SVM-Based Domain Adaptation Machine Learned Models for the Automatic Classification of Disaster-Related Tweets

Beverly Estephany Parilla-Ferrer, Patrick Carl A. Austria, Benedict Lorenzo F. Bueno, Victor Nathaniel S. Rabara, Remigio B. Ramirez, Romina Annie C. Rea

***Abstract*—**Twitter is a well-known microblogging service used by many users to quickly acquire disaster-related information. Since there is an influx of tweet during the occurrence of a disaster event, the need to automatically filter informative disaster-related tweets is imperative. In this study, the researchers used the Habagat tweets and tweets gathered during the Zamboanga Siege event as training corpora for training two Support Vector Machine (SVM) machine classifiers. The first machine classifier is used to filter informative tweets while the second model is used to classify informative tweets as Donation, Traffic update, Weather update, Class suspension, Rescue and Relief and River level/condition. These subclasses were based on the latent topics extracted from the corpus of informative tweets using Principal Component Analysis(PCA) and Latent Dirichlet Allocation(LDA). Several experiments using different preprocessing techniques, vector space of unigrams/bigrams with TF-IDF weightings on balanced and unbalanced data were applied to test and train the classifiers. SVM and Expectation Maximization (EM) algorithms were used for the training and testing of the machine classifiers. The models were evaluated based on the performance metrics of accuracy, precision and recall. Experimental results show that for a binary classifier, using unigram with a corpus of unbalanced labeled corpus returned best classification results and applying bigram to a balanced corpus gave high performance results. Furthermore, the multiclass model yielded better performance results with unigrams as features**.**

Index Terms—Disaster, Domain Adaptation, SVM, Expectation Maximization, Machine Learning, Tweets

I. Introduction

Social media has been used by the masses to disseminate information throughout the globe [22] and also served as a backchannel for communicating information [25] besides traditional media. These social media services play a significant impact on the life of millions of users [3]. A subset of social media known as microblogging allows followers to continuously learn what a followed account holder is thinking even if they have never met the person in real life [22],[30].

Twitter is a microblogging site that allows its subscribers to broadcast information feeds, known as a tweet, which is limited to 140 characters [18]. It has emerged as a platform for providing information, particularly during disasters [7],[8],[15],[32]. It generates various updates of useful information that can lead to a reliable disaster response mechanism, which can result to precautionary measures, early preparations and reduce damages to property or loss of lives before, during and after the disaster. During the events of disasters, broadcasted tweets are rapid and continuous that resorting to manual classification of tweets may seem impossible since humans can only do a limited amount of tasks at a time [3],[28]. Thus, there is a need to develop an automatic classifier that classifies disaster-related tweets into either informative that give relevant information about a current situation in a specific area affected by the disaster, or uninformative which may just be expressions of emotions or opinions which do not give helpful information[30]. The necessity to automatically classify and filter informative tweets will contribute to a disaster response mechanism that can help first responders and concerned citizens hence, the minimization of damage to property and loss of lives during a disaster.

Several interests in research take advantage of the data being provided by disaster-related tweets. On the area of mining, classifying and predicting disaster-related tweets, one of the learning algorithms used to train an automatic classifier is Support Vector Machine (SVM). Classifier models produced by SVM give good classification results with high accuracy and robustness[2]. Although there are several studies involving the evaluation SVM-based classifiers that filter and predict disaster-related tweets, all of them only dealt with training and testing classifiers on a single-source domain. The work of [7] mentioned that a classifier trained using SVM and employing domain adaptation on disaster-related tweets is not yet presented. This motivated the researchers to develop an SVM-trained model focused on binary classification of tweets as informative or uninformative from datasets of different sources to employ domain adaptation. Another objective was to determine the latent topics that can be found in the informative tweets corpus using Principal Component Analysis (PCA) and Latent Dirichlet Allocation (LDA). Based on the latent themes, a second level model to sub-classify the informative tweets corpus was developed.

