Analysis of Sentiments amongst Tweets Concerning Disasters-A Deep Learning Approach

Venugopal.S
School of Mathematics and Computing Sciences,
Department of Computer Science
Rani Channamma University
Belagavi-591156, Karnataka, India
jalihal@gmail.com

Siddu.P.Algur
School of Mathematics and Computing Sciences,
Department of Computer Science
Rani Channamma University,
Belagavi-591156, Karnataka, India
siddu_p_algur@hotmail.com

Abstract:
Finding the most prevalent and pertinent user opinions on a certain topic is key to the effectiveness of sentimental analysis. In this study, a framework is created combining data pre-processing techniques, a blend of machine learning, statistical modelling, and deep learning methodologies, to assess user sentiments on Twitter regarding natural catastrophes. When using text data for natural language processing, features can be retrieved piecemeal and compared, but without taking context or the entire sentence into account, the sentiment may be misunderstood. This research suggests utilising a combination of deep learning and machine learning models to evaluate pre-labelled Twitter feeds on disasters to understand the feelings surrounding a certain tragedy. In our study, we glean information from tweets about various disasters that reflect the emotions expressed during the tragic events as reported by Twitteratti. To use machine learning and deep learning algorithms, data is cleaned, purified, and optimised. We provide several machine learning and deep learning models and assess their performances on many metrics in addition to exploratory data analysis.

Keywords: Disaster analysis, Deep learning, Machine Learning, Exploratory data analysis

I. INTRODUCTION

Twitter is one of the most well-known microblogging platforms, allowing users to post and receive 140-character messages known as tweets. It provides resources for conducting empirical research on the traits of interpersonal interactions. Thoughts on a variety of topics, such as politics, the influence of brands, disasters, elections, etc., may be found in information obtained from Twitter. The introduction of machine learning methodologies enabled decision-makers to ensure efficient solutions to a variety of problems. One technique for sentiment analysis on Twitter is machine learning.Machine learning is useful for resolving NLP (Natural Language Processing) problems since it uses general learning algorithms and a sizable sample of data to learn the categorization rules. Natural disasters harm both the built environment and the populace. Social media is a common place for people to communicate their feelings about such occurrences, which allows academics to gauge the general public's feelings.Due to their crowded population and highly interconnected infrastructure, urban areas are particularly susceptible to natural calamities. Big data has the potential to change how we approach vulnerability analysis, early warning systems, monitoring, and evaluation of catastrophe risk reduction. Research on sentiment analysis and human movement has already made use of the vast amounts of crowd-sourced data obtained from Twitter. These two study fields allow disaster managers to make decisions from the ground up and play more important roles in disaster aid. It has been shown that sentiment analysis of brief social media posts is a practical method for determining the dynamic polarity of catastrophe-related sentiments, improving decision-making about resource aid, humanitarian activities, and disaster recovery, and gathering particular information. Sentiment analysis of brief social media postings has grown to be a significant part of disaster relief and urban resilience. Sentiment analysis of brief social media posts has become more and more relevant. It is a helpful tool for disaster management to acquire precise information, determine the dynamic polarity of attitudes throughout a disaster, and make better decisions on resource demands, support, and humanitarian initiatives. Finding answers to pervasive problems requires accurate analysis. To better understand public feelings regarding large data and disasters, machine learning and deep learning algorithms play a crucial role in finding, analysing, and evaluating public sentiments. As a result, successful models are propagated with perceptual technologies. We make an effort to implement several deep learning and machine learning models and assess their capabilities. The next level involves using algorithms for machine learning and deep learning coupled with functions like SIGMOID, ADAM, and others to produce outcomes that are even more honed. The most suitable deep learning models for sentiment analysis of disaster data are evaluated.

