An Adoptive Learning Process for Social Media Text data Analysis for Disaster Management

Praveen Sharma and Dr. Deepika Pathak
Department of Computer Application,
Dr. A. P. J. Abdul Kalam University, Indore (M.P.) - 452016
Corresponding Author Email: praveendadhich.tj@gmail.com

Article Info
Page Number: 10153-10165
Publication Issue:
Vol. 71 No. 4 (2022)

Article History
Article Received:
12 September 2022
Revised: 16 October 2022
Accepted: 20 November 2022
Publication: 25 December 2022

Abstract
Social media is a source of low cost and publically available news and information. Social media users are contributing the information during the disaster events are valuable for timely disaster response. Therefore, mining help related data in early time of social media content is a valuable to understand the disasters situation. In this paper, we proposed an adoptive learning process to analyze the social media data to recognize the disaster event and it's intensity by using unsupervised learning and sentiment analysis. First, the labeled data is used to train a modified Fuzzy C Means clustering. Additionally, the centroids are updated with incorporating reveled new keywords. The improved centroids are useful for identifying help related social media contents. Next, the sentiment score are used to identify the negative sentiments for identifying the severity of the disaster. In order to classify the text according to their sentiments we utilize the SVM and ANN classifiers. Finally, the results of the proposed model is measured and compared with the baseline model. The results revel that the proposed adoptive learning technique is efficient and enhancing the accuracy with the time.

Keywords: Machine Learning Algorithm, Supervised Learning, Smart Grid Applications, Review, Experimental Study, Performance Comparison.

I. INTRODUCTION

The social media connect the people around the world. The entire world can access the shared information easily and in less time. Therefore, the social media is become an information tool during disaster situations. During disasters people can use the social media platform to request help. Therefore it becomes a tool to distribute the information about the disaster events and its severity. The governmental agencies and emergency responders can utilize this information for providing timely relief to the disaster victims. Therefore monitoring, mining and extracting help related information accurately and efficiently during disaster.

The ML based techniques can be used for all the phases of disaster management, in the early stages for prediction of disaster. During disaster this will be used for providing the awareness and precautions. And during post disaster the social media can be used to figure out the area of affect, location and requirements. In this paper, we are motivated to propose semi supervised learning approach for social media content analysis and discovery of the disaster events more specifically the help related post. In this approach the machine learning algorithms are

employed for accurately and efficiently data identification. Additionally perform the sentiment analysis to identify the severity of disasters.
The proposed method is based on adoptive learning process, which utilizes the unsupervised learning technique for learning with the prior social media labeled data. This process is finding the more relevant keywords and updates them for future unlabelled data recognition. The proposed adoptive learning technique is discussed and their working is described. Therefore, in this paper we first provide a review of the recent articles based on ML and social media content based disaster management systems. Next, we provide an overview of the unsupervised learning algorithms. Additionally the working process is also described. Finally, the proposed algorithm is evaluated and compared with the traditional approaches to identify the help related request during the disaster events. Finally the conclusion and future work has been discussed.

II. RELATED WORK

In this section, we provide the available solutions for the rapid response in disaster management problems based on Machine Learning (ML)by recent researchers. The importance of disaster management is evident by the increasing number of natural and manmade disasters such as Irma and Manchester attacks.
In this section, we provide the available solutions for the rapid response in disaster management problems based on Machine Learning (ML) by recent researchers.
M. Aqib et al [1] have developed a disaster management system to evacuation strategies. They extend earlier work by using deep learning to predict urban traffic behavior. They first apply deep learning in disaster management and use road traffic by UK Department for Transport. The results show the effectiveness of disaster management and prediction of traffic behavior.
V. Linardos et al [2] aims to provide an overview of the studies, presented since 2017, focusing on ML and DL methods for disaster management. Main the areas of disaster and hazard prediction, risk and vulnerability assessment, disaster detection, early warning systems, disaster monitoring, damage assessment and post-disaster response.
Social media serves as a part of the crisis response regardless of the kind of disaster, whether it is. ML approaches extract disaster indicating posts, which is useful. L. Dwarakanath et al [3] reviews and classifies into three phases – early warning and detection, post-disaster coordination and response, damage assessment. This review helps in choosing topics to automated approaches for actionable information classification and disaster coordination in emergency.
H. S. Munawar et al [4] is developed an understanding about the flood risks, explore the existing systems for managing the risks and flood management. Floods as a disaster will be viewed with analysis of the threats. Identified problem and working towards a solution for managing the damage and destruction. The disaster response is also investigated with its limitations. Finally, a ML and image processing based solution is proposed. The system is formulated to overcome the limitations and present a robust and reliable model.

