Natural Disaster Application on Big Data and Machine Learning: A Review

1st Rania Rizki Arinta
Magister Teknik Informatika
Universitas Atma Jaya Yogyakarta
Yogyakarta, Indonesia
raniarinta@gmail.com

2nd Andi W.R. Emanuel
Magister Teknik Informatika
Universitas Atma Jaya Yogyakarta
Yogyakarta, Indonesia
andi.emanuel@uajy.ac.id

***Abstract*— Natural disasters are events that are difficult to avoid. There are several ways of reducing the risks of natural disasters. One of them is implementing disaster reduction programs. There are already several developed countries that apply the concept of disaster reduction. In addition to disaster reduction programs, there are several ways to predict or reducing the risks using artificial intelligence technology. One of them is big data, machine learning, and deep learning. By utilizing this method at the moment, it facilitates tasks in visualizing, analyzing, and predicting natural disaster. This research will focus on conducting a review process and understanding the purpose of machine learning and big data in the area of disaster management and natural disaster. The result of this paper is providing insight and the use of big data, machine learning, and deep learning in 6 disaster management area. This 6-disaster management area includes early warning damage, damage assessment, monitoring and detection, forecasting and predicting, and post-disaster coordination, and response, and long-term risk assessment and reduction.**

Keywords*— *natural disaster, review, big data, machine learning

I. INTRODUCTION

Natural disasters are events that result from natural processes that cannot be predicted. Natural disasters can also cause loss of life or damage property and economic losses. According to data provided by world health organizations from 1900 to 2018, there are around 14 million for all types of disasters [1]. Children are the most vulnerable groups who get direct impacts of disasters. Disaster victims come from various countries, and disasters have been threatening the lives of millions of children [2].

One effort to minimize disasters is to provide a program for reducing the risks of natural disasters. Disaster risk reduction is a concept of how communities reduce damage and victims affected by disasters. One example of risk reduction according to the Sendai framework [3] made by UNISDR (United Nations International Strategy for Disaster Reduction) is understanding the risk of the disaster. According to Goswami's research [4] states that the aim of disaster management is minimizing victims, can save victims promptly, evacuate people to save places, Reconstruct the damages immediately, Offer first aid instantly.

The technology today is big data, machine learning, along with deep learning. According to Hashem's research stated that big data is a set of techniques and technology that requires a new form of integration to find large hidden values derived from datasets that are complex, diverse, and of a massive scale. The advantage of using big data can determine the pattern obtained from data analysis and the creation of hidden information [5].

In this study, we will review and focus more on Disaster management phases, which focus more on the use of machine learning in the area of Disaster management phases, which consists of the data source used and the model/algorithm used. Because to find out whether the previous research solved the problem in the prediction area and early detection we must know the data source used already has 5'v characteristics, namely Velocity, Volume, Value, Variety, and Veracity. The performance level of the model made is good or not from the level of accuracy, precision, recall, and the execution time. The propose of this study to give an insight and the use of big data, machine learning, and deep learning from 6 disaster area which is early warning damage, damage assessment, monitoring and detection, forecasting and predicting, and post-disaster coordination, and response, and long-term risk assessment and reduction. But also give the data source that the previous research

II. RELATED WORK

Based on existing research, Manzhu Yu researches on natural disasters which are divided into several parts. One of them is categorizing the articles based on major data sources which consist of satellite imagery, aerial imagery, and videos from unmanned Detection and Ranging (LiDAR), simulation, spatial data, crowdsourcing, social media, and mobile. Disaster management phases consist of 4 main parts Mitigation/prevention, Preparedness, Response, and Recovery. But in Manzhu's research, they divided the 4 parts of the Disaster Management Phases into 6 phases, which include early warning damage, damage assessment, monitoring and detection, forecasting and predicting, and post-disaster coordination, and response, and long-term risk assessment and reduction. Based on Goswami's research, he conducted a review of natural disasters in a case study in India. This research classifies the objectives of the tasks, namely Prediction, Detection, and Disaster management strategies.

Based on this data analysis, we could determine the pattern and solve the problem. By predicting the event or maybe use it for early detection. To solve this kind of problem, we need to use not only data sources but also the type of model/ algorithm to train and test the data. Most of the review focuses on the only area on the disaster management phase and the data source from the result of Manzhu study he didn't explain about the performance on the model/algorithm in the reviews. The most important part of the early detection and prediction is to analyze the data source and the model they use to solve the problem.