II. RELATED WORK

In contemporary enterprises, machine learning and deep learning models are pervasive. With the quick development of new algorithms, less expensive computing, and more readily available data, the number of AI use cases has been growing tremendously. Model improvement is one of the main obstacles in all of these ML and DL projects across various industries. The performance and accuracy of the models are the fundamental elements of success. Model performance is mostly a technical issue, and deployment doesn't make sense for some machine learning and deep learning use cases if the model isn't precise enough for the specific business use case. Establishing a solid baseline model is crucial for improving any deep learning or machine learning model. A good baseline model takes into account all the technical and business requirements, evaluates the data engineering and model deployment pipelines, and acts as a standard for developing additional models. The specific application, the type of dataset, and the business domain all have an impact on the baseline model selection. Gradient Boosted Decision Trees are often utilized in production for several regression and classification-based applications. Frameworks have been proposed [1] to perform sentiment analysis of disaster data available on Twitter. Methods have been suggested [2] [21]for combining emotions & improving catastrophe relief through analysis and subject modelling which would make it possible to identify topics associated with trends in sentiment. Sentiment analysis is employed in research to identify how individual regions may have reacted differently compared to others and during all phases of the catastrophe. Techniques for partitioning and clustering are propagated [3] and additional clustering methods are proposed that would devise a different technique to assess twitter sentiments. In addition to examining the temporal relationship between sentiment and human mobility, including the dynamic effects of the earthquake over time, studies [4] to examine hypotheses of spatial characteristics of sentiment before, during, and after a disaster have revealed a significant negative correlation between sentiment levels and earthquake intensity levels and demonstrated that sentiment tends to cluster in space in distinct disaster intensity zones [5]. The effects of citizen equity on disaster situational awareness and harnessing machine learning and geotagged Twitter data [6][7] concerning sentiment analysis throw light on how evolution and analysis of public sentiments during a disaster. Studies [8] have shown that a machine learning system can be trained to recognise negative, neutral, or positive sentiments, according to data analysis. Additionally, it is feasible to categorise tweets that include damaging or high-accuracy transportation-related information. The adaptation of various text mining techniques[9] has successfully revealed the public's thoughts and sentiments concerning the CoViD-19 pandemic. Models have been built using studies [10] based on the svmRadial, C5.0, and NB classifiers. These models [11] for ensemble classification by including dictionary classifier results in the feature set producing improvised results. Studies [12] to analyze the Twitter response of users to disasters at global, country, and county levels recorded a variety of sentiments amongst users and victims. To better handle crises, both government and non-government organisations can benefit from the research [13] on a largely unexplored area of catastrophic scenarios for crowd sentiment detection. Analysing the emotions in texts with social research has been done on media data[14] emphasizing the cruciality of plans for the future that include actions and methods to make a topic emotion detection with outcomes of the contrasting techniques for examining social sensitivity and highlighting encouraging elements. Review and analysis [15] of various classifications on disaster datasets have concluded similarities between English and Spanish sentiments with TAN and BF TAN offering interesting qualitative information to historically and socially comprehend the main features of the event dynamics, even if there are no sufficient training examples. The Syrian civil war and the ensuing refugee crisis, two of the most terrible and urgent situations in the world right now, were the focus of some sentiment analyses[16] using Twitter data. The sentiments revealed different emotions based on language, region and type of disaster. Studies [17] on how tweets are connected and what topics are trending have revealed how people are participating in the crisis by sharing accurate information for everyone's awareness, calling on one another for assistance, and adhering to the precautionary measures imposed by the authorities. These studies have helped to understand what is happening and how people are reacting to the current situation, which will also help implement policies that would be beneficial to the economy. For classification, multilayer perceptron models have also been suggested[18]. Of tweets about resource requirements and resource availability during disaster availability and others, who have demonstrated improved outcomes in all the test cases and demonstrated its viability for using data from an unknown disaster. Social media can be a useful data source to examine people's worries amid a natural disaster, according to research [19][23]. A framework that may be utilised both during and after natural catastrophes is suggested as a technique to help disaster management teams uncover the public's issues quickly and construct a better real-time crisis management plan. concentrating on the examination of hashtag-containing tweets, Researchers have been able to pinpoint the primary issues that the world's population is worried about concerning the environment, public health, and sustainable development of the planet[20].

III. METHODOLOGY

Experimental setup:

Data Collection: The majority of the dataset was gathered from internet archives and the microblogging website Twitter using its API and URLs. Even though Twitter makes the data publicly available, it is only accessible for seven days, and real-time access is only available for 1% of all public tweets; neither protected accounts nor direct messages are accessible. In contrast, several online archives, like Kaggle, archive.com, DART, etc., offer a sizable resource of Twitter data.