Table 1 Review summary

Ref	Contribution	Technique used	Algorithm	Dataset
[1]	Disaster management system	VANET, cloud computing, and simulations to evacuation strategies	deep learning to predict urban traffic behavior	road traffic through the UK Department for Transport
[2]	Survey	Overview of studies based on ML and DL for disaster management	Deep learning	Recent ML and DL applications for disaster management have been analyzed.
[3]	Survey	ML approaches for disaster response using social media	ML in social media	three phases – early, post-disaster and response, assessment
[4]	Flood control and disaster management	understanding of flood risks, managing risks and flood	ML and image processing based solution	Advancements in the disaster management and gaps
[5]	datasets of Brazilian disasters	include 9 types of disaster	geographical, demographic and socioeconomic	Multipurpose dataset
[6]	People activities and behavior compared to normal condition	detecting the emergency and its degree	Random Forest (RF), IBK, Bagging, J48 and MLP	Smartphone and Smartwatch Activity and Biometric Dataset
[7]	analyze social media posts to assess the footprint of damage caused	hot spot detection	Latent Dirichlet Allocation with spatial and temporal analysis	US Geological Survey
[8]	Quality of OpenStreetMap (OSM) data	understand who contributed data when and how	unsupervised ML	OSM history data
[9]	Unlabelled and labeled data and	domain adaptation approach	Naïve Bayes, with Self-Training strategy	Unlabelled and labeled data tweets

[10]	Prediction models of heavy rain damage	reduce the damage through proactive management	decision trees, bagging, random forests, and boosting	Seoul Capital in the Korea
[11]	multisource data fusion	decision level fusion approach	SVM and Probabilistic Neural Network(PNN)	UAVSAR, RapiEye, RADARSAT-2
[12]	fear-sentiment	textual analytics and visualizations	Naïve Bayes and logistic regression	pandemic specific Tweets
[13]	help related requests during disaster	disaster response through sentiment analysis	ML algorithm	social media data for disaster response and recovery
[14]	Survey	highlights challenges and presents techniques to deal with social media messages	ML Algorithms	social media content
[15]	Survey	Leaving and learning	-	-
[16]	Flood inventory and flood causative factors	Integrating multiple models for classification approach	statistical, ML, and multi-criteria decision analysis	Remote sensing data, Mike-11 hydrological model and multi source data.
[17]	prediction of fissuring hazards	Modeling the fissure hazard	GLM, MARS, CART, random forest (RF), and SVM	Simulation based data

R. Veloso et al [5] present datasets of Brazilian disasters from Jan 2003 to Feb 2021, through government and institutes reports. The datasets include types of disaster, and number of affected people during 18-year for municipalities. Data on geographical, demographic and socioeconomic aspects are provided. The data can be used for supervised and unsupervised ML techniques. It is useful for the visualization. Data can be used for optimization related problems. Describe two real-world cases for the location, costs and distances.

If people are in emergency, then their activities and behavior will be different from normal situations. To detect emergency, Human Activity Recognition (HAR) can be used. Emergency Management System (EMS) can manage the situation. S. Nanda et al [6] use algorithms as