The following table 1 is a table of previous researchers that focuses on the area topic in disaster management and natural disaster using big data and machine learning approach.

TABLE I. RELATED WORK ON THE PREVIOUS REVIEW RESEARCH PAPER

Categories	Data source	Citation
• Early warning damage • Damage assessment • Monitoring and detection • Forecasting and predicting • Post-disaster coordination and response • Long-term risk assessment and reduction	• Satellite imagery • Aerial imagery and videos from unmanned Detection and Ranging (LiDAR) • Simulation, spatial data • Crowdsourcing • Social media	[6]
• Prediction, • Detection • Disaster management strategies	• Hydrological data, • Meteorological data • twitter	[4]

Based on previous research such as the table above in the study, Manzhu Yu divides into 6 categories, and Goswami divides into 3 groups. Based on these categories, Manzhu's research was more detailed because the category was taken based on the concept of Disaster management phases. In this study, we will adopt 6 categories that have been created by Manzhu, and combine them with Goswami research to provide models/techniques that have been given along with the data used.

III. METHOD

The paper that will be chosen in the first review process is a paper that publishes on international proceedings or an article published in international journals. So from that, the article to be taken must use English. After that, the second requirement is the paper chosen is a paper published in the last five years. That is a paper published from 2014 until 2019. After that, the article was taken based on science direct[7], springer open-source [8], IEEE [9], google scholar [10], and research gate [11]. The search process for the paper by considering the abstract keyword is related to any natural disaster. The article must also discuss the fields of big data, machine learning, or deep learning. The following figure 1 explains the flow chart on how to choose the article that will be reviewed in this study.

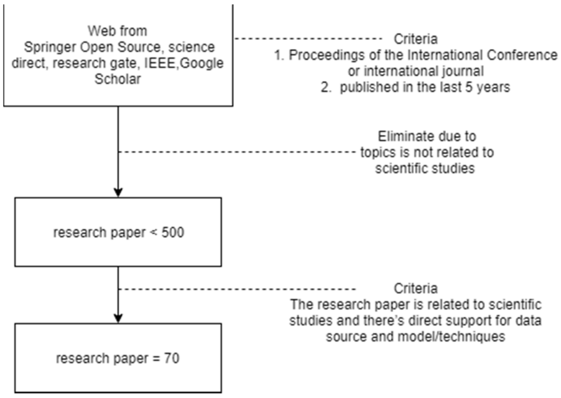

Fig. 1. Step Of Experimental Method

In figure 1 it explains the steps to getting the journal so that it is first selected based on the appropriate topic. Paper taken must be in the area of big data, machine learning, and deep learning. The chosen paper must have data source information used in the study. Along with the model/technique/algorithm used. In this paper, we will classify reviews based on frequently used data sources, often used algorithms so that it can compare and provide evaluations of datasets that are accurately used in big data research, machine learning, or deep learning. Besides that, it categorizes based on Disaster management phases that have been modified by human management, which consists of 6 categories [6]. There are 6, namely early warning damage, damage assessment, monitoring and detection, forecasting and predicting post-disaster coordination, and response, and long-term risk assessment and reduction. Its long-term risk assessment and reduction function is for long-term disaster risk reduction. Forecasting and Predicting used to predict natural disasters that will occur. Monitoring and detection to detect or monitor disasters. Early warning Damage is used to receive relevant and timely information. Damage Assessment is an actual disaster evaluation event caused by a natural disaster event. Post-disaster recovery planning is defined as developing a set of strategies to help rebuild after a disaster occurs

IV. RESULT

After choosing the paper that matches the criteria above, then it is to group the article based on the categories that have been made by Manzhu. After that, classify the review paper on the data source used during natural disaster research. After that, it is classifying review paper based on the model/ algorithm used in the study as with the Goswami study. In Manzhu's study, there are six categories of Disaster management phases. These categories are, among others, Long-term risk assessment and reduction, Forecasting and Predicting, Monitoring and detection, Early Warning Damage, Damage Assessment, Post-disaster Coordination, and Response. The following are the results of the paper following the six categories in Table 2.

