IMPROVING THE EFFICIENCY OF REAL TIME CRISIS MANAGEMENT USING DEEP LEARNING

ASHISH MOHOD[1], NIKHIL BHIVGADE[2], SHUBHAM GANATRA[3], RAHUL GAJBHIYE[4], TUSHAR GIRADKAR[5]

[*1]*Professor, Department of Computer Science and Engineering, Priyadarshini JL College of Engineering, Nagpur, Maharashtra, India.*

[*2,3,4,5]*UG Students, Department of Computer Science and Engineering, Priyadarshini JL College of Engineering, Nagpur, Maharashtra, India.*

ABSTRACT

The most common tool used during disasters is social media. All the people post tweets, posts and filmland on social media showing their support for disaster victims. Crisis Management as a subject basically deals with the operation of coffers and information as far as a disastrous event is concerned and also how effectively and seamlessly one coordinates these coffers. Crisis operation, at the individual and organizational position, deals with issues of planning, coordinating, communication, and threat operation. As the association and operation of coffers and liabilities for dealing with all the philanthropic aspects of extremities, in particular preparedness, response, and recovery to reduce the impact of crisis. Equally important is the knowledge about colourful transnational and public agencies involved in Crisis relief and philanthropic backing. Effectiveness of any tool is understood during the time, when it's utmost demanded, and social media is that tool if used duly can help numerous people during disasters. It contains numerous useful information to help victims during disasters including their position, needs and demands which can be passed on to disaster operation authority to escalate the process of disaster relief.

Keywords: *Disaster Management, Tweet Classifier, Neural Architecture, Natural Language Processing.*

1.INTRODUCTION

People from exigency department spend lot of time trying to organize the world but what part of our life is truly under our control. To come out from extremity situations like natural disasters and extremities, the success of extremity handling operations can be determined by numerous factors like early warning, robust plan, speed of response, effective communication etc. The mortal gets is among the most precious bones and thus, the occasion to cover people's gets in real time during a extremity would be a strong advantage. The crucial task of extremity operation is to fete specific conduct that need to be carried out before forestalment-crisis and after an extremity passed. To execute these tasks efficiently, it's helpful to use data from social media (SM) as social media has come a common channel in responding to numerous exigency situations. Two- way discussion is the core idea of social media (2). Information collected from social media can be employed as one of the essential sources from which exigency askers and social media content judges can prize meaningful information that help to classify and rank new enterprises or give further detail about observed issues (3). When affected or viewer people borrow social media to shoot announcements and updates, ask for or offer help, or report the situation around them, they contribute to information aqueducts that, both resides, and exigency askers depend on during a disaster (4). Because from natural disaster like Hurricane, cataracts, afflictions situation like nimbus contagion or any purposeful manmade conduct, no part of the world is defended from disaster When a extremity strike, it makes chaos and people collect information from the sources most incontinently accessible to them those in their immediate terrain, musketeers, and families by phone, broadcast cautions, radio, TV, online communities, and social media. Grounded on this information, people decide whether to snappily void from the disaster area, seek sanctum, or prepare in a way that keeps them down from threat. Crisis operation department also uses social media to collect, cover and circulate precious information to inform the public thus, it's fruitful to use social media contents to support extremity operation in real time. The main thing to use stoner-generated contents of social media is to distinguish precious information from inapplicable one. The proposed system use bracket as the demarcation system. The classifier plays the part of a filtering ministry. With the help of the stoner, it recognizes the important social media particulars (e.g., tweets), that are related to the event of interest. The named particulars are used as signals to identify sub-events. Then an event is the extremity, whereas sub-events are the matters or content generally bandied (i.e., hotspots like flooding, collapsing of islands, etc. in a specific area of a megacity) during a extremity. These sub-events can be linked by aggregating the dispatches posted on social media networks describing the same specific content (48), (51). This exploration suggests a result to give a timely announcement of dangerous events and to estimate people's response, through the real-time analysis of dispatches on Twitter. Contextual data about an ongoing extremity will be combined with literal data about once heads and reused using deep literacy to give decision, indispensable conduct, and an estimated measure of their impacts. The main purpose of this exploration is to give a deep analysis of nesting goods during a extremity and to develop styles to anticipate them. Proposed work will be carried out by analysing other analogous fabrics and using bettered empirical approach for extremity operation by analysing social media data using deep literacy. The compass of social media computing has expanded extensively, with nearly every branch of software exploration and practice explosively feeling its impact. The term social media computing can be defined as computational facilitation of social studies and mortal social dynamics as well as the design and use of information and communication technologies that consider social environment (16). Whereas the term deep literacy refers to the automated discovery of meaningful patterns in data. Now a day deep literacy is one of the fastest growing areas of computer wisdom, with far- reaching operations. As state, deep literacy can be defined as "computational styles using experience (training data) to ameliorate performance or to make accurate prognostications." Eventually, a disaster is an unforeseen, calamitous event that seriously disrupts the functioning of a community or society and causes individual, community, government, and on-governmental agencies losses that exceed the community's or society's capability to manage using its own coffers. Social media information is frequently irreplaceable incontinently after a unforeseen onset exigency or disaster. It plays a vital part not only after the disaster, but during its entire life cycle

