Analytics and Evolving Landscape of Machine Learning for Emergency Response

Minsung Hong · Rajendra Akerkar

Abstract The advances in information technology have had a profound impact on emergency management by making unprecedented volumes of data available to the decision makers. This has resulted in new challenges related to the effective management of large volumes of data. In this regard, the role of machine learning in mass emergency and humanitarian crises is constantly evolving and gaining traction. As a branch of artificial intelligence, machine learning technologies have the out-standing advantages of self-learning, self-organization, and self-adaptation, along with simpleness, generality and robustness. Although these technologies do not perfectly solve issues in emergency management, and have been showed to can greatly improve the capability and effectiveness of emergency management. The purpose of this chapter is to discuss a hybrid crowdsourcing and real-time machine learning approaches to rapidly process large volumes of data for emergency response in a time-sensitive manner. We review the application of machine learning techniques to support the decision-making processes for the emergency or crisis management and discuss their challenges. Additionally, we discuss the challenges and opportunities of the machine learning approaches and intelligent data analysis to distinct phases of emergency management. Based on the literature review, we observe a trend to move from narrow in scope, problem-specific applications of data mining and machine learning to solutions that address a wider spectrum of problems, such as situational awareness and real-time threat assessment using diverse streams of data. In particular, this chapter also focuses on crowdsourcing approaches with machine learning to achieve better understanding and decision support during a disaster, and we discusses the issues on the approaches in terms of data analysis. Several examples of the tweet related to emergency are discussed to more deeply contemplate the issues.

Minsung Hong
Vestlandsforsking, Box 163, NO-6851 Sogndal, Norway
Tel.: +47-908-03-233
E-mail: msh@vestforsk.no

Rajendra Akerkar
Vestlandsforsking, Box 163, NO-6851 Sogndal, Norway
Tel.: +47-916-85-607
E-mail: rak@vestforsk.no

1 Introduction

In the contemporary society, a variety of emergencies take place more and more frequently. Necessarily, a considerable number of emergency incidents have threatened to human life, environmental protection, social stability, and even political relationship of all countries around the world [1]. In this regard, sociologists of emergencies have been working to define emergency for decades. There is a broad consensus that emergencies are social phenomena, characterized by a disruption of routine and of social structure, norms, and/or values. It implies that the severity of a emergency is more related to the extent of the disruption of social life as aspects of governments, business, and individuals, than the measurable magnitude of the hazard [2]. Therefore, the negative effects of emergencies emphasize the need to improve the emergency management capability and strengthen the security for all countries in the world.

1.1 Emergency Management

The definition of emergency management can be extremely broad and all-encompassing. Unlike other, more structured disciplines, it has expanded and contracted in response to events, congressional desires, and leadership styles [3]. Some representative definitions in the literature are as follows:

- According to definition of the Federal Emergency Management Agency (FEMA) in USA, the process of emergency management consists of preparing for, mitigating, responding to, and recovering from an emergency when a disaster arises [4].
- More modern emergency management involves processes to apply modern technologies and management methods to effectively and efficiently monitor, response to, control, and process events, by integrating various social resources and analyzing scientifically the cause [5].
- A simple definition for emergency management is "a discipline that deals with risk and risk avoidance." Risk represents a wide range of issues, and the range of situations that might possibly involve emergency management or the emergency management system is vast. This supports the premise that emergency management is essential to the security of everyones daily lives and should be integrated into daily decisions and not just called on during times of disasters [3].

In short, emergency management is a complex and multifaceted task that involves a variety of management activities from managers and stakeholders when emergency is not only arising but also the before and after of emergency, so as to prevent the occurrence of unexpected events, to reduce the social damages, and to mitigate the impacts. Based on the definitions of the emergency management,
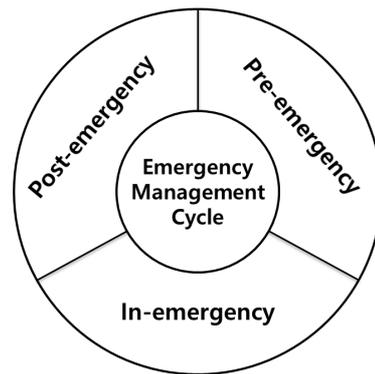
Fig. 1 The lifecycle of emergency management

the evolution of an emergency can be distinguished as three stages, namely pre-emergency, in-emergency, and post-emergency, as shown in Figure 1. Chen et al. described emergency management as a '4R' process, namely reduction, readiness, response and recovery. Reduction is referred to the pre-emergency phase, readiness and response belong to the in-emergency phase, and recovery is referred to the post-emergency phase. In each phase, the outcome of decision-making impacts substantially the evolution of events and the effectiveness of emergency management [5].

As aforementioned, emergency management is a multifaceted process to prevent, reduce, respond to, and recover from the impact of the emergency on the society. Because of the scale of events, emergency response requires the participation and cooperation of multiple organizations (e.g., government, public and private). This emphasizes the need for efficient and effective decision support systems, as it is practically impossible for a human decision maker to understand and manage the complexity of the situation. Instead, problems such as situational awareness [6] and building a common operating picture, shared among multiple actors who often have only partial view of the situation, are becoming some of the most urgent needs of emergency management [7].

However, the emergency data used to these decision support systems arises the problems of delivering repetitive information and information overload [8]. Therefore, to improve the capability and effectiveness of emergency management, machine learning techniques have been proposed.

1.2 Machine Learning

Emergency management is concerned not only with predicting the course and consequences of disasters, but also mitigating those undesired consequences. This process is undoubtedly a challenging task by the unprecedented volumes of data (e.g., forecast, news, web pages, data of social network service, and sensing data) and the pressure of time [7]. Machine learning techniques have been proven to successfully support the decision making processes in managing many complex problems. In that sense, emergency management is no exception; however, it presents a variety

of challenge to machine learning techniques for the emergency management. In this section, we briefly introduce the category of machine learning.

Machine learning has progressed dramatically over the past two decades, from laboratory curiosity to a practical technology in widespread commercial use [9]. Machine learning and data mining often use the same methods and overlap significantly, but while machine learning focuses on prediction, data mining concentrates on the discovery of (previously) unknown properties in the data. According to depending on whether there is labeled instance which is consists of label and data, Machine learning are typically classified into three categories as follows:

- **Supervised learning:** Supervised machine learning makes predictions about future instances using externally supplied instances that consist of values and a label. It's goal is to build a concise model of the distribution of class labels, and then a classifier based on the model is used to assign class labels to the testing instances [10].
- **Unsupervised learning:** Unsupervised learning is inferring directly the properties of this probability density without the help of externally provided instances providing correct label or degree-of-error for each observation [11]. There are representative algorithms like Apriori algorithm, K-means, and so on.
- **Reinforcement learning:** Reinforcement learning deduces labels of instances with a dynamic environment. There are two main strategies. The first is to search in the space of behaviors in order to find one that performs well in the environment, such as genetic algorithms, and the second is to use statistical techniques and dynamic programming methods to estimate the utility of taking actions in states of the world [12].

1.3 Scope and Organizations

This chapter focuses on the application of machine learning techniques to support the decision-making processes for the emergency or crisis management. We start with the data-driven methodologies within the frameworks of machine learning and their roles and challenges in supporting different phases of emergency management. We then discuss the characteristics of disaster data akin to 5 Vs of big data and summary various applications cases of big data analysis. Next, with respect to emphasizing the advance of the social media, we focus on reviewing the crowdsourcing approaches with machine learning in emergency management, and issues of the approaches are discussed in terms of the data analysis. Last, several examples of the tweet related to emergency are discussed to more deeply contemplate the challenges and opportunities.

Existing survey papers for machine learning techniques in emergency management have reviewed according to categories of the machine learning, or have considered only some part of emergency tasks. Whereas, in this chapter, we review the approaches of machine learning, have been proposed from 2010 to current, along each task of emergencies. Therefore, we believe that readers can easily find topics related to their interests and compare with existing approaches to little more concretely grasp potentiality of their methods.

2 Applications of Machine Learning in Emergency Response

As discussed earlier, within the present-day emergency management, the immediate and accurate decision making more and more relies on the capability of data analysis and processing. Therefore, there is an urgent need to enhance the machine learning functionality of emergency management, such as, to develop scalable and real-time algorithms for time-sensitive decisions, to integrate structured, unstructured, and semi-structured data [1]. In this section, we attempt to introduce the tasks of machine learning in each phase of emergency management and review the challenges and benefits of various machine learning techniques for the emergency management.

2.1 Machine Learning Techniques for Emergency Management Cycles

Successful emergency management requires a variety of tasks based on various technologies of machine learning within across the board three phases mentioned in Section 1.1. Figure 2 shows the tasks related to machine learning for each phase of the emergency management as follows: (1) predicting the occurrence of potential events and discovering the early warning signs; (2) during the emergency, detecting the events occurred and tracking change of the incidents, and recognizing situations of people, supply, and so on; (3) evaluating the loss caused by incidents and the execution of response, and simultaneously adjusting volunteer efforts based on crowdsourcing to recover from an emergency.
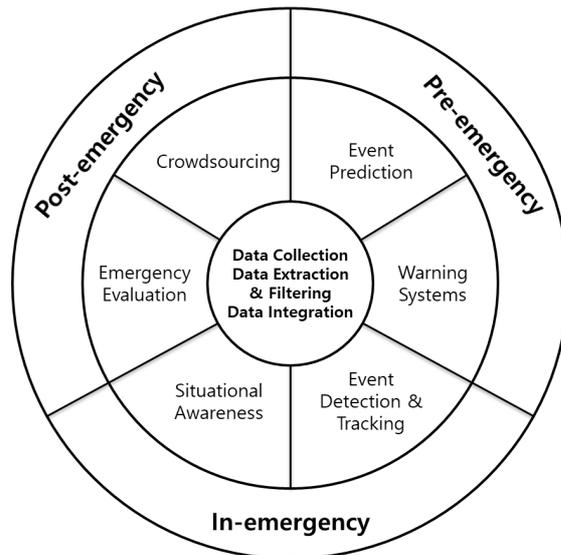

Fig. 2 Tasks related to machine learning in emergency management

– **Event Prediction:** it is forecasting emergency using technology and interpretation methods, and can be achieved by extracting features or pattern.

Although there is no prediction method with perfect accuracy, early detection of natural disasters reduces hazards in nearby locations [13].