II. Related Works

The effectiveness of Twitter as an information source attracted the attention of many researchers and had studies published on various disciplines such as engineering, information technology, education, etc. The following sub-sections mentioned several related research works on the classification of text and prediction on various domains,

Manuscript received November 13, 2016.

B.E.P Ferrer is a faculty in the Information Technology/ Computer Science Department of the School of Computing and Information Sciences of Saint Louis University Baguio City, Philippines 2600

P.C.A Austria, B.L.F Bueno, V.N.S Rabara, R.B Ramirez, R.A.C Rea are students of the Information Technology/Computer Science Department of the School of Computing and Information Sciences of Saint Louis University Baguio City, Philippines 2600

proposed applications and architectures, and the implementation of domain adaptation on text classification.

2.1 Text Classification

Arguments from [21] says that SVM acknowledges properties of text in terms of classification and is therefore well-suited for text classification problems. There are studies that deal with creating automatic classifiers using SVM. Study of [7] classified a corpus of tweets mined from the earthquake in Haiti in 2010 using SVM classifiers. The researchers of [30] used a learned SVM model trained from manually labeled datasets as either relevant or irrelevant. Three automatic tweet classifier model using Naïve Bayes, Maximum Entropy and SVM learning algorithm with distant supervision was developed by [12]. The work of [16] used a Naïve Bayes classifier in Weka to classify tweets using the Joplin tornado dataset. In addition, another work by [17] focused on the training of classifiers using Conditional Random Fields (CRF) for the classification of disaster-related information from the Joplin tornado and Hurricane Sandy datasets. A related study on classifying tweets that contribute to situational awareness(SA) was probed by [33].

2.2 Twitter Applications and Architectures

The work of [7] developed the iPhone application Enhanced Messaging for the Emergency Response Sector (EMERSE) that classifies tweets and text messages regarding Haiti Disaster Relief operations for relief workers and the people of Haiti. A sample of 2,116 text messages were taken from the Ushahidi Text Message Corpus. Five-fold cross validation with feature representations like Bag-of-words, Feature abstraction, Relief Feature Selection (FS) and LDA, together with keywords and SVM were used to train and test the classifier. Experimental results show that the Feature Abstraction outperforms the other feature representations in terms of F-measure. A system architecture employing Natural Language Processing (NLP) and data mining techniques for extracting information from tweets during disasters was developed by [34]. Feature extraction consisted of word unigrams, word bigrams, word length, hashtag count, number of username mentions, retweets and replies. The tweets were classified using two machine learning algorithms: Naïve Bayes and Support Vector Machines (SVM). The classifiers were trained and tested using 10-fold cross validation on the training data, with classification algorithms Naïve Bayes and SVM achieved an accuracy of 86.2% and 87.5% respectively. The application TweetTracker by [23] analyzed the keywords and the location of a tweet to increase the situational awareness of first responders during a disaster. A data-mining tool named Tweedr that extracts relevant disaster-related tweets for relief workers was developed by [1]. The tweets were mined with the usage of keyword and geographical queries from the social media aggregator Gnip. A sample of 1,073 tweets were annotated manually employing two tasks which were binary classification and disaster-related token extraction resulting to a total of 793 positive examples. Several classification algorithms were used in the experiment specifically Naïve Bayes, Logistic Regression, Decision Trees, Supervised Latent Dirichlet Allocation (sLDA) and K-Nearest Neighbors. Experimental results in 10-fold cross validation showed that Logistic Regression is the most reliable across measures F-Score of 65%, accuracy of 86%, and AUC at 88%.

2.3 Domain Adaptation on Text Classification

On the area of domain adaptation being employed in text classification, the following sub-sections are focused on the sentiment analysis and text classification.