2.REALTED WORK

The problem addressed here is related to several topics like Learning Vector Quantization (LVQ) classification and multiple prototypes, online learning for classification, active learning with budget planning, as well as social media analysis a short overview of above mention topics is presented in the following.

2.1 Multiple Prototype Classification and LVQ Classification:

In prototype-based classification approach data items mapped to a vector representation here data points are classified via prototypes considering similarity measures. Prototypes are selected based on items related/similar to them. A Rocchio classifier explain by C. manning etl. [37]. It illustrates the model of single prototype-based classifier that help to discriminates between two classes, e.g., "relevant" and "irrelevant". In real life scenarios, it is often not possible to describe the data with a single prototype-based classifier due to nature of the data and thus multiple prototype classifiers (i.e., several prototypes) are needed. Self-Organizing Maps (SOM) introduced by T. Kohonen [32] are an unsupervised version of prototype based classification, also known as Learning Vector Quantization (LVQ). In this case, prototypes are initialized (e.g., randomized) and adapted. SOM was also used for SM analysis in the context of crisis management to identify important hotspots [50]. LVQ has been applied to several areas, e.g., robotics, pattern recognition, image processing, text classification etc. [20], [32], [62]. LVQ - in the context of similarity representation, rather than vector-based representation - is analyzed by Hammer et al. [25]. Bezdek et al. [6] review several offline multiple prototype classifiers, e.g., LVQ, fuzzy LVQ, and the deterministic Dog-Rabbit (DR) model. The latter limits the movement of prototypes. In contrast to the previous approaches, Bouchachia [8] proposes an incremental supervised LVQlike competitive algorithm that operates online. It consists of two stages. In the first stage (learning stage), the notions of winner reinforcement and rival repulsion are applied to update the weights of the prototypes. In the second stage (control stage), two mechanisms, staleness and dispersion are used to get rid of dead and redundant prototypes.

Here we propose a quantization multi-prototype algorithm, where the winning prototype is adapted based on the input. In particular, the algorithm relies on online active learning.

2.2 Social Media Analysis for Crisis Management

Recent research studies SM from several technical perspectives. Here we describe existing SM analysis frameworks mostly in the context of crisis management, S. Vieweg and A. Hodges [30], [63] describe the Artificial Intelligence for Disaster Response (AIDR) platform, where persons annotate incoming tweets (similar to Amazon Mechanical Turk). The tweets are then used to train classifiers to identify more informative or relevant tweets. AIDR allows to classify incoming tweets based on different information categories, e.g., casualties, advises, damage report etc. L. Chen et al. [15] performed analysis of tweets related to Flu to identify topics for predicting the Flu-peak. V. K. Neppalli et al. [42] work on sentiment analysis for social media related to Hurricane Sandy. This work illustrate that sentiment of users is related to the distance of the Hurricane to the users. Twitcident described by F. Abel et al. [1] is a framework to search and filter Twitter messages through specific profiles (e.g., keywords). Terpstra et al. [61] show the usage of Twitcident in crisis management. Tweak-the-Tweet introduced by Starbird et al. [60] defines a grammar which can be easily integrated in tweets and therefore automatically parsed. Also, TEDAS described by Li et al. [34] is a system to detect high-level events (e.g., all car accidents in a certain time period) using spatial and temporal information. J.Yin et al. [67] design a situational awareness platform for SM. Tweets are analyzed based on bursty keywords to identify emergent incidents. Ragini et al. [52] combine several techniques to identify people in danger. They examined rule based classification and several machine learning approaches, like SVM, for hybrid classification. Additional information on social media analysis in different crises can be found in Reuter and Kaufhold [53]. Daniel Pohl et.al. proposes a novel active online multiple prototype classifier, called AOMPC it identifies relevant data related to crisis. This algorithm labeled ambiguous unlabeled data using closest prototype or by asking users feedback but the limit of intervention of user is based on budget strategy. AOMPC was evaluated using two type of data ie. Synthetic and Colorado floods and Australia Bushfires data from Tweeter data to understand the behavior of an algorithm. Due to the importance of SM, it is our aim to support emergency management when using the content of SM platforms. Currently, there are systems with crowd-sourcing platform characteristics, but no procedure like online active learning along with dynamic budget strategies and preservation of only required clusters (deleting the cluster after they lose their importance) is available to directly involve emergency management personnel in filtering relevant information.