Random Forest (RF), IBK, Bagging, J48 and MLP on Smartphone and Smartwatch Activity and Biometric Dataset and found RF is the best algorithm with accuracy 87.1977%.
Disaster management suffers from high or low temporal lags and spatial resolution. B. Resch et al [7] analyze social media posts to assess the footprint of damage caused by natural disasters through ML for semantic information with spatial and temporal analysis for hot spot detection. The results demonstrate that earthquake footprints reliably and accurately identified. Relevant topics identified without a priori knowledge. Able to generate a damage map to indicates losses. The results using statistical measures complemented by US Geological Survey and shows that produces valid and reliable outputs.
When reference data is not available to assess the quality of OpenStreetMap (OSM) data, its metadata can be used. A. Madubedube et al [8] applied unsupervised ML for analyzing OSM history data to understand who contributed when and how. Even no statements given about the quality of the data, provide insight into the quality. Most of the data in Mozambique was contributed. Results revealed a new class: contributors who were new and attracted by HOT mapping events during disaster relief. They provide suggestions for contributors. Supervised learning relies on labeled data, which is not readily available for disaster. While labeled data available for a prior source, supervised classifiers learned only from the source disaster not perform well. H. Li et al [9] propose a domain adaptation approach, which learns from unlabelled data, in addition to source labeled data. This approach uses the Naïve Bayes, with an iterative Self-Training strategy. Results on identifying tweets relevant to a disaster are show that the domain adaptation classifiers are better.
Prediction models of heavy rain damage using ML were developed for the Seoul Capital in the Korea. C. Choi et al [10] used data on the occurrence of heavy rain damage from 1994 to 2015. The ML model was developed by decision trees, bagging, random forests, and boosting. Result of the prediction performance, the AUC value of the boosting using meteorological data from the past 4 days was the highest at 95.87%. By using the prediction model we predict the rain damage for each region.
To multisource data fusion data is combining in a stacked vector. The decision level fusion approach combines statistical and ML techniques. B. Gokaraju et al [11] use this approach to calculate results of the disaster management. Dataset of levee landslide, Tornado disaster, synthetic aperture radar sensor and multispectral sensor datasets is used. The results of fusion outperformed the non-data fusion techniques.
Corona pandemic has manifested in the form of fear, fueled by incomplete and inaccurate information. It is a need to understand informational crisis and gauge sentiment, for appropriate messaging and policy making. J. Samuel et al [12] identify public sentiment with the pandemic specific Tweets and R. They demonstrate insights into the progress of fear-sentiment, using textual analytics and visualizations. They provide an overview of ML methods, and compare in classifying Tweets. The classification accuracy of 91% with the Naïve Bayes and for logistic regression 74% with shorter Tweets and methods are weak for longer Tweets.
Social networks are used for communications and help related requests during disaster. The sentiment of people during and after the disaster determines the success of the disaster response and recovery. J. R. Ragini et al [13] propose an approach for disaster response

through sentiment analysis. The model collects disaster data from social networks and categorizes them according to the needs. The categorized data are classified through ML algorithm. Features like, parts of speech and lexicon are used to identify strategy. The results show that lexicon based approach is suitable for needs. The methodology is the real-time categorization and classification of social media data for disaster response and recovery.
Social media Information in the early hours are valuable for responders and decision makers, to gain situational awareness and plan relief. Processing social media content to obtain such information involves challenges, of parsing brief messages, handling overload, and prioritizing information types. M. Imran et al [14] highlights challenges and presents techniques to deal with social media messages, during crisis scenarios.
Ministry of Home Affairs, India decided to enforce a complete shutdown. This has affected all sectors including the education. A. Khattar et al [15] survey to understand the day to day living, activities, learning styles, and mental health of students during crisis and assess how they are adapting to the e-learning and managing social lives.
M. Rahman et al [16] proposes an approach by statistical, ML, and multi-criteria decision analysis. Flood inventory and flood causative factors were prepared from different sources. The flood inventory was divided, where 334 locations were used for training and the 141 locations for testing. Using the AUROC, predictive power of the model was tested. The results revealed that LR model had the highest success rate (81.60%) and prediction rate (86.80%). And the best combination was used for generating a flood hazard map. The LR-FR model had predictive power of 88.10%.
Excessive withdrawal of groundwater, and natural resources, has been introduced as land subsidence and earth fissuring. Fissuring is turning into the disasters. B. Choubin et al [17] proposing ML models for prediction of fissuring hazards. The simulated annealing was used to identify features, and the generalized linear model (GLM), multivariate adaptive regression splines (MARS), classification and regression tree (CART), random forest (RF), and SVM have been used for prediction. Results indicated that all the models had accuracy (> 86%) and precision (> 81%).

