text Mining: A Survey. *International Journal of Computer Engineering and Technology (IJCET),* 4 (2), 2016, pp. 535-542.