3.METHODOLOGY

Initially we will be reviewing various methods for calamity and crisis management. These methods will be evaluated in terms of computational complexity, accuracy of evaluation, delay analysis and other output parameters. Once these methods are evaluated, then different parameters will be found out that can help in crisis and calamity management. These parameters will be evaluated by finding out different system efficiency values. The parameters for which the output efficiency is good, those will be used for further processing. Upon identification of these parameters, we will be developing a reinforcement learning algorithm that will take into consideration all these parameters to evaluate calamity prediction and management. The parameters will further be improved with the help of an artificial intelligence that will perform intelligent feature selection and mapping to obtain highest performance. Our proposal is different from the approaches which we have mentioned in our literature review. We are targeting wider sources to understand people need during the crisis and evaluate the social impact using the notion of deep learning. The following methodologies that are incorporated within our system. The overall goal of the system is to facilitate post-disaster relief and increase disaster coordination on social media. This system performs four essential tasks – (i)Pre-processing of tweets, (ii) Filtering the tweets, (iii) Classification of need and availability tweets and (iv) Extracting relevant information from tweets. This model is designed to execute each of the above four tasks in an automated fashion. We brief on the specific methodology involved in each stage of our model.

Figure 3.1: Different Stages in Model

Figure 3.1 depicts the many stages of data modification that will occur in our model. The approaches listed below are included in our system. The system's overall purpose is to make post-disaster assistance coordination activities easier by utilising the large amount of information available on social media. This system performs four key functions: I twitter pre-processing, (ii) tweet filtering, (iii) need and availability tweet classification, and (iv) extracting useful information from tweets. This concept is intended to automate each of the four tasks mentioned above. In the following subsection, we go over the specific methodology used for each of these sub-tasks.

3.1 Pre-processing

We removed URLs (but not email addresses), mentions, brackets, 'RT,' and other non-ASCII characters like #, &, ellipses, and Unicode characters corresponding to emojis from the tweet content using standard pre-processing techniques. Stop words that have no meaning, such as I, me, myself, and the, have also been eliminated. After that, words like earthquake, nepalearthquake, and earthquakenepal were eliminated to aid classification in subsequent steps. Tweets before and after pre-processing are shown in Figures 3.1.1 and 3.1.2.


God save victims of earthquake.. #PrayForNepal #HelpNepal #NepalEarthquake #NepalQuake #savenepal #HelpNepalChildren
We met a group of Malaysians who experienced the earthquake first hand during their trek to Annapurna… https://instagram.com/p/2Qvy2sJ_1s/
Random Act of G,distributing foods, Earthquake victims Nepal@brucepoontip @G_JulieFitzG @Manishsinghvns @shiva_nepal @nytimes need to be distributed to the needy nd #earthquake victims .. #EarthquakeNepal ..