2.3.1 Sentiment Analysis

Domain adaptation to perform sentiment classification of tweets was performed in the study of [29]. Given a source dataset from human labeled tweets, in addition to blips from Blipper and IMDB reviews as target labeled data, they used Feedback EM and Rocchio SVM to improve the classifier for the target. The work of [26] used Lingpipe classifier from a mined data collection of topical and sentiment documents specifically from three streams namely Reviews, Blogs, and Twitter for each of five categories, which are games, movies, music, phones, and restaurants. The results show that combining the training data with other streams further boosts performance, while combining training data from different categories produced classifiers outperforming their native counterparts. Cross-stream experiments involved Single-Source Model Adaptation and further explored other cross-stream adaptation experiments like Two-Source Mixed Model, Three-Source Mixed Model, and Three-Source Voting Model.

2.3.2 Web Documents and Tweet Classification

The study of [13] explored domain adaptation on disaster-related tweets by labeling tweets from the Boston Marathon Bombings as target corpus while using labeled data from the Hurricane Sandy as source corpus. Two approaches were performed. The first approach utilized the Naïve Bayes algorithm to train and test the classifier using the source data. The second approach utilized domain adaptation Naïve Bayes to learn classifiers from labeled source data and unlabeled target data. A proposed domain adaptation algorithm used Naïve Bayes and Expectation-Maximization (EM) for categorizing text documents was presented[9]. Deep Learning system based on Stacked De-noising Auto-Encoders (SDA) that perform unsupervised feature extraction for domain adaptation of sentiment classifiers was conducted[11]. SDA's performance evaluation was compared with Structural Correspondence Learning introduced by [5], Multi-label Consensus Training (MCT) by [24] and Spectral Feature Alignment (SFA) by [27]. Experimental results show that SDA transfer outperforms all others by 11 out of 12 cases applied on an Amazon dataset consisting of four domains. The study of [10] introduced a framework called Domain Adaptation Machine (DAM) for performing domain adaptation on multiple sources. Two domain adaptation methods referred as FastDAM and UniverDAM were introduced. FastDAM uses LS-SVM sparse target classifier for a fast label prediction while UniverDAM makes use of source domains for the generalization ability of the classifier. In the study of [16], the researchers emphasized domain adaptation for the identification of information nuggets using conditional random fields (CRF). The classifiers used training data from the Joplin 2011 tornado as source and a 10% portion of training data from

Hurricane Sandy as target. Experiment results showed that using only the source data yield a significant drop in the detection rate, while significantly not affecting recall.

III. Methodology

A. Data Gathering and Preparation

The researchers gathered tweets from Twitter using a program written in Python utilizing the Tweepy library. The collected corpora is composed of 3,000 tweets from a man-made disaster known as the Zamboanga Siege of 2013 and 1,500 tweets coming from non-disaster related events. These corpora together with the 665,176 tweets from the Habagat flooding of Metro Manila on 2012 composed the whole corpus used in our experiments.

Random samples with sizes of 10,000 and 1,507 were selected for manual annotation from the Habagat and Zamboanga corpora respectively. The selected tweets were processed using an electronic spreadsheet software and after removing the occurrence of duplicates in each sample, the researchers were left with 4,191 Habagat tweets and 1,507 Zamboanga Siege tweets.

Manual annotation for both the Habagat and Zamboanga corpora were done by three (3) annotators. The annotators were tasked to label the tweets as informative or uninformative. In cases where the annotators did not agree on the label of a tweet, discussion and a process of voting was done to assign the final label for the tweet. Fleiss Kappa was computed to determine reliability between the agreements of the annotators.

After the process of manual annotation, it was observed that the raw form of the tweets contained noises in the forms of URLs, twitter usernames, the "RT" term and special characters. In order to remove the said noises, the corpora were fed into a java application developed by the researchers to undergo noise reduction processes. Afterwards, the corpora underwent a process of duplicates removal and had the resulting sizes of 3,792 and 1,507 for the Habagat and Zamboanga corpora respectively. The final corpora were composed of both original tweets and retweets without duplicates.