TABLE II. DISASTER MANAGEMENT PHASES

Categories	Citation
Long-term risk assessment and reduction	[12],[13],[14],[15][16]
Forecasting and Predicting	[17],[18],[19],[20],[21][22][23][24][25][26]
Monitoring and detection	[27],[28],[29],[30],[31],[32],[33],[34],[35],[36][37][38][39][40][41][42][43][44][45][46]

Early warning Damage	[47],[48],[49],[50][51][52][53][54]
Damage Assessment	[55],[15],[56],[15],[55],[56][57][58]
Post-disaster Coordination and Response	[59],[60][61][62][63]

Based on the results of table 2, part of the research is often done in the area of monitoring and detection. The monitoring and detection section is a component that is usually researched because based on the results of the paper to obtain data sources in the process of monitoring and detection is straightforward. Data sources that are often used are twitter data and satellite data. The data retrieval process is easy to take because it is open source data so that the research is a research that is often carried out in the area of a natural disaster.

Research that is rarely done is post-disaster coordination and response. The big data and machine learning in this area are still very uncommon. Because to conduct this area of topics, the expert has to actually examine the disaster area and carry out the infrastructure process from the damaged areas from natural disasters. Based on the results of table 2, it can be concluded that the most suitable area for disaster risk in the natural disaster area is Monitoring and detection and Forecasting and Predicting. the area is the most appropriate area to be applied in a natural disaster area because of the domain is the most suitable area to be used in the area of natural disasters because the data provided has already had 5'v characteristics, namely Velocity, Volume, Value, Variety, and Veracity.

After categorizing according to the 6 categories, then classify based on data sources used in research on big data, machine learning, and deep learning. The following data source is used for research in table 3.

TABLE III. TABLE DATA SOURCE FOR THE DISASTER AREA

Data source	Citation
Satellite imagery	[35][30][64][32][28][42][26][45]
Social media	[47][65][66][56][55][15][67][21][49] [31][36][37][16][57][23][61][52][41][62][53][46][54][63]
Crowdsourcing data	[65][17][21][20][29][16][22][39][58][24][25][44]
Aerial imagery	[59][38]
Online News report	[51][43]

In table 3, there is a data source used by previous research. The data sources are satellite imagery, social media, crowdsourcing data, and aerial imagery. In the above results, the data source that is often used is social media, especially Twitter. For data sources that use Twitter more, they use analysis sentiment, as Ghazaleh's research uses data source social media for sentiment analysis [68]. Zahra's research uses twitter data sources such as Ghazaleh's research for analysis sentiment [60]. On crowdsourcing data that is often taken is data provided by USGS (United States Geological Survey). One paper that uses these data is Mendoza's research [21] and Resch's research [15].The satellite image is also used as a data source. The satellite image came from NASA like the Gokaraju study [30]. Besides that, there is satellite imagery that uses SAR (Synthetic Aperture Radar) to obtain data sources such as Wieland's research [32]. Base on the results in table 3 concluded that the data source that is often used is social media data, especially Twitter. social media data, for example, twitter to access data on tweets is accessible. Based on Nguyen's research, he uses steaming API for extracting the tweets data with a keyword "quake," "tsunami," and "earthquake." The study also explains that they could download 18gb twitter bases on "quake," "tsunami," and "earthquake" keyword from September 11 to December 25, 2016 [36]. Most of the satellite data, crowdsourcing, and areal image took at least two years to generate the data to understand and predict the pattern from the image. For example, in Thibaut's research, he uses data crowdsourcing from the Oklahoma Geological Survey (OGS), which only generate data 2021 seismic events from 2014 to 2016 [29]. We could conclude that social media data which is twitter data is easy to extract with the help of streaming API, not only it's easy but the volume of data is enormous we could get 18gb data on a tweet with a small period of time which in this case it's 4 month rather using crowdsourcing data which take approximately 2 years. After classifying the paper based on the data source, then classify the article based on the model / technique / algorithm used in the study. There are several types of models used. Table 4 discusses the model/algorithm used for the disaster area.