Figure 3.1.1: Tweets before pre-processing

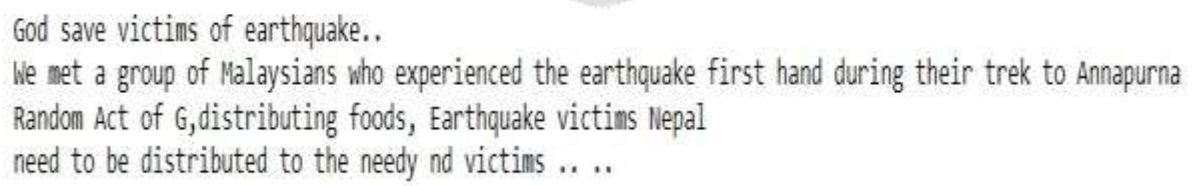
God save victims of earthquake..
We met a group of Malaysians who experienced the earthquake first hand during their trek to Annapurna
Random Act of G,distributing foods, Earthquake victims Nepal
need to be distributed to the needy nd victims .. ..

Figure 3.1.2: Tweets after pre-processing

3.2 Filtering the tweets

In this step we wanted to make sure we only have disaster tweets and remove all non-disaster and purely sympathy tweets. A set of words are taken like need, lack, give, urgent, donate, food, water, hospital, blood. We used word similarity in SPACY. The words in tweet were converted into tokens and compared with the set of words, if the score is above our threshold then the tweets are taken otherwise it is rejected.

3.3 Classification of needs and availability tweets

Convolutional Neural Networks (CNN) have been found to function admirably in the order of calamity related tweets (Caragea et al., 2016; Nguyen et al., 2017). Subsequently, we utilize the CNN of (Kim, 2014) as a standard model. We work on 300-layered word-embeddings and fix the element guides to 100 aspects. We execute convolutional channels with piece size 5 individually, with step 1. At long last, we apply max-pooling prior to going it through a completely associated layer and SoftMax with negative log-probability (NLL) misfortune. We explore different avenues regarding arbitrarily instated embeddings as well as various types of pre-prepared embeddings. In this stage we can utilize both CNN and BERT model to order the tweets into need and accessibility tweets. The outcomes we get in this stage can contain a few blunders as no brain design model gives wonderful exactness, yet we attempt to get greatest F1 score utilizing these models. The last objective of this stage is to go about too prepared classifier for the tweets.

3.4 Extracting pertinent data from tweets

After grouping of tweets, the following stage is to separate essential data from tweets, utilizing Named Entity Tagging each tweet is changed over into various substances like Numbers, Verbs, Nouns, Adjectives or Organization Name. In Need tweets we will remove asset name, area, contact number and email whenever referenced. In Availability tweets we will remove benefactor name, amount, asset, area, contact number or email. To remove contact number and email we will utilize design coordinating. Fig 3.4.1 shows how tweets take care of applying Named Entity labeling on them for extraction of data.

Source: We consider as suitable assets, two kinds of words - (I) formal people, places or things that are labeled as people or things by a Named Entity Recognizer, and (ii) formal people, places or things that are kid hubs of reliance parsing - gave they have not been distinguished beforehand as 'Particular' or 'Plural'during the check stage.

Name: In each tweet in the event that element is of type "Individual" or "Organization" that is removed as name of individual or association.

Quantity: For every asset separated, we distinguish whether it is gone before by a numeric token. The numeric token might be the orthographic documentation of a number (e.g., '100') or may semantically address a number (e.g., 'hundred'). We appoint the numeric token as the amount of the specific asset.

Contact and E-mail ID: We utilize normal articulations to recognize contacts relating to email-ids and telephone numbers.

Figure 3.4.1: Named Entity Tagging on tweets

5. CONCLUSION

We proposed a system for resource operation during a disaster situation, in this model we tried to apply different neural architecture ways and resources to help the victims of disaster. This model is truly effective in relating what type of resources are demanded and where to shoot them during extremity. The main thing of this model is to escalate the response time during disaster and increase the speed of operations. Due to pre- knowledge of number of resources demanded at position, the resource distribution operation is escalated multitudinous times. In future we want to add geo locales and break the queries raised during disaster important hastily and wanted to assign a particular time to each query. The current system allows only limited stoners to pierce platform at same time but in future we want to design a multi- stoner platform. We also want to add support for all type of languages.

6. ACKNOWLEDGEMENT

First and foremost, we would like to express our heartfelt gratitude to Prof. Ashish Mohod of Priyadarshini JL College of Engineering for providing us with excellent advice and helpful recommendations during this project.