III. PROPOSED WORK

The proposed work is aimed to develop a machine learning model which is able to perform the following task:

- Identify the valuable tweets indicating disaster situation
- Extracting the new keywords for incorporating with the learning system
- Performing the sentiment analysis to rank the severity of the disaster

In this context we need to develop three different modules for designing the required model. The model requires an appropriate data set as the primary component. The dataset is obtained from the Kaggle [18], the dataset contains 5 attributes and 7503 instances of data. The attributes are ID, keyword, location, text and target. The ID is unique values for identifying each instance. Thus, we eliminate this attribute. Additionally, the keyword and location has a lot of missing values therefore we also removed these attributes. After data preprocessing it only contains the text and target attribute defined by 0 and 1. Here, 0 indicates normal situation and 1 used for disaster event.

The aim of text preprocessing is to reduce the noise of the contents such as abbreviations, stop words and special characters. In this analysis the hash-tags are very valuable and can be utilized as potential keywords. Therefore, in order to preprocess the text we eliminate the stop words, special characters and abbreviations. Additionally preserve the hash-tag associated keywords. Further, we need a feature selection process to transform text into Vector for utilizing with the machine learning algorithms. Thus, we utilize the Term Frequency and Inverse Document Frequency (TF-IDF). The TF-IDF is defined using the following formula:

$$TF = \frac{count\ of\ a\ term\ in\ a\ document}{total\ term\ in\ document}$$

and

$$IDF = \log\left(\frac{N}{df(t)}\right)$$

Where $df(t)$ is Document frequency of a term t, and N is Number of documents containing the term t. Finally for selecting text features we calculate the weights of each term using:

$$w = tf * IDF$$

The weights are used for selecting potential keywords from the social media text and transformed the data into a vector. In the proposed model these three initial steps are used for transforming the text into a vector therefore we utilize these steps as the common component of the model. Next we providing the details about the initial objective of this paper, thus we discuss them one by one and then provide a complete system.

A. Identify the valuable tweets indicating disaster situation

The key issue is to identify the tweets which are potentially belongs to the natural disaster. But practically we have limited prior labeled data available. Therefore, we need to utilize an unsupervised learning technique. The unsupervised learning utilizes the random centroids and optimizes the centroids in an iterative manner. The figure 1 demonstrates the working of the given model.

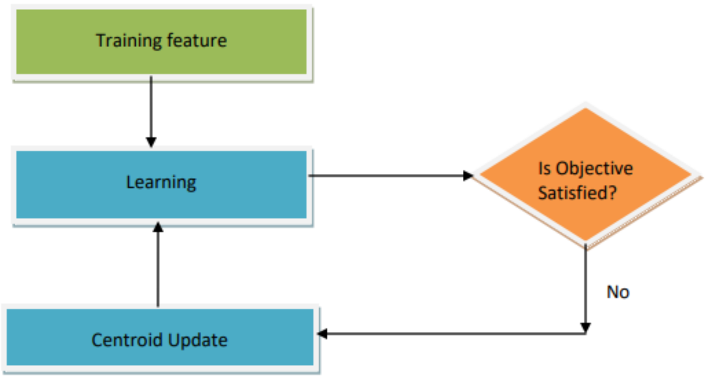


Figure 1: working of unsupervised learning process

In recent study we had compared K-Means and FCM algorithm. Among them we have found the FCM is better than k-means algorithm in their accuracy. Therefore in order to implement the required technique we utilized FCM algorithm for the accurate data extraction. FCM is working similar as K-means algorithm. But the initialization of the algorithm requires a random set of data instance as the centroid. The similar number of centroids is chosen as the number of groups or clusters required. In this presented work we are providing the initial centroid and update the centroid by modifying the update process. Before, initiating the optimization process the initial centroid selection process is given in next section.

B. Extracting the new keywords for incorporating with the learning system

In this work we need a single centroid, initially then we extend the centroids. Let, the C is the initial centroid and defined by:

$$C = \{null\}$$

In tweeter in most of cases for indicating a trending event users are utilizing some common has tags. Therefore, from the training samples we first select keywords with hash tags. Let the total keywords obtained by the hash tags H is given by:

$$H = \{k_1, k_2, \ldots, k_m\}$$

Where, k_i is the ith keyword selected through hash tags.