B. Data Preprocessing

Preprocessing techniques such case transformation, tokenization, stemming, generation of term sequences (n-gram), bag-of-words and term frequency – inverse frequency (TF-IDF) to create the vector space. Case transformation is the process of changing a character into either its uppercase or, in this study, lowercase form. Tokenization is breaking a text into words which are called as tokens. Stemming is the process of removing a word's suffix to get its basic or root form. Generation of n-gram is generating a set of co-occurring words from a given a text. Bag-of-words takes all the words in a corpus and represents them as a bag of elements which will be used as the main feature set. TF-IDF is a weighting scheme which assigns weights to each element of the feature set based on their frequencies on the whole corpus.

C. Classifier Training and Testing

The two SVM based classifiers produced by this research were trained and tested using 10-fold cross validation. In a 10-fold cross validation, the corpus is divided into ten exclusive subsets which approximately have the same sizes. Nine-folds will serve as training corpora and the remaining 1-fold will be used as testing data [30]. The training and testing were done ten times until each of the ten folds had been used to train and test the model.

D. Informative Tweets Themes Discovery

All the informative tweets from the Habagat corpus which totaled to 1,245 were collected to discover the possible themes present in them. To find the topics present in the informative tweets corpus, the researchers used two algorithms namely Principal Component Analysis (PCA) and Latent Dirichlet Allocation (LDA). PCA is a feature reduction algorithm that groups together similar attributes into a single group known as principal components. On the other hand, LDA is a topic modeling algorithm that finds the latent or hidden topics present in a corpus and outputs the keywords that identifies each observed topic.

E. Domain Adaptation Approach

The setting of domain adaptation considers two different but related corpora, a source domain D_S provides training data for the classifier model and a target domain D_T that gives test data instances [11]. In this study, the Habagat corpus was used as D_S to train the classifier and the Zamboanga corpus as D_T where the learned model was tested. All tweets from D_S were labeled but the tweets from D_T were all unlabeled. The labeling setup for both the corpora from D_S and D_T were made to address the fact that in real situations when a disaster occurs there are no readily available target labeled data, thus the classifier should be able to categorize based on what it had learned from D_S and apply it to the incoming target tweets.

The researchers used EM to implement domain adaptation. EM is an iterative algorithm for estimating the maximum likelihood or maximum a posteriori (MAP) values of parameters in statistical models which rely on unlabeled data. EM is divided into two processes, the expectation (E) step which determines the possible distribution of the unobserved data using the current parameter settings, and the maximization (M) step where the values of the parameters are re-estimated based on the results of the E step. The parameter estimates from the M step will then be used in next E step. EM was used in this study to determine the labels of tweets coming from D_T. The tweets labeled by EM were then added to the training corpus to retrain and re-estimate the parameters of the classifier.

Two approaches were done to implement EM in the study. The first approach was with the use of EM clustering which takes a corpora and groups it into *k* clusters. The second approach was done by creating a custom implementation of the EM algorithm. First, a classifier was trained using the corpus from D_S. The learned model was then used to perform the E-step which was to determine the labels of the tweets coming from D_T. All classifications that had a classification confidence value of 0.9 and above were removed from the D_T and added to D_S. The M-step made use of the modified D_S to retrain and re-estimate the parameters of the classifier. This EM process was repeated until a specified number of iteration or until D_T became empty.

F. *Classification Performance Measures*

The performance measures that were used in the study for the evaluation of the Support Vector Machine classification algorithm were accuracy, precision, and recall.

IV. RESULTS AND DISCUSSION

The discussions for the results are divided into three subsections which will give detailed explanations about the discovery of informative tweets themes, EM clustering experiments and custom EM implementation results. All corpora mentioned in the next subsections came from same sources, the labeled corpora were from the Habagat corpus and all unlabeled corpora were combinations of tweets from Zamboanga Siege and non-disaster related events.

4.1 EM Clustering Approach

The target domain Zamboanga corpus consisting of 1,507 tweets after undergoing preprocessing with unigram for term sequence generation created a vector space with a feature set of size 3,568. The vector space was used by the EM clustering algorithm to subdivide the whole corpus into two clusters.