The team members would like to express their gratitude to the Priyadarshini JL College of Engineering's Department of Computer Science and Engineering for providing the essential support and resources.

7.REFRENCES

[1] F. Abel, C. Hauff, G.-J. Houben, R. Stronkman, and K. Tao, “Semantics + Filtering + Search = Twitcident. Exploring Information in Social Web Streams,” in Proc. of the 23rd ACM Conf. on Hypertext and Social Media. ACM, 2012, pp. 285–294.

[2] U. Ahmad, A. Zahid, M. Shoaib, and A. AlAmri, “Harvis: An integrated social media content analysis framework for youtube platform,” Information Systems, vol. 69, pp. 25 – 39, 2017.

[3] G. Backfried, J. Gollner, G. Qirchmayr, K. Rainer, G. Kienast, G. Thallinger, C. Schmidt, and A. Peer, “Integration of Media Sources for Situation Analysis in the Different Phases of Disaster Management: The QuOIMA Project,” in Eur. Intel. and Security Informatics Conf., Aug 2013, pp. 143–146.

[4] BBC News Europe. (2012, Aug.) England Riots: Maps and Timeline. [Online]. Available: http://www.bbc.co.uk/news/ uk-14436499

[5] H. Becker, M. Naaman, and L. Gravano, “Learning Similarity Metrics for Event Identification in Social Media,” in Proc. of the Third ACM Int’l Conf. on Web Search and Data Mining, ser. WSDM ’10. NY, USA: ACM, 2010, pp. 291–300.

[6] J. Bezdek, T. Reichherzer, G. Lim, and Y. Attikiouzel, “Multiple Prototype Classifier Design,” IEEE Trans. on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 28, no. 1, pp. 67– 79, Feb 1998.

[7] M. Biehl, B. Hammer, and T. Villmann, “Prototype-based Models in Machine Learning,” Wiley Interdisciplinary Reviews: Cognitive Science, vol. 7, no. 2, pp. 92–111, 2016.

[8] A. Bouchachia, “Learning with Incrementality,” in Proc. of the Int’l Conf. on Neural Information Processing, 2006, pp. 137–146.

[9] A. Bouchachia, “Incremental Learning with Multi-Level Adaptation,” Neurocomputing, vol. 74, no. 11, pp. 1785–1799, 2011.

[10] A. Bouchachia and C. Vanaret, “Incremental Learning Based on Growing Gaussian Mixture Models,” in 10th Int’l Conf. on Machine Learning and Applications and Workshops (ICMLA), vol. 2, Dec 2011, pp. 47–52.

[11] A. Bouchachia etl,“GT2FC: An Online Growing Interval Type-2 Self-Learning Fuzzy Classifier,” IEEE Transactions on Fuzzy Systems, vol. 22, no. 4, pp. 999–1018, 2014.

[12] M.-R. Bouguelia, Y. Bela¨ıd, and A. Bela¨ıd, “An Adaptive Streaming Active Learning Strategy based on Instance Weighting,” Pattern Recognition Letters, vol. 70, pp. 38 – 44, 2016.

[13] M. Buscher and M. Liegl, “Connected Communities in Crises,” ¨ in Social Media Analysis for Crisis Management, H. Hellwagner, D. Pohl, and R. Kaiser, Eds. IEEE Computer Society Special Technical Community on Social Networking E-Letter, March 2014, vol. 2, no. 1.

[14] G. Cavallanti, N. Cesa-Bianchi, and C. Gentile, “Tracking the best Hyperplane with a simple Budget Perceptron,” Machine Learning, vol. 69, no. 2-3, pp. 143–167, 2007.

[15] L. Chen, K. S. M. Tozammel Hossain, P. Butler, N. Ramakrishnan, and B. A. Prakash, “Syndromic Surveillance of Flu on Twitter Using Weakly Supervised Temporal Topic Models,” Data Mining and Knowledge Discovery, vol. 0, no. 3, pp. 681–710, May 2016.

[16] T. M. Cover and J. A. Thomas, “Entropy, Relative Entropy and Mutual Information,” in Elements of Information Theory. New Jersey: A John Wiley & Sons, 2006.