But the vector H has the following complexities:

1. Duplicate words.
2. Word with similar meaning and different or incomplete spelling.

Duplicate words are initially reduced by:

$$UD = \{H: k_i \notin UD\}$$

Additionally, for reducing the words which have the similar spelling we utilized the levenshtein distance L. Now, if the distance among two keywords in the set of keywords UD has the similarity greater then threshold $T = 0.75$. Therefore,

$$R = \begin{cases} if(L(k_i, k_{i+1}) > 0.75) \ \ then\ add\ to\ R \\ \qquad\qquad else\ remove\ from\ list \end{cases}$$

After refining the keywords from the selected hash tags we get a vector R. The R can be defined as:

$$R = \{k_1, k_2, \ldots, k_n\}$$

Here we need a similar size of centroid vector as the length of selected feature. Let the length of social media content feature is p, then we partitioned the centroid vector which has the length of n. Therefore centroid C,

$$C = \begin{cases} if\ p \geq n \ \ then\ C = pad(R) \\ if\ P < n\ then\ C = part(R) \end{cases}$$

Where, $pad(R)$ is the sequence padding if the length of centroid is less than the feature length, and $part(R)$ is creating part of centroid into two or more parts based on sequence length of p.

Now we utilize the centroid C and feature vector $X = \{x_1, x_2, \ldots, x_p\}$ obtained by TF-IDF feature selection technique. in this context when we apply the FCM algorithm then we need to minimize the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \parallel x_i - c_j \parallel$$

Where, m is a real number greater then 1, here we assumed $m = 2$, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the ith element of data, c_j is the centroid of the cluster, and ||*|| is the similarity between data and centroid.

FCM is optimizing the objective function, which is given in above equation and for computing membership u_{ij} we can use:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left[\frac{\parallel x_i - c_j \parallel}{\parallel x_i - c_k \parallel} \right]^{\frac{2}{m-1}}}$$

Additionally to update the centroids we use the following function:

$$c_j = \frac{\sum_{i=1}^{N} u_{ij} . x_i}{\sum_{i=1}^{N} u_{ij}^m}$$

After applying the clustering algorithms the relevant tweets indicating the disaster events. After computing the disaster event we need to rank the neediness of people therefore we perform the sentiment analysis for ranking the tweets.

Performing the sentiment analysis to rank the severity of the disaster

In this work we used a sentiment score calculation library namely **V**alence **A**ware **D**ictionary and s**E**ntiment **R**easoner (VADER) [19]. In order to calculate the sentiment score the library provides a function "polarity_scores". The polarity score is results in four different scores for a sentence or paragraph. The type of score and the value range is given in table 3.

Table 2 Sentiment score

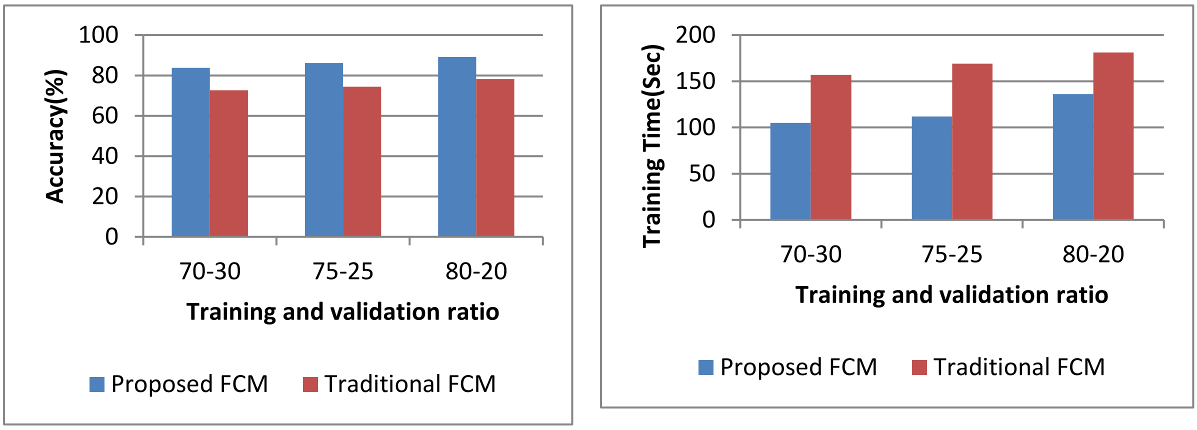
S. No.	Sentiment type	Score range
1	Negative	0 - 1
2	Positive	0 - 1
3	Neutral	0 – 1
4	Compound	-1 - +1

In this work, we utilize the compound sentiment score for representing the sentiment of the tweets. Additionally we map the compound sentiment score into the need of people. The mapping function is denoted as M.

$$M = \begin{cases} if\ score \geq 0\ \ response\ received \\ if\ score < 0\ and > -0.5\ response\ required \\ if\ score < -0.5\ then\ immediate\ response\ required \end{cases}$$

IV. RESULTS ANALYSIS.

This section describes the experimental evaluation and obtained results. In this context, the different set of training and validation is prepared i.e. 70-30, 75-25 and 80-20. Additionally the proposed approach is compared with the baseline FCM algorithm. in order validate the performance accuracy and training time of the model is considered.