- **Warning Systems:** to detect impending emergency can give that information to people at risk, and enable those in danger to make decisions and early take action [14]. These systems have improved drastically in recent years but they are not perfect yet.
- **Event Detection & Tracking:** Most systems based on machine learning during crises start with detecting and tracking events. The events are mainly associated with a specific time and location [15]. However, due to the online nature of collected data, events may or may not be necessarily associated with physical locations.
- **Situational Awareness:** it provides more deep recognition of events in emergency using social media data related to specific information (e.g., caution, advice, donations, casualties and damage) and smart-phones which typically mount various sensors such as camera, GPS, and accelerometer [7].
- **Emergency Evaluation:** it is one of critical and complex tasks in emergency management [1]. In post-emergency, the activity outcomes (e.g., loss of resources, recoverability, performance and social influence) for current emergency should be measured to suppress the deterioration of next emergency.
- **Crowdsourcing:** this task is a sourcing model in which organizations use predominantly advanced Internet technologies to harness the efforts of a virtual crowd to perform specific organizational tasks [16]. It may allow to immediately collect the statuses and requirements of people after an emergency, and analyzed and categorized the data collected to support relief operations.

Like this, there are various tasks for emergencies, and machine learning techniques have been applied into each task to improve the effective and efficient emergency management. Here, the approaches of machine learning will be reviewed for each task.

2.2 Event Prediction

There were many tries used Neural Networks techniques for predicting emergency. Shah and Ghazali were proposed Improved Artificial Bee Colony (IABC) algorithm to improving the training process of Multilayer Perceptron (MLP) in order to overcome local minimal and slow convergence of ordinary backpropagation (BP) [17]. And the IACB-MLP has showed that it is outperforms than conventional BP for forecasting earthquake magnitude with time series data in California. Also, to predict magnitudes of earthquake and the impending event following the occurrence of pre-seismic signals, Moustra et al. evaluated the performances of Artificial Neural Networks (ANNs) with various types of input data from the region of Greece [18]. In their study, a feed-forward MLP type Neural Network were implemented using the BP learning algorithm for training. A feed-forward BP algorithm has been also applied into development of a time-dependent surrogate model of storm surge [19]. As experimental results, storm surge was predicted by the 92 trained networks for approaching hurricane climatological and track parameters in a few seconds. Other approach based on ANN has been studied to forecast probabilities of occurrence and re-occurrences of earthquake in the region of Chile by Reyes et

al., and they used input values (e.g., the b-value, the Bath's law, and the Omori-Utsu's law) which are strongly correlated with seismicity [20]. The occurrences have been judged by threshold values which are adjusted for obtaining as few as false positives as possible. In addition, a combination an ANN and Genetic Algorithm (GA) has been proposed to predict 1-day-ahead Monsoon flood by Sahay and Scrivastava [21]. Four wavelet transform-genetic algorithm-neural network models (WAGANN) have been developed and evaluated for forecasting flows in two Indian Reviers, the Kosi and the Gandak. In their experiments, WAGANN models predicted relatively reasonable estimates for the extreme flows and showed little bias for underprediction or overprediction.

A variety of clustering methods has been also applied into the prediction task for emergency management. An approach for the prediction of the seasonal tropical cyclone activity over the western North Pacific has been developed to provide useful probabilistic information on the seasonal characteristics of the tropical cyclone tracks and vulnerable areas [22]. In a developed model, the fuzzy c-means clustering has been used to forecast tropical cyclone tracks and density over the entire basin. From an experiments, seven patterns were founded to draw a map of the seasonal track density of tropical cyclone. Moreover, the k-means clustering technique has been combined with the statistical regression techniques for the inducement the weather phenomenon in forecasting the cloudburst [23]. The approach clusters atmospheric pressure according to areas of strong relative humidity for discovering weather patterns. To predict wildfire risk using weather data, Context-Based Fire Risk (CBFR) model has been developed based on clustering and ensemble learning techniques by considering the inherent challenge arising due to the temporal dynamicity of weather data [24]. These two machine learning techniques are used to anomaly detection. A particle swarm optimization algorithm-based clustering method with abnormally high-dimensional data has been also proposed to forecast earthquake [25]. A model analyzes relationships between earthquake precursor data and earthquake magnitude, and an average distance between clusters is set as the evaluation function of the particle swarm optimization clustering algorithm. Experimental results indicate that this model can effectively and validly predict the earthquake magnitude in accordance with the earthquake precursor data than k-means algorithm model. Additionally, an ant-colony clustering algorithm has been introduced in earthquake prediction by Shao et al [26]. Measure parameters include spatial entropy, mean-fit and un-similar for clustering analysis. As their experiments, it showed that their algorithm could achieve better results than the traditional k-means algorithm to forecast of earthquake like the swarm optimization algorithm-based clustering method.

Decision trees which have often fast and accurate performance in machine learning, have been combined with other techniques, to predict emergencies. To predict disaster before it's occurrence, there are many studies which combine the decision tree with various machine learning techniques such as Regression [27], hidden Markov Model [28], association rule learning [29,30] and fuzzy logic and particle swarm optimization [31]. In particular, decision tree techniques have been also applied into prediction of surroundings (e.g., flood susceptible areas [32] and landslide susceptible areas [33]) for emergency situation.

2.3 Warning Systems

Large magnitude emergency such as earthquake and flood to kill and injure tens to hundreds of thousands of people, inflicting lasting societal and economic disasters. Early warning could provides seconds to minutes of warning, allowing people to move to safe zones and prepare activities like automated slowdown and shutdown of transit and other machinery [34].

For early detection and warning of emergency in environments with wireless sensor networks (WSNs), Bahrepour et al. have tried to consolidate a general decision tree with the reputation-based voting method [35,36]. In their works, early event warning of emergency are fulfilled with distributed event detection. As experimental results with wild and residential fire datasets, it was showed that their approach not only achieves a high detection rate but also has a low computational overhead and time complexity. In addition, a Random Forest (RF) based decision tree was applied into analysis of the potential factors affecting the satellite signal to announce the flood by Revilla-Romero et al [37]. They investigated various satellite data for 322 rivers in Africa, Asia, Europe, North America and South America. Their experiments shown that mean discharge, climatic region, land cover and upstream catchment area are the dominant variables which determine good or poor performance of the measurement sites.

ANN techniques have been applied into warning emergency with more various perspectives. Kong et al. used a smartphone-based seismic network which consists of smart device contains accelerometers [34]. The ANN was used to separate data collected from personal smartphone sensors into activities of the earthquake and human to warn earthquake. Additionally, to set an early warning threshold level of dam, a continuous monitoring of long-term static deformation based on three ANN approaches (i.e., the static neural network, the dynamic neural network and the auto-associate network) was proposed [38].

Krzhizhanovskaya et al. developed a flood early warning system to monitor sensor networks installed in flood defenses (e.g., dikes, dams and embankments), detect abnormalities in sensor signals, calculate dike failure probability, and simulate possible scenarios of dike breaching and flood propagation [39,40]. In the warning system, k-means clustering and Neural Clouds (NC) based classification has role to detect abnormality of sensor parameters in critical pre-failure conditions. Social media data that contains social concerns of people was used for early warning system by Avvenuti et al [41]. Their system applies classification techniques provided by Weka to distinguish Twitter messages into "useful" and "not useful", and several machine learning techniques are utilized to temporal and spatial analysis of messages. There is also an other study considers social media data. Fersini et al. implemented a decision support system using machine learning and natural language processing to effectively detect and warn the earthquake [42]. On a real Twitter dataset, their system has shown outperformed results to identify messages related to the earthquakes and critical tremors.

2.4 Event Detection & Tracking

The Support vector machine (SVM) is a supervised learning model that defines a kernel function able to transform the data to a high dimensional feature space

when the data can be separated by linear models. SVM is designed for binary classification in nature, but can also solve the multi-class classification problems through one-against-one or one-against-all strategy. SVM was found be effective in emergency rescue evacuation support system for detecting and tracking a sudden incident [43,44]. Pohl et al. taked advantage of the Self-Organizing Map (SOM) and the Agglomerative Clustering (AC) for sub-event detection that operate on Flickr and YouTube data [45]. As their mentions, multimedia data may be of particular importance to detect and track emergency event. Therefore, Vector Space Model (VSM) was also utilized to represent and annotate the media data before applying clustering techniques. Their experiments showed that social multimedia in the context of emergency is worth using for detecting sub-events. Similarly as, a method of cross-media analytic was introduced to detect and track emergency events by using the clustering, the sentiment analysis, and the keyword extraction [46]. Moreover, the semantic expansion and sentiment analysis were adopted to quantify public sentiment time series.

Song et al. developed a model of human behavior that takes into account several factors have been founded through empirical analysis between human mobility and emergency to detect and monitor human emergency behavior and their mobility during large-scale emergency [47]. For the model, they used Hidden Markov Model (HMM) to model dependency between human behaviors in emergency. For the Great East Japan Earthquake and the Fukushima nuclear accident, the efficiency of the behavior model was evaluated. HMM for speech recognition technique was applied into detecting earthquakes and tracking volcano activities [48]. To fit the model parameters in to earthquake detection, Beyreuther et al. introduced state clustering into their model to refine the intrinsically assumed time dependency. As experiments in during around four months, their earthquake detector of single station HMM showed that it can achieve similar detection rates as a common trigger in combination with coincidence sums over two stations. Akin to this, an approach, which used audio data, proposed to identify anthropogenic disasters by Ye et al [49]. In their approach, acoustic events are detected and learned using the dictionary learning and the spherical k-means clustering. And detected events then are classified into specific sounds (e.g., screaming, shouting, gun shout and explosion) by a clustering technique based on the hierarchical regularized logistic regression model. Experimental results with an audio dataset showed the effectiveness of the proposed hazard sound recognition method. Singh et al. investigated Twitter posts in a flood and proposed an algorithm to identify victims asking for help. To categorize the posts into high or low priority tweets, the SVM, the gradient boosting and the RF are applied. Furthermore, inferring users location using the Markov model uses historical locations of users. In their experiments, the proposed algorithm worked with its classification accuracy of 81%, and location prediction accuracy of 87% [50]. In addition, Caragea et al. developed en enhanced messaging for the emergency response sector, as a reusable information technology infrastructure, to detect and track emergency [51]. They focused on correct classifying messages during disasters by using the SVM classifier. Furthermore, the bag-of-words (BoW) approach, the feature abstraction, the feature selection and the Latent Dirichlet Allocation (LDA) were applied into feature representation as inputs for learning the classify. Besides, various techniques of machine learning have been used for analysis of social media data to detect and track emergencies [52–54].