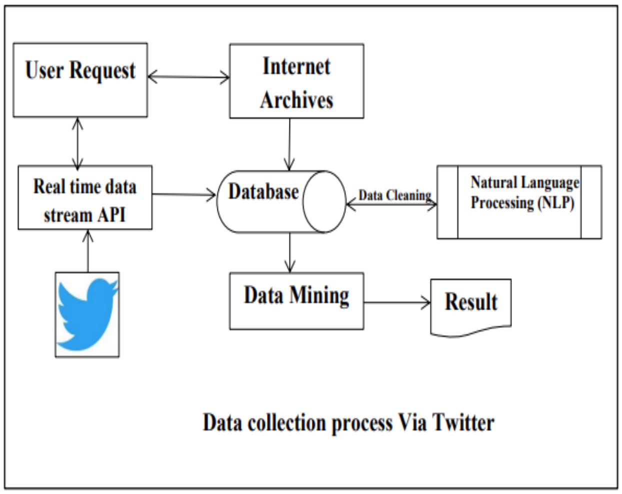


Figure 1: Twitter data collection process

For the current experimental configuration, information from Twitter, Kaggle, and archive.org was combined. The process of data collection is illustrated in Fig-1.
The collected data sample comprises test and training parts. Of the 12745 sampled tweets, 10196 were considered to be the training data and the remaining 2549 tweets were treated as test data. The exploratory data analysis of the data comprises label analysis, feature scrutiny and sentence length analysis. The data is then cleansed by removing URLs, handle tags, HTML tags, stopwords & stemming and useless characters. The cleaned data is labelled true or false based on the lexical analysis and the machine learning and deep learning analysis of the labelled data is performed to analyze the results. The proposed architecture of the process is depicted in Fig-2.

Exploratory Data Analysis:
Exploratory data analysis (EDA) is the process of doing preliminary analyses on data to find patterns, identify anomalies, test hypotheses, and double-check assumptions with the aid of summary statistics and graphical representations. Fig-3 reveals the number of examples and their proportions whereas the sentences and their proportionate lengths can be viewed in Fig-4. The EDA also explores the number of characters in the segregated disaster and non-disaster specific tweets as depicted in Fig-5.

Data Cleaning:
The data cleaning process comprises of Removal of

- URL
- Handle tags
- Emojis
- HTML tags
- Stopwords
- Stemming
- Useless Characters

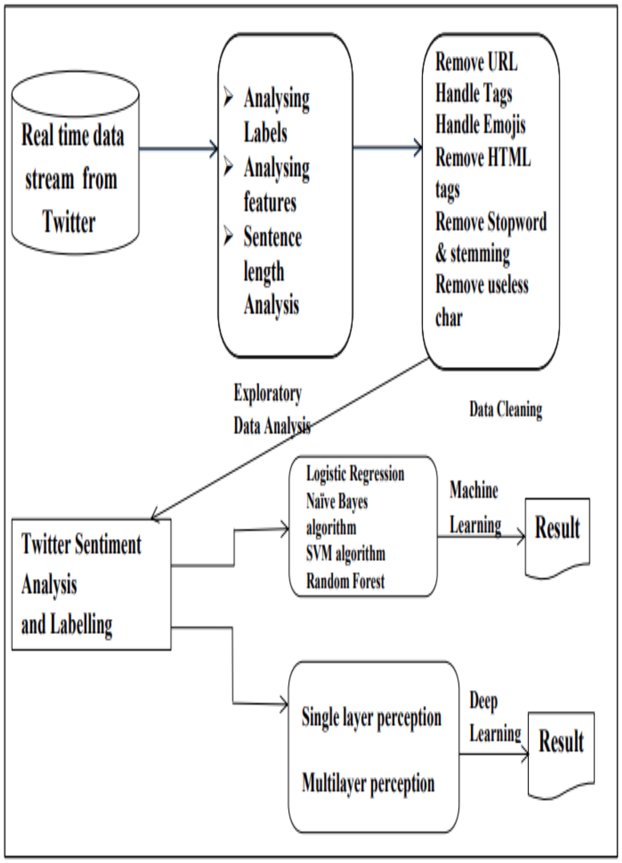
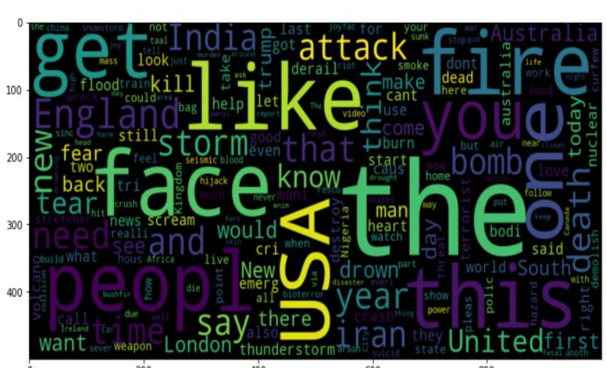