[17] K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer, “Online Passive-Aggressive Algorithms,” J. Mach. Learn. Res., vol. 7, pp. 551–585, Dec. 2006.

[18] S. Dashti, L. Palen, M. P. Heris, K. M. Anderson, S. Anderson, and S. Anderson, “Supporting Disaster Reconnaissance with Social Media Data: A Design-Oriented Case Study of the 2013 Colorado Floods,” in Proc. of the 11th Int’l Conference on Information Systems for Crisis Response and Management, University Park, Pennsylvania, USA, 2014.

[19] O. Dekel, S. Shalev-Shwartz, and Y. Singer, “The Forgetron: A Kernel-Based Perceptron on a Fixed Budget,” in NIPS. MIT Press, 2005, pp. 259–266.

[20] A. Denecke, H. Wersing, J. Steil, and E. Korner, “Online figure- ¨ ground segmentation with adaptive metrics in generalized LVQ,” Neurocomputing, vol. 72, no. 7-9, pp. 1470 – 1482, 2009.

[21] S. Denef, P. S. Bayerl, and N. Kaptein, “Social Media and the Police - Tweeting Practices of British Police Forces during the August 2011 Riots,” in Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI), Paris, France, May 2013.

[22] N. Dufty, “Using Social Media to build Community Disaster Resilience,” The Australian Journal of Emergency Management, vol. 27, no. 1, pp. 40–45, 2012.

[23] M. Freeman and A. Freeman, “Bonding over Bushfires: Social Networks in Action,” in IEEE International Symposium on Technology and Society (ISTAS), June 2010, pp. 419–426.

[24] J. a. Gama, I. Zliobait ˇ e, A. Bifet, M. Pechenizkiy, and A. Bou- ˙ chachia, “A Survey on Concept Drift Adaptation,” ACM Comput. Surv., vol. 46, no. 4, pp. 44:1–44:37, 2014.

[25] B. Hammer, D. Hofmann, F.-M. Schleif, and X. Zhu, “Learning Vector Quantization for (dis-)similarities,” Neurocomputing, vol. 131, pp. 43 – 51, 2014.
[26] S. Hao, J. Lu, P. Zhao, C. Zhang, S. C. H. Hoi, and C. Miao, “

Second-Order Online Active Learning and Its Applications,” IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 7, pp. 1338–1351, July 2018.

[27] S. Hao, P. Hu, P. Zhao, S. C. H. Hoi, and C. Miao, “Online Active Learning with Expert Advice,” ACM Trans. Knowl. Discov. Data, vol. 12, no. 5, pp. 58:1–58:22, 2018.

[28] S. R. Hiltz, B. van de Walle, and M. Turoff, “The Domain of Emergency Management Information,” in Information Systems for Emergency Management. Armonk, New York: B. van de Walle, M. Truoff and S. R. Hiltz, 2010, vol. 16, pp. 3–19.
[29] D. Ienco, A. Bifet, I. Zliobait ˇ e, and B. Pfahringer, “Clustering ˙ Based Active Learning for Evolving Data Streams,” in Discovery Science, ser. Lecture Notes in Computer Science, J. Furnkranz, ¨ E. Hullermeier, and T. Higuchi, Eds. Springer Berlin Heidelberg, ¨ 2013, vol. 8140, pp. 79–93.

[30] M. Imran, C. Castillo, J. Lucas, P. Meier, and S. Vieweg, “AIDR: Artificial Intelligence for Disaster Response,” in Proc. of the Companion Publication of the 23rd Int’l Conf. on World Wide Web, ser. WWW Companion ’14, April 2014, pp. 159–162.

[31] Y. Ishikawa, Y. Chen, and H. Kitagawa, “An On-Line Document Clustering Method Based on Forgetting Factors,” in Research and Advanced Technology for Digital Libraries, ser. Lecture Notes in Computer Science, P. Constantopoulos and I. T. Solvberg, Eds. Springer, 2001, vol. 2163, pp. 325–339.

[32] T. Kohonen, “The Self-Organizing Map,” Proc. of the IEEE, vol. 78, no. 9, pp. 1464 –1480, Sep 1990.

[33] V. I. Levenshtein, “Binary Codes Capable of Correcting Deletions, Insertions and Reversals,” in Soviet physics doklady, vol. 10, 1966, p. 707.