2.5 Situational Awareness

The recent advances of mobile devices that are capable of wireless communications and have sufficient computing power have caught attention of researchers and practitioners. The devices can serve as automated sensors (typically equipped with GPS, motion sensors, etc.), and are capable of relatively high quality imaging and video recording. Therefore, they can be used into enhance situational awareness by gleaning various information to produce accurate results [55]. Furthermore, the advance of social networking services (e.g., Twitter, Instagram, Flicker, Facebook, and others) allows people post their needs and gather the timely-relevant information. Tweets as one of these messages were also investigated to detect possible seismic events, to compare and contrast the people behavior during emergency and to extract useful information using several extraction techniques [8]. Like this, social media is also a potential source for situational awareness in emergency management. Recent disasters, such as the Hurricane Sandy of 2012, the Typhoons Haiyan or Hagpuit in 2013-2014, or the Nepal earthquake in 2015 have shown that information provided by eye-witnesses through social networking services can greatly improve situational awareness.

Alam et al., for situational awareness during cyclone emergency, proposed a social media image processing pipeline, which includes a noise filter and a damage assessment classifier [56–58]. The Convolutional Neural Network (CNN) applied into filtering out irrelevant image content, and the perceptual hashing technique was employed for image de-duplication. Additionally, CNN technique was also used for situational awareness during heavy rainfall by Li et al [59]. They focused on using social remote sensing data for emergency response. The classification results obtained for the central parts of Wuhan and Shenzhen demonstrated the effectiveness of the CNN method considered for monitoring the heavy rainfall event that happened in both cities.

Shen et al. proposed a method to retrieve events based on event-specific hashtags preliminarily collected for situational awareness of emergencies [60]. In their experiments, the SVN showed best performance for extracting and classifing hashtags from data in Twitter. Then, the hashtags were used to collect relevant messages from not only Twitter but also other social media platforms. The SVM was also introduced to extract features and classify texts [61]. Raginia et al. proposed a hybrid method for segregating and classifying the texts obtained from people who are at risk in the affected region for situational awareness of emergency. The results showed that the text classification algorithm can help the emergency responders to locate the people at risk in real-time.

For understanding situations in disaster response, Li et al. proposed a domain adaptation approach, which learns classifiers from unlabelled target data, in addition to labelled data [62]. A Naive Bayes (NB) classifier and an iterative self-training strategy were adopted for tweet classification. Their experiment results showed that the domain adaptation classifiers are better as comparing with the supervised learning using only labelled data. Ramchurn et al. proposed an emergency management system called HAC-ER for situational awareness from large streams of reports posted by members of the public and trusted organizations [63]. They combined the independent bayesian classifier combination with the gaussian process to remove errors and to predict locations of events in affected areas by emergencies. Additionally, Imran et al. presented human-annotated Twit-

ter corpora collected during 19 different crises and compared supervised learning techniques such as the SVM, the NB and the RF in terms of the utility of annotations [64]. Given the complexity of the multiclass classification of short messages, it was indicated that all three classifiers have decedent results. To discover important topics from Twitter and provide useful information of situation awareness during emergency, Yin et al developed an online incremental clustering algorithm that automatically groups similar tweets into topic clusters [65]. They also adapted optimization techniques (i.e., burst detection, text classification, online clustering, and geotagging) to deal with real-time, high-volume text streams. It includes an early indicator identification of unexpected incidents, an impact exploration of events and an incidents evolution monitoring.

In general, to classify social media data may be tedious and time consuming task, since the collected data are not in the form of a labelled data. Therefore, Pandey and Natarajan utilized the semi-supervised machine learning approach [66, 67] to avoid the classification process and concurrently obtain useful information in situational awareness. In addition, they also introduced an interactive map to grasp the vulnerable areas during a emergency. Whereas, a reinforcement learning technique was introduced to map dynamic situations in emergency. Sadhu et al. proposed a Multi-Agent Reinforcement Learning (MARL) framework implemented as a mobile application and a backend server [68]. Via both simulations and real experiments, an evaluation of the framework in terms of effectiveness in tracking random dynamicity of the environment was performed.

2.6 Emergency Evaluation

Trekin at el. applied CNN into developing a method of change detection on remote sensing imagery to improve time effciency of assessment of damaged buildings in disaster affected area [69]. Also, a deep learning-based framework for rapid regional tsunami damage recognition using post-event on synthetic radar imagery was proposed [70]. They applied the SqueezeNet network (as a CNN type) architecture into a selection algorithm, and a recognition algorithm with a modified wide residual network was developed to classify the damaged regions. Via experiments on Tohoku earthquake in 2011 and tsunami area, it was showed that the proposed framework is fast in model training and prediction calculations. The potential of CNN features was also explored for an online classification of satellite image to detect structural damages by Vetrivel et al [71]. A feature extraction and classification process are carried out at an object level, where the objects are obtained by over-segmentation of satellite images. The proposed framework outperformed a batch classifier with lesser time and memory requirements. As other usage case of CNN for natural emergency evaluation, analysis of images posted on social media platforms using the CNN was proposed [72]. Experimental results indicated that the domain-specific fine-tuning of deep CNN outperforms Bag-of-Visual-Words (BoVW). In addition, high classification accuracy under both event-specific and cross-event test settings demonstrated that their approach can effectively adapt deep-CNN features to identify the severity of destruction from social media images taken after a disaster strikes. Additionally, Attari et al. also introduced a Nazr-CNN, a deep learning pipeline for an object detection and fine-grained classification in aerial images for assessing and monitoring damage [73]. In here, a

hidden layer of a CNN was used to encode the popular BoVW of the segments generated from the first component in order to help discriminate between different levels of damage. Moreover, BP neural network as a kind of multilayer feed-forward network was used in evaluation of city emergency management system for disaster event by Jiang and Li [74].

Cervone et al. proposed a methodology that leverages data harvested from social media for collecting remote-sensing imagery during disasters [75,76]. The images are then fused with multiple sources for the damage assessment of transportation infrastructure. In this method, DT was used to classify entire scenes acquired. They also evaluated the proposed methodgology with considering Colorado floods in 2013 [76]. Zhang el as. proposed a machine learning framework to assess post-earthquake structural safety [77]. In this framework, Classification and Regression Tree (CART) and RF were implemented to map damage patterns to classified structural safety states. For assessment of sensitive area for landslide at the Pauri Garhwal in India, RF and CART were also compared with Logistic Model Trees (LMT) and Best First Decision Trees (BFDT) [78]. The results showed that the RF model has the highest predictive capability followed by the LMT, BFDT and CART models, respectively. It was showed that although all four methods have shown good results, the performance of the RF method was the best for landslide spatial prediction.

An assessment model based on RF was adopted to evaluate regional flood hazard [79]. The risk assessment method was implemented in Dongjiang River Basin, China. In addition, the SVM technique was used for risk assessment as a comparison, as well as an analysis of index importance degree. The spatial distributions of the RF and SVM assessment maps showed a similar correlation coefficient, was indicated that the classification capacity of the two methods is similar in the majority of cases. Joshi et al. introduced a methodology for detection of damage post disasters by examining the textural features from high resolution aerial imagery [80]. The proposed technique considered DT, NB, SVM, RF, Voting Classifier and Adaptive Booster, and were compared to identify damaged regions from aerial images using only pre-event images as the input. As a result, the RF-based classifier comparatively had higher accuracy than other classifiers. Yoon and Jeong applied Cubist and RF techniques into assessment what vulnerability indicators are statistically associated with disaster damage in Korea, and found twelve indicators to evaluate vulnerability of 230 local communities to disasters [81].

Zanini et al. proposed a procedure based on Fuzzy Logic (FL) for the evaluation of interactions between existing buildings and urban roadway networks after a seismic event [82]. The methodology was applied to the Municipality of Conegliano in Italy in the potential seismic damage scenario. Their experiments showed it is able to evaluate the network link functionality reductions caused by building damages, through the estimation of the residual road width, without the necessity of carrying out expensive and detailed surveys on the analyzed area. Izadi at el. proposed a neuro-fuzzy approach based on the GA and the SVM for the semi-automatic detection and assessment of damaged roads in urban areas using preevent vector map and both pre and post-earthquake QuickBird images [83]. Experimental results showed the efficiency and accuracy of the Neuro-Fuzzy systems for road damage assessment. Resch et al. introduced an approach based on analyze social media posts to assess footprints and the damage caused by natural disasters through combining LDA for semantic information extraction with spatial

and temporal analysis for hot spot detection [84]. Furthermore, they provided a damage map that indicates where significant losses have occurred. Their experiments showed that earthquake footprints can be reliably and accurately identified in our use case. Nadi and Edrisi introduced a Markov decision process as a multi-agent assessment and a response system with reinforcement learning designed to ensure the integration of emergency response and relief assessment operations [85]. Experiments indicated that the use of the proposed approach in assessing network conditions and true demand during search and rescue operations can decrease death tolls.

2.7 Crowdsourcing

Volunteers provide information and resources to the affected people and this process has been facilitated by social media in recent years [13]. In this regard, for some years now, both researchers and practitioners in the areas of emergency management have been exploring the role of crowdsourcing in collecting, processing, and sharing information [86]. Although there are various roles of crowdsourcing in tasks of emergency management, in this section, we focus on it's usages in tasks for post-emergency, especially related relief activities. The others will be investigated and discussed in Sect .5. First we start with reviewing studies focused on crowdsourcing for post-emergency without considering machine learning techniques.

Landwehr and Carley have reviewed how social media is used in disaster by individuals, first responders, and disaster researchers. They have also introduced a variety of software tools that can be used by analysts to work with social media and have discussed of several different directions in which some of the research on social media usage in disaster is currently heading [87]. As aiming at efficiently harnessing crowdsourcing in remote assistance in real-time, Yang et al. designed and developed a crowdsourcing disaster support platform [88]. They considered three unique features as follows: selecting and notifying individual requests, providing collaborative working functionalities, improving answer credibility through "crowd voting." In addtion, Dubey et al attempted to develop a theoretical framework which can assist relief activities using valuable information derived using comprehensive crowdsourcing framework in environments with Internet of things [89]. They have conducted extensive review of articles published in reputable journals, magazines and blogs by eminent practitioners and policy makers. Murali et al. proposed a multi-platform model to deal with disasters and support relief activities while handling the needs of victims, volunteers and government agencies [90]. Further, they used various techniques, ranging from Natural Language Processing (NLP) to crowdsourcing, for ensuring robustness and scalablility of solution.