Figure 2: Overall proposed architecture

Representing the cleaned data using a word cloud gives a clear outlook of the text content. Term clouds, often referred to as text clouds or tag clouds, operate straightforwardly: the more frequently a certain word appears in a textual data source (such as a speech, blog post, or database), the bigger and bolder it will appear in the word cloud. A grouping of words shown in various sizes is called a word cloud. The more often and how important a word is mentioned in a document, the bigger and bolder it appears. These are excellent methods for extracting the most important portions of textual material, including social media posts and databases, and are also referred to as tag clouds or text clouds. They can also assist business users in comparing and contrasting two different texts to identify phrase overlaps. A sample word cloud of the collected sentiment data is illustrated in Figure 3. The final pre-processing stage involves merging the cleansed data which can then be utilized as an input for Machine learning and deep learning processes.



CPU times: user 1.72 s, sys: 149 ms, total: 1.87 s
Wall time: 1.77 s

Figure 3: Wordcloud of the cleaned data

Machine Learning:

Several machine learning approaches have been propagated to validate the key performance indices such as Micro & macro F1 score, micro averaged F1 score (mean F score) and all the models are compared based on Accuracy, precision, recall, F1-Score and time. We use the following widely used algorithms to validate the data.

- Logistic Regression-A statistical analysis method to predict a binary outcome, such as yes or no, based on prior observations of a data set.
- Naïve Bayes-family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong (naive) independence
    - Gaussian assumes that each class follows a Gaussian distribution
    - Bernoulli-It uses the Bernoulli Distribution as its foundation and only accepts binary data, i.e., 0 or 1. We can presume that Bernoulli Naive Bayes will be the algorithm to apply if the dataset's characteristics are binary.
    - Complement-Instead of estimating the likelihood that an object would belong to a certain class, we determine the likelihood that it will belong to all classes.
- Multinomial-a probabilistic learning method that is mostly used in Natural Language Processing (NLP). The algorithm is based on the Bayes theorem and predicts the tag of a text such as a piece of email or newspaper article.
- Support Vector Machines-a set of supervised learning methods used for classification, regression and outliers detection.
    - RBF Kernel-the default kernel used within the sklearn's SVM classification algorithm
    - Linear Kernel-used when the data is Linearly separable, that is, it can be separated using a single Line. It is one of the most common kernels to be used. It is mostly used when there are a Large number of Features in a particular Data Set.
- Random Forest- mixes the results of various decision trees to arrive at a single conclusion. Its widespread use is motivated by its adaptability and usability because it can solve classification and regression issues.

Deep Learning:

Deep learning is a machine learning method that teaches computers to perform tasks that humans accomplish without thinking about it. A computer model learns to carry out categorization tasks directly from images, text, or sound using deep learning. Modern precision can be attained by deep learning models, sometimes even outperforming human ability. A sizable collection of labelled data and multi-layered neural network architectures are used to train models**.** We adopt different models in combination with deep learning techniques and test the outcomes.

SINGLE LAYER PERCEPTION: There are only two layers of input and output in a single-layer perceptron. The name single layer perceptron refers to the fact that it only has one layer. It doesn't have any secret layers.

MULTILAYER PERCEPTION: MLP is the abbreviation for multi-layer perception. It is made up of dense, completely connected layers that may change any input dimension into the desired dimension. A neural network with numerous layers is referred to as a multi-layer perception. To build a neural network, we combine neurons so that some of their outputs are also their inputs.

We deliberate different models associated with the perceptions using a combination of the below-listed functions:

SIGMOID: The sigmoid function is a mathematical function that has a characteristic that can take any real value and map it to between 0 to 1 shaped like the letter "S". The sigmoid function is also known as a logistic function.

ADAM: a stochastic gradient descent alternative optimization technique for deep learning model training. Adam creates an optimization technique that can handle sparse gradients in noisy situations by combining the best aspects of the AdaGrad and RMSProp algorithms.