[34] R. Li, K. H. Lei, R. Khadiwala, and K.-C. Chang, “TEDAS: A Twitter-based Event Detection and Analysis System,” in IEEE 28th Int’l Conf. on Data Engineering (ICDE), 2012, pp. 1273–1276.

[35] S. Liu, L. Palen, J. Sutton, A. Hughes, and S. Vieweg, “In Search of the Bigger Picture: The Emergent Role of On-Line Photo-Sharing in Times of Disaster,” in Proc. of the 5th Int’l ISCRAM Conf., 2008.

[36] L. Ma, S. Destercke, and Y. Wang, “Online Active Learning of Decision Trees with Evidential Data,” Pattern Recognition, vol. 52, pp. 33 – 45, 2016.

[37] C. Manning, P. Raghavan, and H. Schutze, ¨ Introduction to Information Retrieval. Cambridge University Press, 2008.

[38] S. Mohamad, A. Bouchachia, and M. Sayed-Mouchaweh, “A bicriteria active learning algorithm for dynamic data streams,” IEEE transactions on neural networks and learning systems, vol. 29, no. 1, pp. 74–86, 2018.

[39] S. Mohamad, M. Sayed-Mouchaweh, and A. Bouchachia, “Active learning for classifying data streams with unknown number of classes,” Neural Networks, vol. 98, pp. 1–15, 2018.

[40] B. Mokbel, B. Paassen, F.-M. Schleif, and B. Hammer, “Metric Learning for Sequences in Relational LVQ,” Neurocomputing, vol. 169, pp. 306 – 322, 2015.

[41] B. Mozafari, P. Sarkar, M. Franklin, M. Jordan, and S. Madden, “Scaling Up Crowd-sourcing to Very Large Datasets: A Case for Active Learning,” Proc. VLDB Endow., vol. 8, no. 2, pp. 125–136, Oct. 2014.

[42] V. K. Neppalli, C. Caragea, A. Squicciarini, A. Tapia, and S. Stehle, “Sentiment Analysis during Hurricane Sandy in Emergency Response,” International Journal of Disaster Risk Reduction, vol. 21, pp. 213 – 222, 2017.

[43] A. Olteanu, S. Vieweg, and C. Castillo, “What to Expect When the Unexpected Happens: Social Media Communications Across Crises,” In Proc. of the ACM Conf. on Computer Supported Cooperative Work and Social Computing, 2015.

[44] F. Orabona, DOGMA: a MATLAB toolbox for Online Learning, 2009, software available at http://dogma.sourceforge.net.

[45] F. Orabona, C. Castellini, B. Caputo, L. Jie, and G. Sandini, “Online Independent Support Vector Machines,” Pattern Recognition, vol. 43, no. 4, pp. 1402 – 1412, 2010.

[46] F. Orabona, J. Keshet, and B. Caputo, “Bounded Kernel-Based Online Learning,” J. Mach. Learn. Res., vol. 10, pp. 2643–2666, Dec. 2009.

[47] S.-Y. Perng, M. Buscher, L. Wood, R. Halvorsrud, M. Stiso, L. Ra- ¨ mirez, and A. Al-Akka, “Peripheral Response: Microblogging During the 22/7/2011 Norway Attacks,” Int’l Journal of Information Systems for Crisis Response and Management (IJISCRAM), vol. 5, no. 1, pp. 41–57, 2013.

[48] D. Pohl, A. Bouchachia, and H. Hellwagner, “Online Processing of Social Media Data for Emergency Management,” in Int’l Conf. on Machine Learning and Applications, vol. 2, Dec. 2013, pp. 333 – 338.

[49] D. Pohl, “Social Media Analysis for Crisis Management: A Brief Survey,” in Social Media Analysis for Crisis Management, H. Hellwagner, D. Pohl, and R. Kaiser, Eds. IEEE Computer Society Special Technical Community on Social Networking E-Letter, March 2014, vol. 2, no. 1.

[50] D. Pohl, A. Bouchachia, and H. Hellwagner, “Social Media for Crisis Management: Clustering Approaches for Sub-Event Detection,” Multimedia Tools and Applications, pp. 1–32, 2013.