TABLE IV. MODEL / TECHNIQUES USED

Model/techniques	Citation
SVSA (Support Vector Selection Adaptation)	[28]
Decision Tree	[34][63]
Random Forest	[35][65][69][34][13][61]
Support Vector Machine	[65][67][32][49][28] [34],[13][16][39][58][51][61][42] [25][62][26]
Boosted Regression Trees (BRT)	[35]
Bayesian Networks	[14]
K-Means	[12][39]
Naïve Bayes	[51][61][43]
Naïve Bayes Classifiers	[66][69][49][57]
CNN (Convolutional Neural Network)	[55][36][29][38][44][45][54]
Latent Dirichlet Allocation (LDA)	[15][37][61][41]
ANN (artificial neural networks)	[27][17][20][70][22][42]
C4.5	[51][63]
Neural Clouds	[40]
Logistic Regression	[42]

Based on the results of table 4, there are several types of algorithms used. For the algorithm, Support Vector Machine is used in research to detect changes in single- and multi-temporal X- and L-bands using SAR (Synthetic Aperture Radar) image [32]. Besides that, it is also used in the research of earthquake image classification process. In this study, using SVM to compare whether the SVSA model (Support Vector Selection Adaptation) can provide increased performance at the computational time, along with the new model does not require kernels [28]. Also, the model that is often used is ANN (artificial neural network). The study used ANN for earthquake early detection, also known as precursor [27]. Also, ANN was also used to predict earthquakes magnitude in Tokyo [20]. Base on the result, we conclude that CNN works best for earthquake detection. Not only for the accuracy, but recall, precision, and the execution time is much faster. But the downside of this algorithm/model you need a lot of data. For example, Perol's research uses CNN for earthquake detection. He compares it with the previous study and that his model execution time only took 1 minute 1 second then the other 2 models, which took 9 days 13 hours and 48 hours to execute [29]. Not only in Perol's research but also Nguyen uses CNN for earthquake detection perform better not only in the accuracy, but also recall, pre, and score. They compare the CNN model with SVM. The reason why CNN performs better than SVM it's because of the amount of data they use. Base on that research, they use 18gb tweets of data. SVM performs poorly due to the amount of data. SVM works really well on smaller data rather than big amounts of data. They use the CNN model to avoid underfitting and overfitting problem [36]. The neural cloud is an algorithm that uses the combination of an Advanced K-Means (AKM) clustering algorithm and an extended Radial Basis Functions (RBF). The neural cloud is a cloud detection algorithm. Research by Pyayt *et al.* uses the algorithm for monitoring and flood protection [40].

Based on the result of table 2 we could see the result from figure 2.

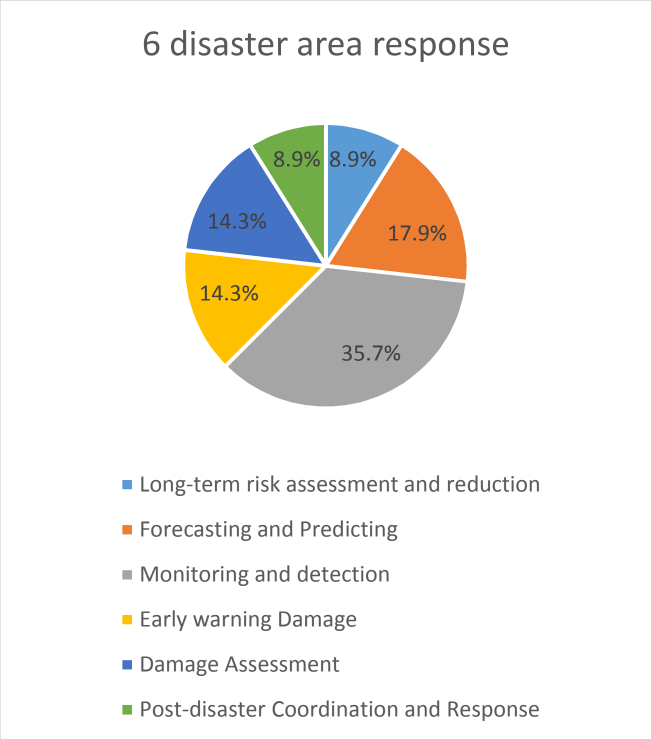

Fig. 2. Result Chart

Figure 2 explains the most common research and the most uncommon research regarding the 6-disaster area. From the result we conclude that for the 6-disaster area the most common area was monitoring and detection. The reason why this area is so common because monitoring a disaster has a big impact on the disaster management phase because if we could monitor the activities of the natural disaster, we could recognize the pattern and eventually we could predict the cause of the natural disaster. From the data source area, the most common research was social media data. The reason why it was so common because the data extraction only needed a twitter API and the data that the twitter produces up to 18 GB of twitter data. The most important part of big data and machine learning was to extract a large amount of data that has 5'v characteristics, namely Velocity, Volume, Value, Variety, and Veracity.