SGD: A straightforward yet highly effective method for fitting linear classifiers and regressors under convex loss functions, such as (linear) Support Vector Machines and Logistic Regression, is stochastic gradient descent (SGD).

RELU: In deep neural networks or multi-layer neural networks, ReLu is a non-linear activation function. In deep neural networks or conventional neural network paradigms, the activation levels have been calculated instead of using the ReLu function.

BATCH NORMALIZATION: The contributions to a layer for each mini-batch are normalised when using the batch normalisation technique to train very deep neural networks. As a result, the learning process is stabilised, and the quantity of training epochs needed to train deep neural networks is drastically reduced.

DROPOUT: Dropout is a training method in which some neurons are ignored at random. They "drop out" at random. This means that any weight updates are not applied to the neuron on the backward pass, and their contribution to the activation of downstream neurons is temporally erased on the forward pass.

IV. Results & Discussion

Low variance and strong bias characterize an underfit model. Regardless of the exact samples in the training set, it cannot learn the problem. A model that is overfitting has a high variance and low bias. When fresh, previously unknown examples or even statistical noise is added to examples in the training dataset, the model performs inconsistently because it learns the training data too well.

By expanding the model's capabilities, we can address underfitting. When a model has a higher capacity, it can fit more different types of functions for mapping inputs to outputs. Capacity describes a model's ability to fit a range of functions. By altering the model's structure, for as by including extra layers and/or nodes, it is simple to increase a model's capacity.

Overfit models are more prevalent since an underfit model can be fixed so quickly. Monitoring the model's performance during training by assessing it on both a training dataset and a holdout validation dataset makes it simple to identify an overfit model. The learning curves, which are line plots of the model's performance throughout training, will reveal a well-known pattern. As evident from the results depicted in Figure 4, it is observed that the effectiveness of machine learning methods cannot be underestimated. As can be seen in Figure-5, RELU with a combination of ADAM and DROPOUT proves to be the best fitting model and SVM continues to be the best in terms of accuracy and training time.

V. Conclusion & Future Work

Both supervised and unsupervised machine learning are possible. Choose supervised learning if fewer data points with well-marked training data are available. For huge data sets, unsupervised learning would typically perform and produce superior outcomes. Deep learning's primary objective is to get better with every new piece of data. This involves having the ability to change its fundamental structure to properly evaluate data. Utilizing consumer analytics, once that network has been completely developed using the test data, it enables more individualized service. As a part of our learning approach various machine learning and deep learning, models were applied to the labelled data and the performances were evaluated. The future work is planned on a larger data set with different combinations of models.

Machine Learning / Deep Learning	Accuracy	Precision	Recall	F1-Score	Time
Logistic Regression	0.8821	0.8699	0.6875	0.7336	183 ms
Gaussian Naïve Bayes	0.6543	0.6183	0.7031	0.5947	909 ms
Bernoulli Naïve Bayes	0.8948	0.8415	0.7649	0.7951	639 ms
Complement Naïve Bayes	0.8571	0.7536	0.7958	0.7712	369 ms
Multinational Naïve Bayes	0.8812	0.8783	0.6801	0.7267	352 ms
RBF Kernel SVM	0.8913	0.8787	0.716	0.7632	10.2 s
Linear Kernel SVM	0.8927	0.8499	0.7447	0.7821	5.79 s
Random Forest	0.8861	0.8497	0.7168	0.7584	5.89 s
Single Layer Perceptron	0.1741	0.0871	0.5	0.1483	45.2 s
Model 1 : SIGMOID + ADAM	0.1741	0.0871	0.5	0.1483	13min 27s
Model 2 : SIGMOID + SGD	0.1741	0.0871	0.5	0.1483	12min 23s
Model 3 : RELU + ADAM	0.8883	0.8187	0.7679	0.7894	13min 25s
Model 4 : RELU + SGD	0.1741	0.0871	0.5	0.1483	12min 33s
Model 5 : SIGMOID + BN + ADAM	0.1741	0.0871	0.5	0.1483	13min 52s
Model 6 : SIGMOID + BN + SGD	0.1741	0.0871	0.5	0.1483	13min 7s
Model 7 : RELU + DROPOUT + ADAM	0.8922	0.8305	0.7683	0.7938	14min 1s
Model 8 : RELU + DROPOUT + SGD	0.1741	0.0871	0.5	0.1483	13min 9s