Using the preprocessing settings, corpora of sizes 10, 20, 40, 50 and 60 were prepared and fed into the EM clustering algorithm. Each of the corpora produced different vector spaces with feature set dimensionality having a direct positive correlation to the size of the corresponding corpus. Experimental results showed that it only took a matter of seconds or minutes to finish the process if the corpus has a size of 50 and below. When the corpus of size 60 was used, the process already took hours if not days. It was observed that the number of attributes or the size of the feature set was the factor that determined the runtime of the EM clustering algorithm. As the size of the feature set increase, the runtime of EM clustering increased by exponential factors.

To give a solution with the issue regarding the high dimensionality of the feature set generated by applying bag-of-words the resulted to long runtime of the EM clustering, the researchers made a list which was used to be the feature set of the vector space. The list was made by first getting the frequency of each word in the whole Habagat corpus. A word was selected to be part of the keyword list by balancing between the word's frequency count and its relevance to identifying an informative tweet. The final list was composed of 503 unique words.

TABLE 1

. LEVEL 1 SVM CLASSIFIER PERFORMANCE WITH 503 KEYWORDS USED AS A FEATURE LIST

Corpora		Term Sequencing	Accuracy	Precision	Recall
Labeled Habagat Corpus	Unlabeled Zamboanga Corpus				
3792	1507	Unigram	80.99%	73.98%	82.58%

Class Recall (Informative): 52.59%
Class Recall (Uninformative): 95.38%

Table 1 shows the performance of the learned model. Since the size of the feature set was reduced to 503 from an original of 3,568, the trained model had lesser factors to base its decisions which explains why its performance have rather low values. In addition, the value of the class recall for the uninformative tweets being 42% higher than that of the informative tweets class recall was due to the fact that the training corpora was unbalanced. Informative tweets only occupied a third of the whole training corpora while uninformative tweets took the remaining two-thirds.

4.1 Custom EM Implementation

The runtime for the experiments done with the custom EM implementation only took several minutes to a couple of hours. Thus many experimental setups were done using this approach. The following subsections present the various experimental set-ups used to train the classifiers.

4.2.1 Level 1 Classifier

Two experimental approaches were used to train the first-level binary classifier under the custom implementation of EM. The first approach was the use of unbalanced corpora that had a size of 3,792 and was composed of 1,245 informative tweets and 2,547 uninformative tweets. Three sets of unlabeled corpora were prepared, with sizes of 1000, 2000, and 3000.

For the second approach, the training corpora was composed of 2490 elements with equal ratio between informative and uninformative tweets partnered with an unlabeled corpora of size 1507.

Each set of unlabeled data was used in different setups and had the main function of being the testing corpora. The preprocessing steps were the same for all the experiments except for the generation of term sequences where the researchers used unigram and bigram for every pair of training and testing corpora.

Table 2 shows the performances of a trained classifier using three different corpora with unbalanced labeled data. Results show that when unigram is applied during the preprocessing step, the classifier performs a little better than using bigram. Table 3 shows a different scenario when the training corpus is composed of balanced labeled tweets. In contrast with the results in Table 2, bigram makes the performance of the classifier better when applied during preprocessing. The highlighted rows in Tables 2, 3 and 4 show the experimental set-ups that yielded best results.

Between the results of using unbalanced labeled corpora, the former gave better performance values. This was due to the fact that the developed model was able to classify more uninformative tweets due to the greater proportion of uninformative tweets in the corpora.