Various techniques of machine learning have been also combined with crowdsourcing to support relief activities in post-emergency. Most of the current systems allow volunteers to directly provide input to them [91]. Hence, many social media posts in the aftermath of disasters might contain useful information. There are several tries to extract the information from social media through a variety of machine learning techniques such as LDA and DT [91], RF and NB classifier [92]. In order to detect potential incidents implicated by victims negative emotions in the post-disaster situation, Bai et al [93]. introduced a structured framework including three phases. The NB, RF, SVM, and KNN techniques were compared

each other in terms of to classify emergency-related messages, and the RF and SVM were outperformed than the others. Harris et al. proposed an approach of the post-phase situational awareness of an earthquake hit area to the rescue task [94]. To ensure the credibility of the crowdsourced data, their system considers the K-means clustering technique and maps coordinates of the calamity area through a short messaging service. In addition, experimentation was carried out to evaluate the time taken to notify via SMS. Imran et al. also applied a clustering technique to the classification of crisis-related messages in microblog streams [95]. Social media messages are clustered together with textual similarity, and human curators annotated the larger clusters first to train the classifer. Liang at el. introduced a semi-supervised learning based cognitive framework to support emergency management through mapping crowdsourced data. The framework first divides the satellite or aerial image into patches leveraging a graph-based clustering approach [96]. The KNN classifier is then used to provide labels for a few patches. With over 50 participants working on three different tasks, their experiments showed that the crowdsourced variant performs well producing noise-tolerate flood maps.

3 Analysis of Emergency Data

Information exchange during and after the disaster periods can greatly reduce the losses caused by the disaster. This is because it allows people to make better use of the available resources and provides a channel through which reports on casualties and losses in each affected area can be delivered expeditiously [97]. Furthermore, the success of a disaster relief and response process is largely dependent on timely and accurate information regarding the status of the disaster, the surrounding environment, and the affected people [13]. Therefore, understanding, analytic, and utilization of data collected in emergency are vital. In this section, we describe data characteristics generated during disaster with considering the 5 Vs of big data and various application cases of Big Data Analysis (BDA) are looked over.

3.1 Big Data in Emergency Management

The role of data in emergency management has been evolving. Nowadays, scientists are facing one of the biggest challenges of managing large volumes of data generated at times of disasters. As a huge amount of emergency-related data is getting generated, traditional data storage and processing systems are facing challenges in fulfilling performance, scalability and availability [98]. Therefore, analytic methods to manage and process data in emergency management are particularly challenged due to the combination of it's unique characteristics as follows:

- **Rapid increase of data** by many number of producers and consumers
- **The timely sensitivity** of detect and response
- **Combination of static and dynamic data** (e.g., maps and crowd emotion) [96,99]
- **Heterogeneous formats**, ranging from raw data (e.g., sensors) to structured data (e.g., metadata) and unstructured data (e.g., multimedia) [86]
- **Various levels of trustworthiness** of the data sources [100,101]

- **Possibility of extracting valuable information** like crowdsoucing, generated by people who are actually at the emergency scene for near real-time [97]

These characteristics similar to "5 Vs (i.e., volume, variety, velocity, veracity, and value)" of big data. Besides, within the present-day emergency management systems, the immediate and accurate decision-making more and more relies on the capability of data analysis and processing especially in the face of big data [1]. Therefore, there is an urgent need to enhance the BDA technologies of emergency management, such as, to develop scalable and real-time algorithms for time-sensitive decisions, to integrate structured, unstructured, and semi-structured data, to deal with the imprecise and uncertain information, to extract dynamic patterns and outline the evolution of these patterns, to work in distributed environment, and to present the multi-scale, multilevel and multi-dimensional patterns through various visualization approaches [102].

BDA was often defined as holistic process to manage, process and analyze the 5 Vs in order to create actionable insights for sustained competitive advantages [103]. In this regard, Mehrotra et al. suggested that BDA can aid to create the next generation of emergency management technologies as it has the potential to mitigate the effects of disasters by enabling access to critical information in real-time [104]. Also, they emphasized that "accurate and timely analysis and assessment of the situation can empower decision makers during a crisis to make more informed decisions, take appropriate actions, and better manage the response process and associated risks." Thus, it is essential to reconsider how data on disasters should be properly and efficiently produced, organized, stored, and analyzed [105]. Here, we briefly discuss the BDA for emergency management as processes separated such as data collection, information extraction, data filtering and data integration.

3.2 Data Collection

Data collection is the process of inserting to a system data, which is coming from multiple heterogeneous sources. Scalability is an important issue in data collection, since the flow of information may be very high during the time of a critical event [106]. Traditionally, it was done in the form of paper reports or questionnaires. The development of IT allowed for use of word processors, spreadsheets, and forms to enter data directly into the databases [7]. Moreover, sensing technologies are also currently undergoing rapid advances, leading to might use of them significantly increases performance of situational awareness [107]. As another increasingly important source, social media is a new way of communication in the course of disasters. A major difference between social media and traditional sources is the possibility of receiving feedback from the affected people [13]. Additionally, for some years now, both researchers and practitioners in the areas of disaster and emergency management have been exploring the role of crowdsourcing in collecting, processing, and sharing information across organizations and affected populations [86].

3.3 Information Extraction and Filtering

The goal of information extraction is to to automatically extract structured information, i.e., categorized and contextually and semantically well-defined data from a certain domain, from unstructured machine-readable documents [106]. Whereas, according to Wikipedia, a data filtering is that removes redundant or unwanted information from an data stream using (semi) automated or computerized methods prior to provide it. Its main goal is the management of the data overload and increment of the semantic information. If all the disaster information were presented to the users, it would cause an overwhelming workload. Therefore, the disaster information should be filtered based on the specific purposes of the users [97]. Especially, in emergency management, the utilization of social media data be extended for various tasks such as warning [41], detection & tracking [52,45], situational awareness [91,58], assessment [76,84], etc. Therefore, it is more emphasized that the data extraction and filtering became the core of emergency management together with effects of social media data. However, the social media data contain texts, images, videos, tags, and so on. Therefore, information is extracted from heterogeneous sources such as social media and monitoring devices. Typically, the disaster information from different sources varies greatly in structure or format. To support further analysis and processing, a specification of common format should be required for disaster information integration. Then, the integrated disaster information in this format can be organized and stored for further processing [97].

3.4 Data Integration

Emergency services sometimes should deal with the massive amount of data arriving through multiple channels such as existing records, sensors, satellite networks or social media [108]. Therefore, one of the biggest challenge in emergency management is to develop a data integration protocol. Data integration is the process of combining data residing at different sources and providing the user with a unified view of these data [109]. This process includes the following tasks: (1) to convert contents of different formats into a standard format; (2) to verify the credibility of various crowdsourcing data sources and attempt to leverage it to produce useful information for disaster decision-making; (3) to map images or texts with their corresponding geolocations to better capture the current situation; and (4) to process and analyze the data from different sources [97].

3.5 Applications for Data Analysis in Emergency

In this section, we introduce several applications and platforms comprehensively deal with data analysis for emergency managements. **Ushahidi** as the first large-scale crowdsourcing system developed to report Kenyan post-election violence in 2008 and since then has been applied into many major disasters such as Hurricane Sandy and Haiti Earthquake [110]. Ushahidi is an open source and free systems which can either be deployed on external servers or on it's hosting system CrowdMap. This system collects emergency-related data from several sources,

web, Twitter, RSS feeds, emails, SMS, and so on. Collected data is then visualized on the map. Last, Ushahidi allows users to filter information based on several types (e.g. supplies or shelter).

Artificial Intelligence for Disaster Response (AIDR) is a free software platform which can be either run as a web application or created [111]. In AIDR, tweets are collected according to pre-selected set of keywords. Prior labeled tweets will be used as the training set of a classifier which labels collected tweets based on the keywords. In the training process, n-grams of tweets are used as features and therefore the classifier is retrained for every new category and disaster.

TweetTracker consists of tracking, analyzing, and understanding tweets related to a specific topic [112]. In the process of tracking the status of and event, data includes keywords, location, and users can be collected using a set of criteria from Twitter, Facebook, YouTube, VK, and Instagram. Fluctuations in the total number of post or frequency of posts with specific words can be analyzed for different time periods. Moreover, keywords, hashtags, links, images, and videos with their frequencies are available to understand tweets by the user. For instance, to better understand the geographic distribution of posts, geo-tagged tweets are presented on a map.

DisasterMapper is a CyberGIS framework to ingest and archive massive amounts of social media data [113]. In this framework, to manage massive social media data, Apache Hive is used as the scalable storage solution. Furthermore, Hadoop platform is used as a scalable distributed computing environment to process social media data. Mahout is leveraged to support big data analytics. It can can automatically synthesize multi-sourced data, such as social media and socioeconomic data, to track disaster events, to produce maps, and to perform spatial and statistical analysis for emergency management.

IDDSS-Sensor is GIS-based software implemented to provide the functions of standard-based access, as well as on-the-fly harmonization, integration and usage of multi-agency sensor information [114]. The software has three layers, namely, the storage, service, and presentation layer. In storage layer, PostgreSQL is used as an open-source object relational database of integration data models. For service layer, 52ºNorth and GeoServer was employed as SOS implementation and as spatial data service for serving static spatial data. As GUI of the system, third layer was developed by JavaScript, and the CESIUM, ExtJS3 and IDDSS were uased for displaying the sensor data and visual indicators.

4 Challenges and Opportunities of machine learning in Emergency

Here, we summarize and look over approaches aforementioned to draw challenges and opportunities of machine learning techniques in emergency management. Table 1 lists approaches and applications in emergency management in terms of target types of emergency, belonged tasks, used technique of machine learning, and data analysis.

This section focuses on issues for machine learning techniques for overall tasks in emergency management. Crowdsourcing with machine learning will be discussed in section 5.