Base on the result of tables 2,3, and 4 we could conclude that big data and machine learning could help overcome the disaster management area which includes early warning damage, damage assessment, monitoring and detection, forecasting and predicting post-disaster coordination, and response, and long-term risk assessment and reduction. From this area, we could choose the right time of data source and model which one performs better for the set of problems that we choose.

V. Conclusion

Based on previous results, it can be concluded that with big data, machine learning, or deep learning can help in 6 areas of Disaster management. These areas are early warning damage, damage assessment, monitoring and detection, forecasting and predicting post-disaster coordination, and response, and long-term risk assessment and reduction. By classifying based on data sources and models/algorithms that are often used by previous researchers can help researchers who will research in the field of a natural disaster from the result we can conclude that the most common area that utilizes big data and machine learning Monitoring and detection. Monitoring and detection is the most common research in the 6 disaster management area due to the increase of data collection. The data collection on monitoring and detection is easy to extract with twitter data. And the amount of data that had been extracted with twitter API was 18gb twitter data. And the data has 5'v characteristics, namely Velocity, Volume, Value, Variety, and Veracity. For the future, we will examine the model/ algorithm not only base on the level of accuracy, precision, recall, and the execution time but also the time needed to build the model itself.

References

[1] "Number of reported disasters by type." [Online]. Available: https://ourworldindata.org/natural-disasters.

[2] Tuswadi and T. Hayashi, "Disaster Prevention Education in Merapi Volcano Area Primary Schools: Focusing on Students' Perception and Teachers' Performance," *Procedia Environ. Sci.*, vol. 20, pp. 668–677, 2014.

[3] "2015_43291_Sendaiframeworkfordrren_Disaster Reducton 2015-2030," 2015.

[4] S. Goswami, S. Chakraborty, S. Ghosh, A. Chakrabarti, and B. Chakraborty, "A review on application of data mining techniques to combat natural disasters," *Ain Shams Eng. J.*, vol. 9, no. 3, pp. 365–378, 2018.

[5] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. Ullah Khan, "The rise of 'big data' on cloud computing: Review and open research issues," *Inf. Syst.*, vol. 47, pp. 98–115, 2015.

[6] M. Yu, C. Yang, and Y. Li, "Big Data in Natural Disaster Management: A Review," *Geosciences*, vol. 8, no. 5, p. 165, 2018.

[7] P. Sciences, "science direct," 2018. [Online]. Available: https://www.sciencedirect.com/.

[8] Springer, "springeropen," *Technolo*, 2018. [Online]. Available: https://www.springeropen.com/journals.

[9] IEEE, "IEEE," 2018. [Online]. Available: http://ieeexplore.ieee.org.

[10] Google Scholar, "Google Scholar," 2018. [Online]. Available: https://scholar.google.com/intl/en/scholar/about.html.

[11] L. Connect, "ResearchGate," 2018. [Online]. Available: https://www.researchgate.net/.

[12] B. Supriyadi, A. P. Windarto, T. Soemartono, and . M., "Classification of Natural Disaster Prone Areas in Indonesia using

K-Means," *Int. J. Grid Distrib. Comput.*, vol. 11, no. 8, pp. 87–98, 2018.

[13] C. Geiß *et al.*, "Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques," *ISPRS J. Photogramm. Remote Sens.*, vol. 104, pp. 175–188, 2015.

[14] J. Wang, X. Gu, and T. Huang, "Using Bayesian networks in analyzing powerful earthquake disaster chains," *Nat. Hazards*, vol. 68, no. 2, pp. 509–527, 2013.

[15] B. Resch, F. Usländer, and C. Havas, "Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment," *Cartogr. Geogr. Inf. Sci.*, vol. 45, no. 4, pp. 362–376, 2018.

[16] R. Kiatpanont, U. Tanlamai, and P. Chongstitvatana, "Extraction of actionable information from crowdsourced disaster data," *J. Emerg. Manag.*, vol. 14, no. 6, pp. 377–390, 2016.

[17] K. M. Asim, F. Martínez-Álvarez, A. Basit, and T. Iqbal, "Earthquake magnitude prediction in Hindukush region using machine learning techniques," *Nat. Hazards*, vol. 85, no. 1, pp. 471–486, 2017.