TABLE II
UNBALANCED LABELED DATA EXPERIMENTS

Labeled Habagat Corpus	Corpora Unlabeled Zamboanga Corpus + Non-disaster-related corpus	Term Sequencing	Iteration	Accuracy	Precision	Recall
3792	1000	Unigram	15	89.56%	87.19%	88.73%
3792	1000	Bigram	15	89.15%	86.04%	88.81%
3792	2000	Unigram	15	89.07%	86.23%	88.40%
3792	2000	Bigram	15	89.04%	85.53%	88.87%
3792	3000	Unigram	15	90.12%	87.80%	89.14%
3792	3000	Bigram	15	89.49%	86.48%	88.89%

TABLE III
BALANCED LABEL DATA EXPERIMENTS

Corpora						
Labeled Habagat Corpus	Unlabeled Zamboanga Corpus	Term Sequencing	Iteration	Accuracy	Precision	Recall
2490	1507	Unigram	15	87.54%	87.50%	87.59%
2490	1507	Bigram	15	88.81%	88.70%	88.87%

TABLE IV
UNBALANCED LABELED DATA WITH 5 THEMES

Corpora		**Term Sequencing**	**Iteration**	**Accuracy**	**Precision**	**Recall**
Labeled Habagat Corpus	*Unlabeled Zamboanga Corpus*					
1083	**100**	**Unigram**	**15**	**94.55%**	**96.73%**	**91.97%**
1083	100	Bigram	15	92.43%	96.27%	87.32%

4.2.2 Latent Themes of Informative Tweets

To proceed with training the Level 2 classifier, the researchers applied both PCA and LDA to the informative Habagat tweets to find the classes for sub- classification. PCA was ran with its parameter set to fixed number having an arbitrary value of 6 and further investigation on the themes was performed. The corpus undergone the same preprocessing steps performed in Section 3.B, with unigrams as the parameter for generating term sequences. LDA was performed using a java application known as Mallet and was set to find 6 latent topics in the corpus.

Results of PCA and LDA returned similar themes for the informative tweets. Upon manual inspection, the themes "Donation" and "Rescue" were found compatible to merge as one subtopic, which the researchers named as Rescue and Relief. The final themes set composed of five elements namely Traffic Update, Weather Update, Rescue and Relief, Class Suspension and River water level.

4.2.3 Level 2 Classifier

In training the second-level multi-class classifier, LibSVM was applied. The labeled corpus was composed of1083 tweets and the unlabelled corpora had a size of 100. The corpus of informative tweets was classified as Class Suspension, Rescue and Relief, Traffic Update, Weather Update and Water Level Condition.

As shown in Table 4, the trained level 2 model for classifying informative tweets into the five categories yielded higher performance values with unigrams as features. This result may be attributed to the specific terms used by Twitter subscribers for a specific class. There are terms which are highly associated to specific classes, e.g. for Rescue and relief – Help, need ; Traffic update – passable; Suspension of classes – no, classes; Weather update – signal, typhoon; Marikina River – level, high. In addition, the application of the EM algorithm may have an effect with the performance results of the trained level 2 classifiers. The iterative step of EM may have benefited the experimental setting utilizing unigram more than the one with bigram. Several experiments which focus on different configurations of EM and rearranging the order of the preprocessing techniques will be performed in the future to verify the said results.

V. CONCLUSION

In this study, tweets collected during the Habagat monsoon and Zamboanga Siege were used to train two SVM classifiers and discover latent topics present in the informative tweets corpus. The first level classifier was created by applying preprocessing techniques to the corpora and then feeding the generated vector space to SVM and EM for the training. Experimental results show that using a combination of unbalanced labeled corpus and a number of unlabeled corpus with unigram gave best performance values for the first classifier. The classes used for the training of the second classifier were discovered by using PCA and LDA. The results showed 2 out of the 6 identified themes were compatible to be merged, thus giving a final theme set with five elements namely Traffic Update, Weather Update, Relief, Class Suspension, and Marikina River. The second level multiclass classifier was trained using the same methods applied on the first classifier. It was shown that the second classifier performs best with unigrams as attributes.

The researchers recommend on gathering more tweets for the target domain to increase the amount of informative tweets that could be used for the domain adaptation approach of level 2 classifier training. In addition, more experiments involving corpora from different disaster related events could be done to test the versatility of the trained classifiers.

REFERENCES

[1] Ashktorab, Z., Brown, C., Nandi, M., & Culotta, A. (2014). Tweedr: Mining twitter to inform disaster response. Proc. of ISCRAM.