Information and Knowledge Learning. For this category, the following challenges are raised as further research issues: (1) Support sustainable annotation and

Table 1 Approaches with Machine Learning for emergency management

Title	Emergency	Task	Machine learning technique	Data	Source	Data analysis
Prediction of Earthquake Magnitude by an Improved ABC-MLP [17]	Earthquake	Prediction	Artificial Bee Colony, Multilayer Perceptron, Backpropagation	Earthquake parameter magnitude	California earthquake data (2011)[a]	– Labeling
Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals [18]	Earthquake	Prediction	Artificial Neural Network	time series magnitude data, Seismic electric signals	Seismological Institute National Observatory of Athens (SINOA), VAN team in Greece Athens earthquake data (1981)	– Labeling
A time-dependent surrogate model for storm surge prediction based on an artificial neural network using high-fidelity synthetic hurricane modeling [19]	Hurricane	Prediction	Artificial Neural Network, Backpropagation	Historical hurricane parameters	Atlantic hurricane database, HURDAT2[b]	– Labeling
Neural Networks to Predict Earthquakes in Chile [20]	Earthquake	Prediction	Artificial Neural Network	Earthquake of magnitude equal or larger to 3.0	Chile's National Seismological Service	– Labeling
Predicting Monsoon Floods in Rivers Embedding Wavelet Transform, Genetic Algorithm and Neural Network [21]	Flood	Prediction	Artificial Neural Network, Genetic Algorithm	Flow data	Monsoon flow (2001-07)	– Labeling
Track-pattern-based model for seasonal prediction of tropical cyclone activity in the western North Pacific [22]	Tropical cyclone	Prediction	Fuzzy C-Means Clustering	Tropical cyclone locations with maximum sustained wind speed greater than 17 ms21	Regional Specialized Meteorological Centers Tokyo-Typhoon Center	– Labeling
The Application of Ant-Colony Clustering Algorithm to Earthquake Prediction [26]	Earthquake	Prediction	Ant Colony Clustering Algorithm	Earthquake magnitude		– Labeling – Anomaly detection
Dynamic and robust wildfire risk prediction system: an unsupervised approach [24]	Wildfire	Prediction	Clustering techniques	Weather data	Fire & Rescue NSW from the Bureau of Meteorology (BoM)	– Labeling – Anomaly detection
The data mining technology of particle swarm optimization algorithm in earthquake prediction [25]	Earthquake	Prediction	Particle Swarm Optimization Clustering Algorithm	Earthquake magnitude, earthquake precursor data		– Labeling – Anomaly detection
Spatial prediction of flood susceptible areas using rule based decision tree and a novel ensemble bivariate and multivariate statistical models in GIS [32]	Flood	Prediction	Rule-based Decision Tree, Frequency Ratio and Logistic Regression Statistical Methods	Rainfall data, spatial database	Landslide-hazard map [115]	– Labeling – Anomaly detection – Mapping and flood susceptibility
GIS-based landslide susceptibility modelling: a comparative assessment of kernel logistic regression, Naïve-Bayes tree, and alternating decision tree models [33]	Landslide	Prediction	Kernal Logistic Regression, Naive Bayes Tree, Alternating Decision Tree	Landslide inventory map, Landslide conditioning factors, Landsat 8 OLI images, precipitation data	Self-collection, NASA[c], U.S Geological Survey (USGS)[d]	– Collecting landslide inventory map – Labeling – To manually ingrate early reports, aerial phtographs, GPSs

[a] http://www.data.scec.org/

[b] http://www.nhc.noaa.gov/data/#hurdat

[c] http://reverb.echo.nasa.gov/reverb/

[d] http://landsat.usgs.gov/landsat8.php, http://www.sxmb.gov.cn/index.php

Table 1 continued

Title	Emergency	Task	Machine learning technique	Data	Source	Data analysis
Clustering technique to interpret Numerical Weather Prediction output products for forecast of Cloudburst [23]	Cloudburst	Prediction	K-Means Clustering	Flow pattern data (temperature according to atmospheric pressure level)	European Center for Medium-range Weather Forecasting	– Labeling
Application of decision trees to the analysis of soil radon data for earthquake prediction [27]	Earthquake	Prediction	Decision Tree	Soil radon data	Soil radon data [116]	– Labeling
A new data mining model for hurricane intensity prediction [28]	Hurricane	Prediction	Feature Weight Learning, Genetic Algorithm, Extensible Markov Model	Historical hurricane data	Atlantic tropical cyclones (1982-2003) National Center for Atmospheric Research NCAR	– Labeling
Computational Intelligence Techniques for Predicting Earthquakes [29]	Earthquake	Prediction	Association Rules, Decision Tree	Location and magnitude of Spanish earthquakes	Spanishs Geographical Institute (SGI)	– Labeling
Novel method for hurricane trajectory prediction based on data mining [30]	Hurricane	Prediction	Association Rules	Hurricanes motion characteristics	Atlantic weather Hurricane/Tropical Data from 1900 to 2008[a]	– Labeling
Spatial Data Mining for Prediction of natural events and emergency management based on fuzzy logic using hybrid PSO [31]	Natural emergency	Prediction	Fuzzy Logic, Particle Swarm Optimization, K-Meas Clustering, Naive bayes classifier	Water level modifications data, spatial-temporal data, text data	Bangladesh Country Almanac (BCA) data set	– Labeling and centroid detection – Filtering inconsistent and unnecessary data – Integrating spatial data and text
Evaluation of the satellite-based Global Flood Detection System for measuring river discharge: influence of local factors [37]	Flood	Warning Detection & Tracking, Evaluation	Random Forest	– Satellite-derived data (surface water extent and floodplains) – River width – Presence of floodplains flooded forest and wetlands – Flood extent – Land cover – Leaf area index – Climatic areas – Presence of river ice – Dam location	– Global Flood Detection System, Global Runoff Data Centre, and South African Water Affairs[b] (1998 to 2010), – SRTM Water Body Database and the HydroSHEDS [117] – Global Lakes and Wetlands Database level 3 [118] – Global Flood Hazard Map [119] – Global Land Cover 2009 [120] – SPOT-VGT[c] – Koppen-Geiger climate map of the world [121] – Circum-Arctic map [122] – Global Reservoir and Dam [123]	– Labeling – Integrating data related to flood
The evaluation of city emergency management system results based on BP neural network [74]	Natural emergency	Evaluation	Backpropagation Neural Network	Results evaluation of City emergency management system		– Survey evaluation data – Labeling

[a] http://weather.unisys.com/hurricane/

[b] http://www.dwa.gov.za/

[c] http://wdc.dlr.de/

Table 1 continued

Title	Emergency	Task	Machine learning technique	Data	Source	Data analysis
Use of wireless sensor networks for distributed event detection in emergency management applications[36]	Fire	Warning	Decision Tree	Residential fire data (temperature, ionization, photoelectric, and CO)	NIST[a]	– Labeling
MyShake: A smartphone seismic network for earthquake early warning and beyond [34]	Earthquake	Warning	Artificial Neural Network	Acceleration data from android phone, GPS	Self-collection, Northern and Southern California Earthquake Data Centers[b], the Center for Engineering Strong Motion Data, and the National Research Institute for Earth Science and Disaster Prevention for Japanese	– Collecting smartphone sensor – Labeling – Filtering frequency band – Integrating locations and acceleration according to time-stamps
Monitoring of long-term static deformation data of Fei-Tsui arch dam using artificial neural network-based approaches [38]	Flood	Warning	Artificial Neural Network (the static neural network, the dynamic neural network, and the auto-associate network)	Residual deformation in Fei-Tsui arch dam (the static deformation, the water level, and temperature distribution)		– Collecting the static deformation, the water level, and temperature distribution – Labeling
Flood early warning system: design, implementation and computational modules [39] Flood early warning system: sensors and internet [40]	Flood	Warning	K-Means Clustering	– streams of sensor data – digital terrain model data	Actueel Hoogtebestand Nederland[c]	– Collecting real time sensor data – Labeling and anomaly detection – Integrating mapping and sensor data
Earthquake emergency management by social sensing [41]	Earthquake	Warning, Evaluation	Bayesian statistics, Peak-detection algorithms, Corrected conditional entropy, classification techniques	Social media data with keywords, multimedial context such as videos and photos	Using Streaming API of Twitter	– Collecting social media data – Labeling – Removing noise of different meaning and past event
Earthquake management: a decision support system based on natural language processing [42]	Earthquake	Warning	Bayesian Model Averaging, Voting mechanism, Classifiers (decision tree, support vector machine, naive bayes, logistic regression, and K-near neighbors	Social media data with keywords, K	Using Streaming API of Twitter	– Collecting social media data – Labeling

[a] http://smokealarm.nist.gov/

[b] www.NCEDC.org, www.SCEDC.org

[c] http://www.ahn.nl/

Table 1 continued

Title	Emergency	Task	Machine learning technique	Data	Source	Data analysis
Development of ERESS in panic-type disasters: Disaster recognition algorithm by buffering-SVM [43] Disaster detection by statistics and SVM for emergency rescue evacuation support system [44]	Natural emergerncy	Detection & Tracking	Support Vector Machine	Sensor data from mobile		– Collecting sensor data from mobile – Feature extraction and labeling
Social media for crisis management: clustering approaches for sub-event detection [45]	Emergency	Detection & Tracking	Self-Organizing Map and Agglomerative Clustering	Social multimedia data (images and videos) - 2011 Mississippi Flood, Oslo Bombing, UK Riots, and Hurricane Irene	Using APIs of Flicker and Youtube[a]	- Collecting social media data – Labeling – Integrating spatial-temporal data, images, and videos
Classifying text messages for the Haiti earthquake [51]	Earthquake	Detection & Tracking	Support vector machine, BoWs approach, feature abstraction, feature selection, and Latent Dirichlet Allocation	Text messages in phone, e-mail, Twitter, or web about the Haiti earthquake	Ushahidi-Haiti[b]	– Labeling – Feature selection to remove irrelevant and redundant feature
Prediction of human emergency behavior and their mobility following large-scale disaster [47]	Natural emergency	Detection & Tracking	Hidden Markov Model, Bayesian filtering, Maximum Entropy Inverse Reinforcement Learning	– Human mobility database (GPS records) from Aug. 2010 to Jul. 2011 – Disaster intensity data (seismic scale and damage level) – Disaster reporting data"	"- The Great East Japan Earthquake and Fukushima nuclear from Japan Government statistical reports – government declarations – Japanese Cabinet Secretariat[c] and news reports from mainstream medias in Japan and all over the world [124]	– Recording GPS of human – Labeling – Filtering to improve behavior model efficient – Mapping and human movements
Constructing a Hidden Markov Model based earthquake detector: application to induced seismicity [48]	Earthquake	Detection & Tracking	Hidden Markov Model	– Geothermal data set of one short period station (DHFO, Mark-L4) for geothermal reservoir monitoring – Speech recordings of one broadband station (KLT, Streckeisen STS-2) for Volcano monitoring	– Geothermal plant in the municipality of Unterhaching in Germany (2008-2010) – Mt. Merapi volcano 1998	– Labeling – Filtering noise on sound recording