[18] N. D. Hoang and D. T. Bui, "Predicting earthquake-induced soil liquefaction based on a hybridization of kernel Fisher discriminant analysis and a least squares support vector machine: a multi-dataset study," *Bull. Eng. Geol. Environ.*, vol. 77, no. 1, pp. 191–204, 2018.

[19] K. M. Asim, A. Idris, T. Iqbal, and F. Martínez-Álvarez, "Earthquake prediction model using support vector regressor and hybrid neural networks," *PLoS One*, vol. 13, no. 7, pp. 1–22, 2018.

[20] G. Asencio-Cortés, F. Martínez-Álvarez, A. Troncoso, and A. Morales-Esteban, "Medium–large earthquake magnitude prediction in Tokyo with artificial neural networks," *Neural Comput. Appl.*, vol. 28, no. 5, pp. 1043–1055, 2017.

[21] M. Mendoza, B. Poblete, and I. Valderrama, "Nowcasting earthquake damages with Twitter," *EPJ Data Sci.*, vol. 8, no. 1, 2019.

[22] F. Martínez-Álvarez, A. Troncoso, H. Quintián, and E. Corchado, "Rainfall prediction: A Deep Learning approach," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9648, no. June, 2016.

[23] O. Aulov and M. Halem, "Human sensor networks for improved modeling of natural disasters," *Proc. IEEE*, vol. 100, no. 10, pp. 2812–2823, 2012.

[24] A. Baldominos, E. Albacete, Y. Saez, and P. Isasi, "A scalable machine learning online service for big data real-time analysis," *IEEE SSCI 2014 - 2014 IEEE Symp. Ser. Comput. Intell. - CIBD 2014 2014 IEEE Symp. Comput. Intell. Big Data, Proc.*, pp. 1–8, 2015.

[25] X. Yu, "Disaster prediction model based on support vector machine for regression and improved differential evolution," *Nat. Hazards*, vol. 85, no. 2, pp. 959–976, 2017.

[26] M. N. Jebur, B. Pradhan, and M. S. Tehrany, "Manifestation of LiDAR-derived parameters in the spatial prediction of landslides using novel ensemble evidential belief functions and support vector machine models in GIS," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 8, no. 2, pp. 674–690, 2015.

[27] B. A. S. Aji, T. H. Liong, and B. Muslim, "Detection precursor of sumatra earthquake based on ionospheric total electron content anomalies using N-Model Articial Neural Network," *2017 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2017*, vol. 2018-Janua, pp. 269–276, 2018.

[28] G. Tas, O. K. Ersoy, and M. E. Kamas, "SUPPORT VECTOR SELECTION AND ADAPTATION FOR CLASSIFICATION OF Istanbul Technical Univ , Informatics Institute , Istanbul , Turkey , e-mail : gulsen@be.itu.edu.tr Purdue University , School of Electrical and Computer Eng ., Indiana , USA , e-mail : ersoy@p," *Computer (Long. Beach. Calif).*, pp. 851–854, 2009.

[29] T. Perol, M. Gharbi, and M. Denolle, "Convolutional neural network for earthquake detection and location," *Sci. Adv.*, vol. 4, no. 2, pp. 2–10, 2018.

[30] B. Gokaraju, A. C. Turlapaty, D. A. Doss, R. L. King, and N. H. Younan, "Change detection analysis of tornado disaster using conditional copulas and Data Fusion for cost-effective disaster management," *2015 IEEE Appl. Imag. Pattern Recognit. Work. AIPR 2015*, no. October, 2016.

[31] T. Sakaki, M. Okazaki, and Y. Matsuo, "Tweet analysis for real-time event detection and earthquake reporting system development," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 4, pp. 919–931, 2013.

[32] M. Wieland, W. Liu, and F. Yamazaki, "Learning change from Synthetic Aperture Radar images: Performance evaluation of a Support Vector Machine to detect earthquake and tsunami-induced changes," *Remote Sens.*, vol. 8, no. 10, 2016.

[33] A. J. Cooner, Y. Shao, and J. B. Campbell, "Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 Haiti earthquake," *Remote Sens.*, vol. 8, no. 10, 2016.

[34] W. Li, N. Narvekar, N. Nakshatra, N. Raut, B. Sirkeci, and J. Gao, "Seismic data classification using machine learning," *Proc. - IEEE 4th Int. Conf. Big Data Comput. Serv. Appl. BigDataService 2018*, pp. 56–63, 2018.