Figure-4: Results of various deep learning models

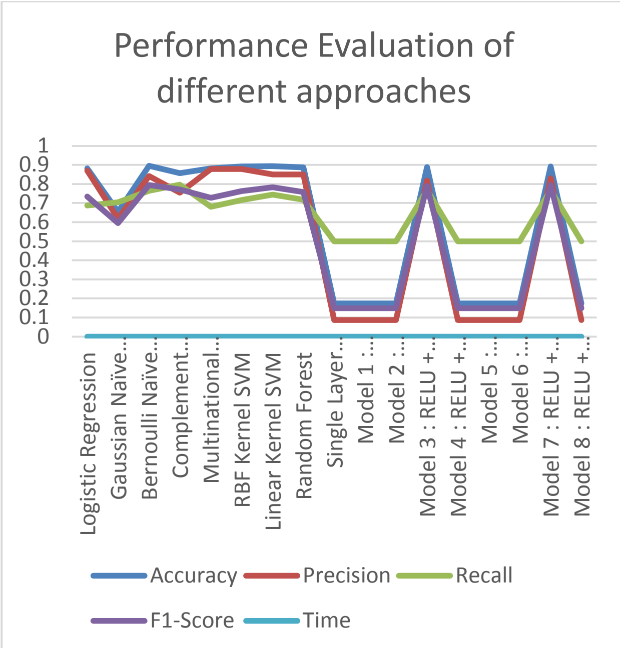


Figure-5: Comparison of performances

References

[1] Shalak Mendon, Pankaj Dutta, Abhishek Behl, Stefan Lessmann, "A Hybrid Approach of Machine Learning and Lexicons to Sentiment Analysis: Enhanced Insights from Twitter Data of Natural Disasters", Information Systems Frontiers, https://doi.org/10.1007/s10796-021-10107-x.

[2] Ike Vayansky, Sathish A.P. Kumar, Zhenlong Li, "An Evaluation of Geotagged Twitter Data during Hurricane Irma Using Sentiment Analysis and Topic Modeling for Disaster Resilience", 2019 IEEE International Symposium on Technology in Society (ISTAS) Proceedings Miriam Cunningham and Paul Cunningham (Eds) ISBN: 978-1-7281-5480-0.

[3] Hima Suresh, Dr Gladston Raj. S, "An Unsupervised Fuzzy Clustering Method for Twitter Sentiment Analysis", 2016 International Conference on Computational Systems and Information Systems for Sustainable Solutions, 978-1-5090-1022-6/16/$31.00 ©2016 IEEE.

[4] Shiyang Liaoa, Junbo Wangb, Ruiyun Yua, Koichi Satob, Zixue Chengb, "CNN for situations understanding based on sentiment analysis of Twitter data", 1877-0509 © 2017 The Authors. Published by Elsevier B.V, Peer-review under the responsibility of the organizing committee of the 8th International Conference on Advances in Information Technology, 10.1016/j.procs.2017.06.037

[5] Türkay Dereli, Nazmiye Eligüzel, Cihan Çetinkaya, "Content analyses of the international federation of red cross and red crescent societies (ifrc) based on machine learning techniques through Twitter", Natural Hazards (2021) 106:2025–2045, https://doi.org/10.1007/s11069-021-04527-w.

[6] Yan Wang, John E. Taylor, " Coupling sentiment and human mobility in natural disasters: a Twitter-based study of the 2014 South Napa Earthquake ", Nat Hazards (2018) 92:907–925,
https://doi.org/10.1007/s11069-018-3231-1.

[7] Wei Zhai, Zhong-Ren Peng, Faxi Yuan, "Examine the effects of neighbourhood equity on disaster situational awareness: Harness machine learning and geotagged Twitter data.", International Journal of Disaster Risk Reduction 48 (2020) 101611, https://doi.org/10.1016/j.ijdrr.2020.101611.

[8] Darcy Reynard, Manish Shirgaokar, "Harnessing the power of machine learning: Can Twitter data be useful in guiding resource allocation decisions during a natural disaster?", Transportation Research Part D xxx (XXXX) xxx–xxxhttps://doi.org/10.1016/j.trd.2019.03.002.

[9] Jasten Keneth D. Treceñe, Ralph Jerico P. Abides, " How did Twitter Users React to the Pandemic? Social Network Analysis of Public Tweets on CoViD-19 Outbreak", Asia Pacific Journal of Multidisciplinary Research.