[2] Auria, L., & Moro, R. A. (2008). Support vector machines (SVM) as a technique for solvency analysis.

[3] Batool, R., Khattak, A. M., Maqbool, J., & Lee, S. (2013, June). Precise Tweet

[4] Classification and Sentiment Analysis. Computer and Information Science (ICIS), 2013 IEEE/ACIS 12th International Conference (pp. 461-466). IEEE.

[5] Blitzer, J., McDonald, R., & Pereira, F. (2006, July). Domain adaptation with structural correspondence learning. In Proceedings of the 2006 conference on empirical methods in natural language processing (pp. 120-128). Association for Computational Linguistics. https://doi.org/10.3115/1610075.1610094

[6] Blitzer, J., Dredze, M., & Pereira, F. (2007, June). Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In ACL (Vol. 7, pp. 440-447).

[7] Caragea, C., McNeese, N., Jaiswal, A., Traylor, G., Kim, H. W., Mitra, P., & Yen, J. (2011). Classifying text messages for the Haiti earthquake. In Information Systems for Crisis Response and Management, ISCRAM.

[8] Cobo, A., Parra, D., & Navón, J. (2015, May). Identifying Relevant Messages in a Twitter-based Citizen Channel for Natural Disaster Situations. In Proceedings of the 24th International Conference on World Wide Web Companion (pp. 1189-1194). International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/2740908.2741719

[9] Dai, W., Xue, G. R., Yang, Q., & Yu, Y. (2007, July). Transferring naive bayes classifiers for text classification. In Proceedings of the national conference on artificial intelligence (Vol. 22, No. 1, p. 540). Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

[10] Duan, L., Xu, D., & Tsang, I. W. (2012). Domain adaptation from multiple sources: A domain-dependent regularization approach. Neural Networks and Learning Systems, IEEE Transactions on, 23(3), 504-518. https://doi.org/10.1109/TNNLS.2011.2178556

[11] Glorot, X., Bordes, A., & Bengio, Y. (2011). Domain adaptation for large-scale sentiment classification: A deep learning approach. In Proceedings of the 28th International Conference on Machine Learning (ICML-11) (pp. 513-520).

[12] Go, A., Bhayani, R., & Huang, L. (2009). Twitter Sentiment Classification using Distant Supervision. CS224N Project Report, Stanford, 1, 12.

[13] Guevara, H. L., Caragea, N. H., Cornelia, C. K., & Tapia, A. S. (2015). Twitter Mining for Disaster Response: A Domain Adaptation Approach. ISCRAM 2015 CONFERERENCE.

[14] Huang, Q., & Xiao, Y. (2015). Geographic situational awareness: mining tweets for disaster preparedness, emergency response, impact, and recovery. ISPRS International Journal of Geo-Information, 4(3), 1549-1568. https://doi.org/10.3390/ijgi4031549

[15] Hughes, A. L., & Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. International Journal of Emergency Management, 6(3-4), 248-260. https://doi.org/10.1504/IJEM.2009.031564

[16] Imran, M., Elbassuoni, S. M., Castillo, C., Diaz, F., & Meier, P. (2013a). Extracting information nuggets from disaster-related messages in social media. Proc. of ISCRAM, Baden-Baden, Germany

[17] Imran, M., Elbassuoni, S., Castillo, C., Diaz, F., & Meier, P. (2013b). Practical extraction of disaster-relevant information from social media. In Proceedings of the 22nd international conference on World Wide Web companion (pp. 1021-1024). International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/2487788.2488109

[18] Jiang, L., Yu, M., Zhou, M., Liu, X., & Zhao, T. (2011). Target-dependent Twitter

[19] Sentiment Classification. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies Volume 1 (pp. 151-160). Association for Computational Linguistics.

[20] Jivani, A. G. (2011). A comparative study of stemming algorithms. Int. J. Comp. Tech. Appl, 2(6), 1930-1938.