[35] S. Park, J. Im, E. Jang, and J. Rhee, "Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions," *Agric. For. Meteorol.*, vol. 216, pp. 157–169, 2016.

[36] V. Q. Nguyen, H. J. Yang, K. Kim, and A. R. Oh, "Real-Time Earthquake Detection Using Convolutional Neural Network and Social Data," *Proc. - 2017 IEEE 3rd Int. Conf. Multimed. Big Data, BigMM 2017*, pp. 154–157, 2017.

[37] M. Imran, S. Elbassuoni, C. Castillo, F. Diaz, and P. Meier, "Practical extraction of disaster-relevant information from social media," *WWW 2013 Companion - Proc. 22nd Int. Conf. World Wide Web*, pp. 1021–1024, 2013.

[38] S. N. K. B. Amit and Y. Aoki, "Disaster detection from aerial imagery with convolutional neural network," *Proc. - Int. Electron. Symp. Knowl. Creat. Intell. Comput. IES-KCIC 2017*, vol. 2017-Janua, no. July 2018, pp. 239–245, 2017.

[39] M. Bria, A. Baban, and D. Dumitrascu, "A Region Thesaurus Approach for High-Level Concept Detection in the Natural Disaster Domain," *An Interdiscip. J.*, vol. 16, no. 3, pp. 423–452, 2013.

[40] A. L. Pyayt, I. I. Mokhov, B. Lang, V. V. Krzhizhanovskaya, and R. J. Meijer, "Machine learning methods for environmental monitoring and flood protection," *World Acad. Sci. Eng. Technol.*, vol. 78, no. 6, pp. 118–123, 2011.

[41] K. Kireyev, L. Palen, and K. Anderson, "Applications of topics models to analysis of disaster-related twitter data," *NIPS Work. Appl. Top. Model. Text Beyond*, p. , 2009.

[42] I. Yilmaz, "Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine," *Environ. Earth Sci.*, vol. 61, no. 4, pp. 821–836, 2010.

[43] M. Scharkow, "Thematic content analysis using supervised machine learning: An empirical evaluation using German online news," *Qual. Quant.*, vol. 47, no. 2, pp. 761–773, 2013.

[44] S. Ahmad, K. Ahmad, N. Ahmad, and N. Conci, "Convolutional neural networks for disaster images retrieval," *CEUR Workshop Proc.*, vol. 1984, pp. 2016–2018, 2017.

[45] K. Ahmad, P. Konstantin, M. Riegler, N. Conci, and P. Holversen, "CNN and GAN based satellite and social media data fusion for disaster detection," *CEUR Workshop Proc.*, vol. 1984, no. September, pp. 15–18, 2017.

[46] M. Maldonado, D. Alulema, D. Morocho, and M. Proano, "System for monitoring natural disasters using natural language processing in the social network Twitter," *Proc. - Int. Carnahan Conf. Secur. Technol.*, no. October, 2017.

[47] D. Buscaldi and I. Hernandez-Farias, "Sentiment Analysis on Microblogs for Natural Disasters Management," *Proc. 24th Int. Conf. World Wide Web - WWW '15 Companion*, pp. 1185–1188, 2015.

[48] G. Shalunts, G. Backfried, and K. Prinz, "Sentiment Analysis of German Social Media Data for Natural Disasters," *11th Int. Conf. Inf. Syst. Cris. Response Manag.*, no. May, pp. 752–756, 2014.

[49] L. Maceda, J. Llovido, and A. Satuito, "Categorization of Earthquake-Related Tweets Using Machine Learning Approaches," no. 1, pp. 229–232, 2019.

[50] B. Truong, C. Caragea, A. Squicciarini, and A. H. Tapia,

"Identifying valuable information from Twitter during natural disasters," *Proc. ASIST Annu. Meet.*, vol. 51, no. 1, 2014.

[51] T. Valero, M. Gómez, and V. Pineda, "Using Machine Learning for Extracting Information from Natural Disaster News Reports," *Comput. y Sist.*, vol. 13, no. 1, pp. 33–44, 2009.

[52] A. H. SALANS, J. A. SCHACK, and L. N. KATZ, "Natural Language Processing to the Rescue? Extracting 'Situational Awareness' Tweets During Mass Emergency," *Circulation*, vol. 2, no. 6, pp. 900–906, 1950.