(A) (B)

Figure 2 shows the performance in terms of (A) Accuracy and (B) Training Time

The accuracy is a ratio of correctly predicted class labels on the total samples are provided for prediction. That can be calculated using the following equation:

$$accuracy = \frac{correctly\ predicted\ samples}{total\ samples} X100$$

Next, for describing the efficiency of the proposed model the training time is calculated. The training time is the amount of time taken to perform training, which is calculated as:

$$training\ time = end\ time - start\ time$$

The aim of the proposed work is to obtain the accurate classification of disaster events and identify the request related to immediate response. In this context, we modify a FCM based unsupervised learning algorithm for categorizing the tweets into disaster and non disaster events. Additionally, we utilized the sentiment scores for identifying the intensity of response request. Figure 2(A) shows the comparative accuracy of both the versions of clustering algorithms for identifying the disaster events and its sentiment. Additionally the experimental observations are also reported in table 3. In this figure X axis shows the sample ratio used for the training and validation and Y axis shows the obtained accuracy in terms of percentage (%). According to the results it is found that the increasing amount of training sample will improve the performance of the FCM algorithm. Additionally the proposed method of FCM training will enhance the performance as compared to traditional approach of the FCM clustering. The main reason behind the enhanced accuracy is that the selection of potential keywords which are indicating the disaster events on the other hand the traditional FCM is completely depends on the TF-IDF based calculated features which are also contains the significant amount of noisy keywords.

Table 3: Performance of Clustering Algorithms

		Accuracy (%)		Training time (Sec)	
S. No.	Training and validation ratio	Proposed FCM	Traditional FCM	Proposed FCM	Traditional FCM
1	70-30	83.8	72.7	105	157
2	75-25	86.2	74.4	112	169
3	80-20	89.1	78.2	136	181

On the other hand the Figure 2(B) demonstrates the training time of both the proposed and traditional FCM clustering algorithms. The training time of both the models are measured here in terms of seconds (Sec). Additionally the observations of the experimental results are given in table 3. The X axis demonstrates the used sample size in terms of training and validation ratio, and Y axis shows the training time of the proposed and traditional model. According to the obtained results we found that the increasing amount of training time will increase the training time. Based on the measured training time the proposed technique requires less amount of training time as compared to traditional FCM approach because the initially created single centroid is updated with the new data. Thus the proposed algorithm just optimizing a single cluster but the traditional FCM algorithm needed to optimize both the clusters. Thus according to the observed results we found that the proposed FCM is more accurate and efficient than the traditional FCM algorithm.

V. CONCLUSION

The social media can be utilized as an information tool in order to manage the natural disasters. During the disaster situations identification of severity and providing timely relief to the disaster victims is one of the essential tasks for preventing losses. In this context, social media is also used for asking the help to others by post. In this paper the main aim is to analyze the social media data for recovering the natural disaster related content. Additionally utilize the identified data to support the disaster victims by mapping the sentiments into required response. In this context, we proposed an unsupervised social media content analysis algorithm. That algorithm modifies the traditional FCM clustering algorithm to minimize the training efforts and enhancing the accuracy of the FCM clustering. The modification involves the initial cluster centroid selection and optimization of the centroids by the potential keywords. The proposed method is compared with the baseline FCM algorithm in terms of identification accuracy and training time. The experimental observations demonstrate that the modification of FCM based on the proposed data modeling enhance the identification accuracy and efficiency of the algorithm.

REFERENCES

[1] M. Aqib, R. Mehmood, A. Albeshri, A. Alzahrani, “Disaster Management in Smart Cities by Forecasting Traffic Plan Using Deep Learning and GPUs”, Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2018, LNICST 224, pp. 139–154, 2018.