[21] Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features (pp. 137-142). Springer Berlin Heidelberg. https://doi.org/10.1007/bfb0026683

[22] Kaigo, M. (2012). Social media usage during disasters and social capital: Twitter and the Great East Japan earthquake. Keio Communication Review, 34, 19-35.

[23] Kumar, S., Barbier, G., Abbasi, M. A., & Liu, H. (2011, July). TweetTracker: An Analysis Tool for Humanitarian and Disaster Relief. In ICWSM.

[24] Li, S., & Zong, C. (2008, October). Multi-domain adaptation for sentiment classification: Using multiple classifier combining methods. In Natural Language Processing and Knowledge Engineering, 2008. NLP-KE'08. International Conference on (pp. 1-8). IEEE.

[25] McCarthy, J. F., & Boyd, D.M. (2005, April). Digital backchannels in shared physical spaces: experiences at an academic conference. In CHI'05 extended abstracts on Human factors in computing systems (pp. 1641-1644). ACM. https://doi.org/10.1145/1056808.1056986

[26] Mejova, Y., & Srinivasan, P. (2012, May). Crossing Media Streams with Sentiment: Domain Adaptation in Blogs, Reviews and Twitter. In ICWSM.

[27] Pan, S. J., Ni, X., Sun, J. T., Yang, Q., & Chen, Z. (2010, April). Cross-domain sentiment classification via spectral feature alignment. In Proceedings of the 19th international conference on World Wide Web (pp. 751-760). ACM. https://doi.org/10.1145/1772690.1772767

[28] Panem, S., Gupta, M., & Varma, V. (2014, November). Structured Information Extraction from Natural Disaster Events on Twitter. Proceedings of the 5th International Workshop on Web-scale Knowledge Representation Retrieval & Reasoning (pp. 1-8). ACM. https://doi.org/10.1145/2663792.2663794

[29] Peddinti, V. M. K., & Chintalapoodi, P. (2011, August). Domain adaptation in sentiment analysis of twitter. In Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence.

[30] Prasetyo, P. K., Lo, D., Achananuparp, P., Tian, Y., & Lim, E. P. (2012, September). Automatic classification of software related microblogs. In Software Maintenance (ICSM), 2012 28th IEEE International Conference on (pp. 596-599). IEEE. https://doi.org/10.1109/icsm.2012.6405330

[31] Starbird, K., Palen, L., Hughes, A. L., & Vieweg, S. (2010, February). Chatter on the red: what hazards threat reveals about the social life of microblogged information. In Proceedings of the 2010 ACM conference on Computer supported cooperative work (pp. 241-250). ACM. https://doi.org/10.1145/1718918.1718965

[32] Sundharivel, K., & Rajkumar, D. U. S. (2014). Real Time Nature Disasters of Data Analysis Using Tweet. Journal of Advanced Engineering Research, 1(1).

[33] Verma, S., Vieweg, S., Corvey, W. J., Palen, L., Martin, J. H., Palmer, M., & Anderson, K. M. (2011, July). Natural Language Processing to the Rescue? Extracting" Situational Awareness" Tweets during Mass Emergency. In ICWSM.

[34] Yin, J., Lampert, A., Cameron, M., Robinson, B., & Power, R. (2012). Using social media to enhance emergency situation awareness. IEEE Intelligent Systems, 27(6), 52-59. https://doi.org/10.1109/MIS.2012.6

Beverly Estephany Parilla-Ferrer is an Electronics & Communications Engineer. She is an Electronics & Communications Engineer and a graduate of Masters of Science in Information Technology from Saint Louis University, Baguio City, Philippines. She is a candidate for the degree of Doctor of Information Technology of the University of Cordilleras**,** Gov. Pack Road, Baguio City, Philippines.

Co-authors: Patrick Carl C. Austria, Benedict Lorenzo F. Bueno, Romina Annie C. Rea, Victor Nathaniel Rabara & Remigio Ramirez are Senior Computer Science students of the School of Computing and Information Sciences of Saint Louis University, Baguio City,,Philippines.