[53] I. Varga *et al.*, "Aid is out there: Looking for help from tweets during a large scale disaster," *ACL 2013 - 51st Annu. Meet. Assoc. Comput. Linguist. Proc. Conf.*, vol. 1, pp. 1619–1629, 2013.

[54] C. Caragea, A. Silvescu, and A. H. Tapia, "Identifying informative messages in disaster events using Convolutional Neural Networks," *Proc. Int. ISCRAM Conf.*, no. May, 2016.

[55] D. T. Nguyen, F. Ofli, M. Imran, and P. Mitra, "Damage Assessment from Social Media Imagery Data During Disasters," no. December, pp. 569–576, 2017.

[56] S. Cresci, M. Tesconi, A. Cimino, and F. Dell'Orletta, "A Linguistically-driven Approach to Cross-Event Damage Assessment of Natural Disasters from Social Media Messages," pp. 1195–1200, 2016.

[57] K. Zahra, F. O. Ostermann, and R. S. Purves, "Geographic variability of Twitter usage characteristics during disaster events," *Geo-Spatial Inf. Sci.*, vol. 20, no. 3, pp. 231–240, 2017.

[58] Z. Huang, J. Zhou, L. Song, Y. Lu, and Y. Zhang, "Flood disaster loss comprehensive evaluation model based on optimization support vector machine," *Expert Syst. Appl.*, vol. 37, no. 5, pp. 3810–3814, 2010.

[59] F. Ofli *et al.*, "Combining Human Computing and Machine Learning to Make Sense of Big (Aerial) Data for Disaster Response," *Big Data*, vol. 4, no. 1, pp. 47–59, 2016.

[60] J. Peterson and B. Argo, "Tweedr: Mining Twitter to Inform Disaster Response," vol. 12, no. 4, pp. 354–358, 2014.

[61] M. Imran, P. Mitra, and C. Castillo, "Twitter as a lifeline: Human-annotated Twitter corpora for NLP of crisis-related messages," *Proc. 10th Int. Conf. Lang. Resour. Eval. Lr. 2016*, pp. 1638–1643, 2016.

[62] J. R. Ragini, P. M. R. Anand, and V. Bhaskar, "Big data analytics for disaster response and recovery through sentiment analysis," *Int. J. Inf. Manage.*, vol. 42, no. May, pp. 13–24, 2018.

[63] M. Avvenuti, F. Del Vigna, S. Cresci, A. Marchetti, and M. Tesconi, "Pulling Information from social media in the aftermath of unpredictable disasters," *Proc. 2015 2nd Int. Conf. Inf. Commun. Technol. Disaster Manag. ICT-DM 2015*, no. November, pp. 258–264, 2016.

[64] Y. Liu and L. Wu, "Geological Disaster Recognition on Optical Remote Sensing Images Using Deep Learning," *Procedia Comput. Sci.*, vol. 91, no. Itqm, pp. 566–575, 2016.

[65] A. Cobo, D. Parra, and J. Navón, "Identifying Relevant Messages in a Twitter-based Citizen Channel for Natural Disaster Situations," 2015.

[66] H. Li *et al.*, "Twitter Mining for Disaster Response: A Domain Adaptation Approach," *12th Int. Conf. Inf. Syst. Cris. Response Manag. (ISCRAM 2015)*, p. [1-7], 2015.

[67] S. Cresci, A. Cimino, F. Dell'Orletta, and M. Tesconi, "Crisis Mapping During Natural Disasters via Text Analysis of Social Media Messages," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9419, pp. 250–258, 2015.

[68] G. Beigi, R. Maciejewski, and H. Liu, "Sentiment Analysis and Ontology Engineering," vol. 639, no. February 2016, 2016.

[69] A. Fikri Aji, I. P. E. S. Putra, P. Mursanto, and S. Yazid, "Can smartphones be used to detect an earthquake? Using a machine learning approach to identify an earthquake event," *8th Annu. IEEE Int. Syst. Conf. SysCon 2014 - Proc.*, pp. 72–77, 2014.

[70] G. Rajguru, Y. S. Bhadauria, and S. Mukhopadhyay, "Estimation of Earthquake Source Parameters Using Machine Learning Techniques," *2018 9th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2018*, pp. 1–7, 2018.