[51] Daniela Pohl etl. , “Online Indexing and Clustering of Social Media Data for Emergency Management,” Elsevier Neurocomputing, vol. 172, pp. 168 – 179, 2016.

[52] J. R. Ragini, P. R. Anand, and V. Bhaskar, “Mining Crisis Information: A Strategic Approach for Detection of People at Risk through Social Media Analysis,” International Journal of Disaster Risk Reduction, vol. 27, pp. 556 – 566, 2018.

[53] C. Reuter and M. Kaufhold, “Fifteen Years of Social Media in Emergencies: A Retrospective Review and Future Directions for Crisis Informatics,” Journal of Contingencies and Crisis Management, vol. 26, no. 1, pp. 41–57, 2018.

[54] T. Reuter and P. Cimiano, “Event-based Classification of Social Media Streams,” in Proc. of the 2nd ACM Int’l Conf. on Multimedia Retrieval, 2012, pp. 22:1–22:8.

[55] T. Reuter, P. Cimiano, L. Drumond, K. Buza, and L. SchmidtThieme, “Scalable Event-Based Clustering of Social Media via Record Linkage Techniques,” in The 5th Int’l Conf. on Weblogs and Social Media, 2011, pp. 313–320.

[56] A. Rosenberg and J. Hirschberg, “V-Measure: A Conditional Entropy-Based External Cluster Evaluation Measure,” in EMNLP-CoNLL, vol. 7, 2007, pp. 410–420.

[57] B. Settles, “Active Learning Literature Survey,” University of Wisconsin, Madison, vol. 52, pp. 55–66, 2010.
[58] E. Shook, K. Leetaru, G. Cao, A. Padmanabhan, and S. Wang, “Happy or Not: Generating Topic-based Emotional Heatmaps for Culturomics using CyberGIS,” in 2012 IEEE 8th Int’l Conf. on E-Science, oct. 2012, pp. 1 –6.

[59] J. Smailovic, M. Gr ´ car, N. Lavra ˇ c, and M. ˇ Znidar ˇ si ˇ c, “Stream- ˇ based Active Learning for Sentiment Analysis in the Financial Domain (in press),” Information Sciences, April 2014.

[60] K. Starbird and J. Stamberger, “Tweak the Tweet: Leveraging Microblogging Proliferation with a Prescriptive Syntax to Support Citizen Reporting,” in Proc. of the 7th Int’l ISCRAM Conf., Seattle, USA, May 2010.

[61] T. Terpstra, A. de Vries, R. Stronkman, and G. L. Paradies, “Towards a Realtime Twitter Analysis during Crises for Operational Crisis Management,” in Proc. of the 9th Int’l ISCRAM Conf., Vancouver, April 2012.

[62] M. F. Umer and M. S. H. Khiyal, “Classification of Textual Documents using Learning Vector Quantization,” Information Technology Journal, vol. 6, no. 1, pp. 154–159, 2007.

[63] S. Vieweg and A. Hodges, “Rethinking Context: Leveraging Human and Machine Computation in Disaster Response,” Computer, vol. 47, no. 4, pp. 22–27, Apr 2014.

[64] S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen, “Microblogging During Two Natural Hazards Events: What Twitter May Contribute to Situational Awareness,” in Proc. of the Int’l Conf. on Human Factors in Computing Systems, ser. CHI ’10. NY, USA: ACM, 2010, pp. 1079–1088.

[65] I. Zliobait ˇ e, A. Bifet, B. Pfahringer, and G. Holmes, “Active ˙ Learning with Drifting Streaming Data,” IEEE Trans. on Neural Networks and Learning Sys., vol. 25, no. 1, pp. 27–39, Jan 2014.

[66] I. H. Witten, E. Frank, and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques. Elsevier, 2011.

[67] J. Yin, A. Lampert, M. Cameron, B. Robinson, and R. Power,, “Using Social Media to Enhance Emergency Situation Awareness ,” IEEE Intelligent Sys., vol. 27, no. 6, pp. 52–59, 2012.

[68] D. Pohl, A. Bouchachia and H. Hellwagner , “Active Online Learning for Social Media Analysis to support Crisis Management” , IEEE Trans. Knowledge and Data Engineering vol. 32 no 8 pp.1-14 2019