[a] http://flickrj.sourceforge.net/, https://developers.google.com/gdata/

[b] http://haiti.ushahidi.com

[c] http://www.cas.go.jp/

Table 1 continued

Title	Emergency	Task	Machine learning technique	Data	Source	Data analysis
Leveraging cross-media analytics to detect events and mine opinions for emergency management [46]	Natural emergency	Detection & Tracking	Clustering techniques	Microblog data of an emergency event		– Crawling microblog data such as online news and forums by Python – Extracting relevance keywords – Filtering document by negative sentiment analysis – Integrating cross-media data
Audio Data Mining for Anthropogenic Disaster Identification: An Automatic Taxonomy Approach [49]	Man-made Emergency	Detection & Tracking	K-Means dictionary learning, Clustering technique based on hierarchical regularized logistic regression model	Audio clips of emergency event from four channel	– BBC Sound Effects Library – Urban-Sound8K datasets [125] – Sound Classification [126] – Sound effects from internet sources[a]	– Labeling
Event classification and location prediction from tweets during disasters [50]	Emergerncy	Detection & Tracking	Markov Chain technique, Support vector machine, gradient boosting, and random forest	Social media data with Geo-tagging according to keywords	Using Streaming API of Twitter and Tweepy python library (Nov.-Dec. 2015 and Jul.Aug. 2016)	– Collecting social media data – Labeling – Filtering noise and redundancy
Earthquake shakes Twitter users: real-time event detection by social sensors [52] Tweet analysis for real-time event detection and earthquake reporting system development [53]	Earthquake	Prediction, Warning, Detection & Tracking	Support vector machine, Kalman filtering and particle filtering	Tweets, GPS and registered location of user,	Using Streaming API of Twitter	– Collecting social media data – Labeling – Mapping emergency event and tweets
Image4Act: Online Social Media Image Processing for Disaster Response [56]	Natural emergency	Situational awareness, Crowd-sourcing	Convolutional Neural Networks, Perceptual Hashing	Images on social media data with Geo-tagging according to keywords	Using Streaming API of Twitter	– Collecting social media data – Labeling – Filtering irrelevant and duplicate images
Automatic Image Filtering on Social Networks Using Deep Learning and Perceptual Hashing During Crises [57]	Natural emergency	Evaluation	Convolutional Neural Networks, Perceptual Hashing	Images from Twitter	Typhoon Ruby (2014), Nepal Earthquake (2015), Ecuador Earthquake (2016), and Hurricane Matthew (2016) by AIDR platform [127]	– Collecting social media data – Labeling – Filtering de-duplication and relevancy
Processing Social Media Images by Combining Human and Machine Computing during Crises [58]	Natural emergency	Evaluation	Convolutional Neural Networks, Perceptual Hashing, transfer learning	Images from Twitter, Flickr, or Instagram and so on.	Typhoon Ruby (2014), Nepal Earthquake (2015), Ecuador Earthquake (2016), and Hurricane Matthew (2016) using Streaming API of Twitter	– Collecting social media data – Labeling – Filtering de-duplication and relevancy – Integrating cross-social media data
Social Media: New Perspectives to Improve Remote Sensing for Emergency Response [59]	Natural emergency	Situational awareness	Convolutional Neural Network	Social media data with specific keywords (in Chinese)	Using Sina Weibo open platform API Weibo data from Jun. 2016 to Jul. 2016 (Wuhan) and May 2014 to May 2014 (Shenzhen)	– Collecting social media data – Labeling

[a] http://sound.natix.org/

Table 1 continued

Title	Emergency	Task	Machine learning technique	Data	Source	Data analysis
Disaster response aided by tweet classification with a domain adaptation approach [62]	Natural emergency	Situational awareness	Naive Bayes classifier, iterative self-training strategy	Social media data with hashtags	– disaster dataset (CrisisLexT6) of tweets [128]	– Labeling
Using social media to enhance emergency situation awareness [65]	Natural emergerncy	Situational awareness	Incremental clustering algorithm	Social media data from Twitter	Using Streaming API of Twitter (2010)	– Collecting social media data – Labeling – Filtering stop words
A machine learning framework for assessing post-earthquake structural safety [77]	Earthquake	Evaluation	Regression Tree, Random Forest	– Response patterns or distribution of engineering demand parameters	– Nonlinear response history analysis	– Labeling
Information retrieval of a disaster event from cross-platform social media [60]	Natural emergency	Situational awareness	Support vector machine	Social media data with hashtags	– Collecting hashtags using Streaming API of Twitter – English language disaster datasets [127–129]	– Collecting social media data – Labeling – Filtering noise – Integrating cross-social media data
Mining crisis information: A strategic approach for detection of people at risk through social media analysis [61]	Natural emergerncy	Situational awareness	Support vector machine	Social media data contained hashtags and geo-tags	– Using Streaming API of Twitter – Followthehashtag [a] – test dataset [129]	– Collecting social media data – Labeling
Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages [64]	Natural emergency	Situational awareness	Naive Bayes, Support Vector Machine, Random Forest	Social media data from Twitter	19 different crises using Streaming API of Twitter (2013-2015)	– Collecting social media data – Labeling – Filtering stop words, URLs, and user-mentions
How social media can contribute during disaster events? Case study of Chennai floods 2015 [66,67]	Natural emergency	Situational awareness	Semi-Supervised Classification	Social media data contained hashtags and geo-tags with specific keywords Web Map service and load satellite data	Using Streaming API of Twitter (2015)	– Collecting social media data – Labeling – Mapping and tweet data
Argus: Smartphone-enabled human cooperation via multi-agent reinforcement learning for disaster situational awareness [68]	Natural emergency	Situational awareness	Multi-Agent Reinforcement Learning (modified version of distributed Q-learning [130])	multimedia data (audio and video)	Using a mobile application implemented by them-self	– Collecting audio and video from user smartphone – Labeling – Generating 3D map
Satellite imagery analysis for operational damage assessment in Emergency situations [69]	Natural emergency	Evaluation	Convolutional Neural Network	Two aerospace images for before and after of emergency	two areas of Ventura and Santa Rosa counties, California, USA, 2017 from Digital-globe within Open Data Program[b]	– Labeling – Mapping and emergency damage
A framework of rapid regional tsunami damage recognition from post-event TerraSAR-X imagery using deep neural networks [70]	Tsunami	Evaluation	Convolutional Neural Network	TerraSAR-X data (high-resolution synthetic aperture radar)	the Pacific coast of the Miyagi prefecture, Japan on March 12, 2011 from the Pasco Corporation	– Collecting acquiring in the StripMap mode with HH polarization – Labeling – Mapping and emergency damage

[a] http://www.followthehashtag.com/features/Twitter-historical-data-recover/

[b] https://www.digitalglobe.com/opendata

Table 1 continued

Title	Emergency	Task	Machine learning technique	Data	Source	Data analysis
Post earthquake disaster awareness to emergency task force using crowdsourced data [94]	Earthquake	Situational awareness, Crowdsourcing	K-Means Clustering	Social media data with spatial and temporal information	Using Streaming API of Twitter	– Collecting social media data – Labeling – Mapping and disaster-hit area
Damage assessment from social media imagery data during disasters [72]	Natural emergency	Evaluation	Convolutional Neural Network	– Disaster-specific image data with specific hashtags and keywords from social media – Google image data with specific queries related to damage	– AIDR [95] Typhoon Ruby/Hagupit 2014, Nepal Earthquake 2015, Ecuador Earthquake 2016, and Hurricane Matthew 2016 – Google search	– Collecting images from web and social media – Labeling"
Nazr-CNN: Fine-Grained Classification of UAV Imagery for Damage Assessment [73]	Natural emergency	Evaluation	Nazr-Convolutional Neural Network	Unmanned Aerial Vehicles (UAV) imagery	Cyclone Pam in Vanuatu in 2015 by collaboration with the Humanitarian UAV Network (UAViators)	– Collecting unmanned aerial vehicles imagery – Labeling – Mapping and emergency damage
Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study [75]	Natural emergency	Evaluation	Decision Tree	Social media data in Twitter Unmanned Aerial Vehicles imagery	September 2013 by using Streaming API of Twitter and Falcon UAV	– Collecting unmanned aerial vehicles imagery and tweets – Labeling – Mapping and emergency damage"
Application and comparison of decision tree-based machine learning methods in landside susceptibility assessment at Pauri Garhwal Area, Uttarakhand, India [78]	Landslide	Evaluation	Random Forest, Logistic Model Trees, Best First Decision Trees and Classification and Regression Trees	Landslide and Non-landslide Data	Interpretation of satellite and Google Earth images	– Collecting images – Labeling – Mapping and emergency damage
Damage identification and assessment using image processing on post-disaster satellite imagery [80]	Natural emergency	Evaluation	Decision Tree, Naive Bayes, Support Vector Machine, Random Forest, Voting Classifier, Adaptive Booster	Aerial images	Using GeoEye1 after the 2011 Christchurch earthquake and 2011 Japan earthquake and tsunami	– Collecting images – Labeling – Filtering noise – Mapping and emergency damage
A New Neuro-Fuzzy Approach for Post-earthquake Road Damage Assessment Using GA and SVM Classification from QuickBird Satellite Images [83]	Earthquake	Evaluation	Neurofuzzy Inference, Support vector machine, Genetic algorithm	QuickBrid (high-resolution satellite images)	Bam in Iran (Dec. 2003)	– Collecting images – Labeling – mapping and road damage
Combining machine-learning topic models and spatio-temporal analysis of social media data for disaster footprint and damage assessment [84]	Earthquake	Evaluation	Latent Dirichlet Allocation	Social media data footprint of earthquake (Experiment)	Napa earthquake on Aug. 2014 using Streaming API of Twitter Official earthquake footprint by USGS	– Collecting social media data for specific time and place – Labeling – Filtering stop words, Unique words, stemming, etc. – Mapping and emergency damage