[10] Sangeeta Rani, Nasib Singh Gill," Hybrid Model using Stack-Based Ensemble Classifier and Dictionary Classifier to Improve Classification Accuracy of Twitter Sentiment Analysis", International Journal of Emerging Trends in Engineering Research,
https://doi.org/10.30534/ijeter/2020/02872020.

[11] Muhammed Ali Sit, Caglar Koylu, Ibrahim Demir,” Identifying disaster-related tweets and their semantic, spatial and temporal context using deep learning, natural language processing and spatial analysis: a case study of Hurricane Irma”, International Journal of Digital Earth, https://doi.org/10.1080/17538947.2018.1563219.

[12] Lei Zou, Nina S. N. Lam, Heng Cai & Yi Qiang,” Mining Twitter Data for Improved Understanding of Disaster Resilience”, Annals of the American Association of Geographers, https://doi.org/10.1080/24694452.2017.1421897.

[13] Ms Harvinder Jeet Kaur, Mr Rajiv Kumar, "Sentiment Analysis from Social Media in Crisis Situations", International Conference on Computing, Communication and Automation (ICCCA2015), ISBN:978-1-4799-8890-7/15/$31.00 ©2015 IEEE.

[14] Yasin Kirelli, Seher Arslankaya." Sentiment Analysis of Shared Tweets on Global Warming on Twitter with Data Mining Methods: A Case Study on the Turkish Language". Hindawi Computational Intelligence and Neuroscience Volume 2020, Article ID 1904172, 9 pages, https://doi.org/10.1155/2020/1904172.

[15] Gonzalo A. Ruz , Pablo A. Henríquez, Aldo Mascareño,” Sentiment analysis of Twitter data during critical events through
Bayesian networks classifiers”, Future Generation Computer Systems, https://doi.org/10.1016/j.future.2020.01.005.

[16] Marjan Kamyab, Ran Tao, Mohammad Hadi Mohammadi, Abdul Rasool, “Sentiment Analysis on Twitter: A text Mining Approach to the Afghanistan Status Reviews”. Association for Computing Machinery. ACM ISBN 978-1-4503-6641-0/18/11,
https://doi.org/10.1145/3293663.3293687.

[17] Nazan Öztürk, Serkan Ayvaz, “Sentiment Analysis on Twitter: A Text Mining Approach to the Syrian Refugee Crisis”, Accepted Manuscript,
https://doi.org/10.1016/j.tele.2017.10.006.

[18] Jasten Keneth D. Treceñe, Ralph Jerico P. Abides, “How did Twitter Users React to the Pandemic? Social Network Analysis of
Public Tweets on CoViD-19 Outbreak”, Asia Pacific Journal of Multidisciplinary Research, P-ISSN 2350-7756 | E-ISSN 2350-8442.

[19] Shivam Behl, Aman Rao, Sahil Aggarwal, Sakshi Chadha, H.S. Pannu, "Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises", International Journal of Disaster Risk Reduction, https://doi.org/10.1016/j.ijdrr.2021.102101 .

[20] Amir Karami, Vishal Shah, Reza Vaezi, Amit Bansal, “Twitter speaks: A case of national disaster situational awareness”, Journal of Information Science, DOI: 10.1177/0165551519828620.

[21] Ana Reyes-Menendez, José Ramón Saura, Cesar Alvarez-Alonso, “Understanding #WorldEnvironmentDay User Opinions in Twitter: A Topic-Based Sentiment Analysis Approach”, International Journal of Environmental Research and Public Health.

[22] Nastaran Pourebrahima, Selima Sultanaa, John Edwardsb, Amanda Gochanourb, Somya Mohantyc, “Understanding communication dynamics on Twitter during natural disasters: A case study of Hurricane Sandy”, International Journal of Disaster Risk Reduction, https://doi.org/10.1016/j.ijdrr.2019.101176.

[23] Saimah Bashir, Shohar Bano, Sheikh Shueb, Sumeer Gul, Aasif Ahmad Mir, Romisa Ashraf, Shakeela, Neelofar Noor, “Twitter chirps for Syrian people: Sentiment analysis of tweets related to Syria Chemical Attack”, International Journal of Disaster Risk Reduction, https://doi.org/10.1016/j.ijdrr.2021.102397.

[24]https://github.com/erYash15/Real-or-Not-NLP-with-Disaster-Tweets