[2] V. Linardos, M. Drakaki, P. Tzionas, Y. L. Karnavas, “Machine Learning in Disaster Management: Recent Developments in Methods and Applications”, Mach. Learn. Knowl. Extr. 2022, 4, 446–473

[3] L. Dwarakanath, A. Kamsin, R. A. Rasheed, A. Anandhan, L. Shuib, “Automated Machine Learning Approaches for Emergency Response and Coordination via Social Media in the Aftermath of a Disaster: A Review”, IEEE Access, VOLUME 9, 2021

[4] H. S. Munawar, “Flood Disaster Management: Risks, Technologies, and Future Directions”, Machine Vision Inspection Systems (Vol. 1): Image Processing, Concepts, Methodologies and Applications, (115–146) © 2020 Scrivener Publishing LLC

[5] R. Veloso, J. Cespedes, A. Caunhye, D. Alem, “Brazilian disaster datasets and real-world instances for optimization and machine learning”, Data in Brief 42 (2022) 108012

[6] S. Nanda, C. R. Panigrahi, B. Pati, A. Mishra, “A Novel Approach to Detect Emergency Using Machine Learning”, Progress in Advanced Computing and Intelligent Engineering, Advances in Intelligent Systems and Computing 1199, 2021

[7] B. Resch, F. Usländer, C. Havas, “Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment”, Cartography and Geographic Information Science, 2018, Vol 45, No 4, 362–376

[8] A. Madubedube, S. Coetzee, V. Rautenbach, “A Contributor-Focused Intrinsic Quality Assessment of OpenStreetMap in Mozambique Using Unsupervised Machine Learning”, ISPRS Int. J. Geo-Inf. 2021, 10, 156

[9] H. Li, D. Caragea, C. Caragea, N. Herndon, “Disaster response aided by tweet classification with a domain adaptation approach”, J Contingencies and Crisis Management. 2018; 26:16–27

[10] C. Choi, J. Kim, J. Kim, D. Kim, Y. Bae, H. S. Kim, “Development of Heavy Rain Damage Prediction Model Using Machine Learning Based on Big Data”, Hindawi Advances in Meteorology Volume 2018, Article ID 5024930, 11 pages

[11] B. Gokaraju, R. A. A. Nóbrega, D. A. Doss, A. C. Turlapaty, R. C. Tesiero, "Data Fusion of Multi-Source Satellite Data Sets For Cost Effective Disaster Management Studies", 978-1-4673-9558-8/15/$31.00 ©2015 IEEE
[12] J. Samuel, G. G. Md. N. Ali, M. M. Rahman, E. Esawi, Y. Samuel, "COVID-19 Public Sentiment Insights and Machine Learning for Tweets Classification", Information 2020, 11, 314;
[13] J. R. Ragini, P. M. R. Anand, V. Bhaskar, "Big data analytics for disaster response and recovery through sentiment analysis", International Journal of Information Management, 42, 13–24, 2018
[14] M. Imran, C. Castillo, F. Diaz, S. Vieweg, "Processing Social Media Messages in Mass Emergency: Survey Summary", International World Wide Web Conference Committee, WWW '18 Companion, April 23–27, 2018, Lyon, France
[15] A. Khattar, P. R. Jain, S. M. K. Quadri, "Effects of the Disastrous Pandemic COVID 19 on Learning Styles, Activities and Mental Health of Young Indian Students - A Machine Learning Approach", Proceedings of the International Conference on Intelligent Computing and Control Systems, IEEE, 2020
[16] M. Rahman, C. Ningsheng, M. M. Islam, A. Dewan, J. Iqbal, R. M. A. Washakh, T. Shufeng, "Flood Susceptibility Assessment in Bangladesh Using Machine Learning and Multi-criteria Decision Analysis", Earth Systems and Environment, 2019
[17] B. Choubin, A. Mosavi, E. H. Alamdarlood, F. S. Hosseinid, S. Shamshirband, K. Dashtekiang, P. Ghamisih, "Earth fissure hazard prediction using machine learning models", Environmental Research 179 (2019) 108770
[18] https://www.kaggle.com/competitions/nlp-getting-started/data?select= train.csv
[19] S. Elbagir, J. Yang, "Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment", Proceedings of the International MultiConference of Engineers and Computer Scientists 2019, IMECS 2019, March 13-15, 2019, Hong Kong