Table 1 continued

Title	Emergency	Task	Machine learning technique	Data	Source	Data analysis
Emergency-relief coordination on social media: Automatically matching resource requests and offers [92]	Natural emergency	Crowd-sourcing	Random Forest, Naive Bayes Multinomial	Social media data with keywords and hashtags	Using Streaming API of Twitter	– Collecting social media data for specific time and place – Labeling
Finding requests in social media for disaster relief [91]	Natural emergency	Crowd-sourcing	Latent Dirichlet Allocation, Decision Tree	Social media data Request data (Experiment)	Using Streaming API of Twitter request data [92]	– Collecting social media data for specific time and place – Labeling – Removing duplicate tweets, re-tweets, etc – Mapping and request tweets
A Weibo-based approach to disaster informatics: incidents monitor in post-disaster situation via Weibo text negative sentiment analysis [93]	Natural emergency	Crowd-sourcing	Random Forest, Naive Bayes, Support Vector Machine, K-Nearest Neighbor	Social media data	Using Sina Weibo open platform API Yushu earthquake dataset (3442 messages), Beijing rainstorm dataset (2517 messages), and Yuyao flooding dataset (3928 messages)	– Collecting social media data – Labeling – Filtering messages of positive and neutral emotion
Ushahidi, or testimony: Web 2.0 tools for crowd-sourcing crisis information [110]	Natural emergency	Warning, Tracking, Crowd-sourcing	Various techniques	Web, Twitter, RSS feeds, emails, SMS, etc.	Using Streaming API of Twitter	– Collecting social media data – Labeling – Filtering noise – Mapping and emergency tweets
TweetTracker: An Analysis Tool for Humanitarian and Disaster Relief [112]	Natural emergency	Tracking, Situational awareness, Crowd-sourcing	Various techniques	Various social data	Twitter, Facebook, YouTube, VK, and Instagram	– Collecting social media data – Labeling – Filtering noise – Mapping and emergency information
DisasterMapper: A CyberGIS framework for emergency management using social media data [113]	Natural emergency	Crowd-sourcing	Various techniques (Mahout)	Social media data	Using crawler based on Twitter4j[a]	– Collecting social media data – Labeling – Mapping and emergency tweets
Identifying valuable information from twitter during natural disasters [131]	Natural emergency	Situational awareness, Crowd-sourcing	Naive Bayes classifiers	Social media data	Using Streaming API of Twitter	– Collecting social media data – Labeling – Removing unnecessary data such as URLs, slang, etc.
Data mining Twitter during the UK floods: Investigating the potential use of social media in emergency management [132]	Flood	Situational awareness, Crowd-sourcing	Naive Bayes classifiers	Social media data	Using Streaming API of Twitter (Nov. 2015 to Jan. 2016)	– Collecting social media data – Labeling – Filtering noise
Crowdsourcing Incident Information for Disaster Response Using Twitter [133]	Natural emergency	Crowd-sourcing	Naive Bayes classifiers, Support Vector Machine	Social media data with specific keywords	Twitter dataset of Hurricane Sandy (Oct. 2012 to Nov. 2012) [134]	– Labeling – Mapping and emergency damage
#ChennaiFloods: Leveraging Human and Machine Learning for Crisis Mapping during Disasters Using Social Media [135]	Flood	Crowd-sourcing	Smooth Support Vector Machine	Social media data with specific hashtags	With hashtage '#ChennaiFloods' using Streaming API of Twitter	– Collecting social media data – Labeling – Filtering stop words, stemming, etc – Mapping and emergency tweets

[a] http://twitter4j.org/en/index.html

classification on heterogeneous, streaming data from the multisource. To achieve this, effective and efficient learning algorithms must be studied considering the context of the emergency and the stakeholders belong to each tasks. (2) Facilitate real-time analysis and discover information across multiple streams through developing high-speed and flexible techniques. (3) Construct customized information extraction methods that can learn by the integration between domain experts and existing system.

Integration with Geographic Information System (GIS). GISs supports to integrate, store, share, and display geographically referenced information. Users (e.g., stakeholders or decision maker) can understand overall situations and discover insight through GIS for emergency. In this regard, integration between a GIS and other components is an important research. There are worth further investigation as follows: (1) automate or crowdsource the linkage construction between information/data and geo-map in real-time. For achieve this, as a prior work, automatic location extraction techniques from disaster data obtained or collected are essential for the real-time processing. (2) Intelligent alerts and location broadcasting when people enter a dangerous area. In this case, sensor-based or vision-based approaches can be integrated as the anomaly detection techniques.

Emergency Data Analysis. As aforemetioned, data in emergency are in general generated from various sources and are heterogeneous in nature. Therefore, effective and efficient methods for data analysis in emergency management should consider discovering the inter-dependencies of data and extracting useful information and knowledge. Additionally, it is also essential to handle rapid increasing amount of data, since disaster data are generated by many of producers and customers. An additional challenge is to integrate data with great diversity which may be cased from heterogeneous sources with different levels of redundancy, accuracy, and uncertainty or may be due to different characteristics of data (e.g., structured/unstructured, real-time streams/static data). In this regards, some interesting research directions include: (1) a unified method for each specific algorithms to collect, extract, filter for heterogeneous, multisource disaster data; and (2) building an analysis method capable of real-time processing.

5 Crowdsourcing in Emergency Management

In this section, combination crowdsourcing and machine learning for various tasks in emergencies are reviewed, and challenges and opportunities are raised in terms of data analysis. Last we more deeply contemplate, with several examples (tweets in Twitter) for challenges and opportunities discussed.

Volunteering is part of how community reacts to emergencies [136] and this process has been facilitated by social media in recent years. Volunteers provide information and resources to the affected people [13]. In this regard, for some years now, both researchers and practitioners, in the field of emergency management, have been exploring crowdsourcing for collecting, processing, and sharing information between stakeholders. Jeff Howe in 2006 fir sly defined the term as "the act of taking a job traditionally performed by a designated agent and outsourcing it to an generally large group of people in the form of an open call [137]". Since Howes definition, an extended range of crowdsourcing researches have been carried out from a number of fields such as computer sciences, management, in-

formation systems and so on. Additionally, as a problem-solving method [138] crowdsourcing has also caught the attention of emerging paradigms such as collective intelligence, human computation, or social computing [86]. With respect to this, Liu has analysed the distinct skills and expertise of different crowds typically involved in emergency management: (1) affected-populations, (2) social medias, and (3) digital volunteer communities [139]. In this framework, affected populations generate local, timely, and direct experiential information, and social media make available unexpected and fortuitous experience. Last, digital volunteers offer their capabilities for processing and managing emergency data. Also, Poblet et al. introduced four different crowd's roles such as sensors, social computers, reporters, and microtaskers. It considered four data types (i.e., raw data, unstructured data, semi-structured data, and structured data) and two involvement types (i.e., active or passive) [86]. There have been recently efforts toward crowdsourcing such tasks in emergency management but it is still challenging. Social media posts come at a fast pace and immense volume. Moreover, it is challenging to collect all the posts which are related to a disaster due to the restrictions by social media services. The collected data contains daily information and is only in part insightful information. Another issue is malicious content such as spam and rumors which can cause panic and stress, especially when produced in large scale using bots [13]. In this regard, one of another area to potentially improvement these challenges is for machine learning methods which have been applied to a variety of tasks in emergency management as mentioned in section 2.

5.1 Crowdsourcing with Machine Learning for emergency management

Here, combination crowdsourcing with machine learning for various tasks in emergencies are reviewed, with exception of approaches for relief task, since the approaches are already reviewed in section 2.7.

Pandey and Natarajan proposed a prototype solution to provide situation awareness during and post the time of a disaster event using semi-supervised machine learning technique based on SVM and creating interactive open street maps for crowdsourcing the user data providing threat and relief information [66]. Their model was evaluated with the data from Chennai flood 2015. Truong et al. developed a Bayesian approach to the classification of tweets during Hurricane Sandy in order to distinguish "informational" from "conversational" tweets [131]. They designed an effective set of features and used them as input to NB classifiers. The NB classifier was also introduced to reduce noise in crowdsourced data related to emergency management [132]. Their approach was assessed with the flood data of Cumbria 2015. As similarily, Imran et al. proposed automatic methods based on NB classifier for extracting information from microblog posts. They also focused on extracting valuable "information nuggets" relevant to disaster response [140]. Kurkcu et al. likewise used NB classifier with TF-IDF to identify keywords from tweets related to emergency [133]. Experimental result indicated that crowdsourced data refined could provide detailed location information of a specific incident along with its intensity, duration and. Anbalagan and Valliyammai proposed a system performs disaster tweet collection based on trending disaster hash tags [135]. They compared Naive-Bayesian and Smooth Support Vector Machine (SSVM) classifications on collected tweets to identify the severity of the

emergency. As a results, SSVM outperformed Naive-Bayesian. Also, emergency geographic map was generated for the affected area through location to interpolation cluster proximity. Balena et al. evaluated supervised classification methods (AdaBoost, NB, RF, SVM, and Neural Networks) to compare their effectiveness and potential for classifying message requests asking for/offering to help in emergency [141]. As a results, the RF and the Neural Network had better performance than the others. Nagy et al. introduced a evaluation of approaches to accurately and precisely identify crowd sentiment in social media data (Tweets) in emergency situation [142]. SVM was used to classify the lexicons linked to the seed. Their technique performed better than Bayesian Networks alone, and the combination with Bayesian networks improved the sentiment detection.

5.2 Challenges with data analysis aspects

5.2.1 Data collection

Scalability issues. Large crises often generate an explosion of social media activity. In case of Twitter, although each message contains 140 characters, is around 4KB by considering the metadata. Furthermore, the significant amount of storage space may be required by attached multimedia objects such as images and videos. Data velocity is a more challenging issue, especially with considering frequent occuring drastic variations. The largest peak of tweets during a natural hazard was measured as 16,000 tweets per minute [1]. Finally, redundancy is in general cited as a scalability challenge. Repeated (e.g., retweets) messages are common in time-sensitive social media, even un-trusted tweet such as rumors and spam might gain more concern due to simply repeating more.
Content issues. Even although microtexts are brief and informal, to analysis this type of text is difficult work due to complexity with technological, cross-lingual and cross-cultural factors. This causes severe challenges to computational methods and can lead to poor and misleading results. Additionally, the texts are also highly heterogeneous with multiple sources and varying levels of quality. Quality itself is important question, encompassing many attributes including objectivity, clarity, timeliness, conciseness, and so on.

5.2.2 Information extraction

Inadequate spatial information. As spatial and temporal information, are two components of an event, most systems encounter challenges to determine geographical information of social media that lack GPS information. In this case, additional information (e.g., geo-tag and locations in user profile) can be used.
Combining manual and automatic annotation. In a supervised learning setting, data labeled through manual works is necessary to training a model, but it may be costly to obtain. This is particularly problematic in emergency that attract concerns of a multilingual population, or for tasks that require domain knowledge, related to affected region or characteristics of emergency events. Also, labeled data

---

[1] During Hurricane Sandy in 2012: http://www.cbsnews.com/news/social-media-a-news-source-and-toolduring- superstorm-sandy/.

are not always reliable and may not be available at the time of the emergency. In this case, a hybrid approach that combines human and automatic annotation can be used. An active learning with the selection of items to be labeled by humans can be applied to improve classification accuracy as new labels are received. When at the time Optimal budget allocation and active learning in crowdsourcing.

5.2.3 Data filtering

Mundane events. People post specific events as well as daily life on social media sites. These data as noise, which creates more challenges for an event detection methods to overcome, should be separated of real-life big events like emergencies. **Rumors, spam, and social bots.** Filtering social media data in crowdsourcing is a necessary process before data usage in any stages of a disaster. The data which is overwhelmed with unwanted content (e.g., rumors, spam, and content created by social bots) does not show real opinion of the crowd. To overcome challenge by rumors, methods have been proposed to automatically detect tweets by using their specific behaviors such as the difference of diffusion process of rumors and normal posts, the number of users and the depth related the diffusion, etc. In case of spam, characteristics of spammers can be used, such as posting numerous messages by one account and the few numbers of reciprocal connections. For solve the bot issues, three major methods have been proposed through manual annotation, using the suspension mechanism of social media sites, or creating lure bots.

5.2.4 Data integration

Describing the events. Creating descriptions or labeling for a detected event are in general challenge tasks. Although major keywords that are frequently posted during the event are presented as a description, it does not constitute a grammatically well-formed. In addition, with other useful data such as maps, images, and video, even this issue will be more complex, but it may be practical and helpful to understand event on the whole.

Domain adaptation. Simply reusing an existing classifier trained on past data does not perform well in practice, as it yields a significant loss in accuracy even when emergencies have a lot of common elements. In machine learning, domain adaptation is a series of methods adapting it to continue to fulfill well on a dataset with different characteristics. Furthermore, this techniques may help to integrate a variety of data from different domain to supplement weaknesses each other. For instance, briefly, sensor-based approach that provides more detail situation can be integrated with crowdsourcing-based methods, which support to detect event in broad area, to more fast and accurate event detection for time-sensitive tasks in emergency.

From these challenges, several interest topics are able to be raised as opportunities such as:

– Annotation: What is the best way to collect and aggregate labels for unlabeled data from the crowd? How can we do the annotation in the most cost-efficient manner? What is the most effective way to collect probabilistic data from the crowd? How can we collect data requiring global domain knowledge via crowdsourcing?

– Time-sensitive and complex tasks: How can we design crowdsourcing systems
  to handle (near) real-time or time-sensitive tasks? How can we deal with work
  dependencies requiring more complexity?
– Data collection for specific domains: How can ML researchers apply the crowd-
  sourcing principles for different domains where privacy and original character-
  istics are at play?
– Reliability, efficiency, and scalability of a system: How can we deal with sparse,
  noisy and large number of label classes such as tagging images for Deep Learn-
  ing based computer vision algorithms? How can we efficiently applied a series of
  useful methods (e.g., optimal budget allocation, label aggregation algorithms,
  and active learning) into crowdsourcing disciplines?

5.3 Example: Crowdsourcing and Machine Learning for Tracking Emergency

In this section, we more deeply contemplate, with several examples (tweets in
Twitter) for challenges and opportunities discussed in the previous section. Table
2 shows the tweet examples which will be discussed in this section.

Table 2 Tweet examples related to emergency

No	Tweet
E1	BIG EARTHQUAKE!!!
E2	Massive quake in Tokyo
E3	#havetoacemysatexam this made me laugh, no idea why but omfg haha http://bit.ly/S6fbBt
E4	ATTENTION ANYONE LOOKING FOR A JOB: FEMA needs assistance for South Jersey! $1000.00/per l (904)797-5338!
E5	#Sandy RUMOR CONTROL: The rumor that @fema is offering $300 cash cards for food is FALSE. http://www.fema.gov/hurricane-sandy-rumor-control
E6	Lets not forget the people in #NepalEarthquake let us all say a prayer for the people in Nepal MrAlMubarak

Discovering location of social media data. Geographical coordinate (known
as geotagging) attached in message is useful for a number of tasks in disaster
response [31]. For instance, it allows to search or verify information about a lo-
cal event, by filtering the messages corresponding to a particular affected region;
further geotagging can also be used for higher-level tasks, such as predict transmis-
sion of infectious disease. Unfortunately, only 2% of emergency-related messages
include GPS coordinates in practice, further, large portion of these messages may
be made through the social bot [143]. For example, a place of the tweet E2 is
able to be speculated from it's content about Tohoku earthquake, while E1 don't
contain explicit coordinates. In this regard, from users profile, some information
such as a home location (freeform text), preferred language and time zone may be
considered for determining the location of the tweet [53].
Trustworthiness of crowdsourced data. Crowdsourced data reflects opinion
of crowd; which sometimes contains more than the credible data. On the other
hand, within crowdsourcing using social media service, rumor, spam and bots

can be critical issues as aforementioned. Among the 8 million tweets according to Boston Marathon Bombing in 2013, around 29% were revealed to be rumors and 51% to be general opinions and comments [144]. This insight shows the perils of using keyword or hashtag-based topic definition. Additionally, as an example of spam (E3), many tweets which generated in areas close to the occurrence of Hurricane Sandy had been related to promoting the material for Scholastic Assessment Test (SAT) together with *#HaveToAceMySATExam* tag [13]. In case of E4, it may be also confirmed by other tweets which notice whether or not a rumor is true like E5. For retrieving the relevant tweets, using the expert knowledge to compose high-precision queries has been emphasized as one possible solution [145]. **Integration between different disciplines.** Images and videos in tweets may open new opportunities to deeply understand situation and event in emergency. Information source such as website URLs, photos, and videos in tweets related to emergency has been found to be around 18% [144]. Once images from the photos and videos are obtained, vision-based analysis can be used to find sources of the multimedia [146]. E6[2] is one example related to Nepal earthquake in 2015, and contains one image for broken apartments. The photo in tweet E6 can be searched by Google Image Search to link with one Youtube video[3] which contain a "Nepal" as a title.

6 Conclusions

In the contemporary society, various emergencies take place more and more frequently and have threatened to human life, environmental protection, social stability, and even political relationship of all countries around the world. Machine learning techniques have been proven to successfully support the decision making processes in managing many complex problems. In that sense, emergency management is no exception; however, it presents a variety of challenge to machine learning techniques for the emergency management.

The purpose of this chapter is to discuss a hybrid crowdsourcing and real-time machine learning approaches to rapidly process large volumes of data for emergency response in a time-sensitive manner. We separated the emergency stage as three phases from various definitions of the emergency, and each phase contains two tasks were applied machine learning techniques. And we then reviewed the application and the approach of machine learning techniques to support the emergency management for the each task, and the challenges and the opportunities were proposed. We described characteristics of data being generated during disaster, and discussed that the data characteristics akin to 5 Vs of big data. In addition, various applications cases of big data analysis were looked over. Moreover, we focused on crowdsourcing with machine learning in emergency management, and the their challenges and opportunities were discussed in terms of the data analysis. Finally, several examples of the tweet related to emergency were discussed to more deeply contemplate the challenges and opportunities.

---

[2] https://twitter.com/saflaher/status/598247877295644674
[3] https://www.youtube.com/watch?v=aNZiLYEr6to

References

1. N. Chen, L. Wenjing, B. Ruizhen, A. Chen, Artificial Intelligence Review pp. 1–38 (2017)
2. C. Castillo, *Big Crisis Data: Social Media in Disasters and Time-Critical Situations* (Cambridge University Press, 2016)
3. G.D. Haddow, J.A. Bullock, D.P. Coppola, *Introduction to Emergency Management* (Butterworth-Heinemann, 2017)
4. F.E.M. Agency, Federal response plan (FRP). Tech. rep., Federal Emergency Management Agency (1999)
5. A. Chen, N. Chen, H. Ni, et al. Modern emergency management theory and method (2009)
6. L. Zheng, C. Shen, L. Tang, T. Li, S. Luis, S. Chen, in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ed. by C. Apté, J. Ghosh, P. Smyth (ACM, 2011), pp. 283–291
7. A.T. Zagorecki, E.J. David, J. Ristvej, International Journal of Emergency Management **9**(4), 351 (2013)
8. S. Akter, S.F. Wamba, Annals of Operations Research pp. 1–21 (2017)
9. M.I. Jordan, T.M. Mitchell, Science **349**(6245), 255 (2015)
10. S.B. Kotsiantis, I. Zaharakis, P. Pintelas, Emerging artificial intelligence applications in computer engineering **160**, 3 (2007)
11. T. Hastie, R. Tibshirani, J. Friedman, in *The elements of statistical learning* (Springer, 2009), pp. 485–585
12. L.P. Kaelbling, M.L. Littman, A.W. Moore, Journal of artificial intelligence research **4**, 237 (1996)
13. T.H. Nazer, G. Xue, Y. Ji, H. Liu, ACM SIGKDD Explorations Newsletter **19**(1), 46 (2017)
14. J.H. Sorensen, Natural Hazards Review **1**(2), 119 (2000)
15. T. Brants, F. Chen, A. Farahat, in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval* (ACM, 2003), pp. 330–337
16. G.D. Saxton, O. Oh, R. Kishore, Information Systems Management **30**(1), 2 (2013)
17. H. Shah, R. Ghazali, in *Developments in E-systems Engineering (DeSE)* (IEEE, 2011), pp. 312–317
18. M. Moustra, M. Avraamides, C. Christodoulou, Expert systems with applications **38**(12), 15032 (2011)
19. S.W. Kim, J.A. Melby, N.C. Nadal-Caraballo, J. Ratcliff, Natural Hazards **76**(1), 565 (2015)
20. J. Reyes, A. Morales-Esteban, F. Martínez-Álvarez, Applied Soft Computing **13**(2), 1314 (2013)
21. R.R. Sahay, A. Srivastava, Water resources management **28**(2), 301 (2014)
22. H.S. Kim, C.H. Ho, J.H. Kim, P.S. Chu, Journal of Climate **25**(13), 4660 (2012)
23. K. Pabreja, International Journal of Computer Science and Information Technologies (IJCSIT) **3**(1), 2996 (2012)
24. M. Salehi, L.I. Rusu, T. Lynar, A. Phan, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2016), pp. 245–254
25. X.Y. Zhang, X. Li, X. Lin, in *Advanced Materials Research*, vol. 989 (Trans Tech Publ, 2014), vol. 989, pp. 1570–1573
26. X. Shao, X. Li, L. Li, X. Hu, in *Advances in Electronic Engineering, Communication and Management Vol. 2* (Springer, 2012), pp. 145–150
27. B. Zmazek, L. Todorovski, S. Džeroski, J. Vaupotič, I. Kobal, Applied Radiation and Isotopes **58**(6), 697 (2003)
28. Y. Su, S. Chelluboina, M. Hahsler, M.H. Dunham, in *Proceedings of the 2010 IEEE International Conference on Data Mining Workshops (ICDMW)* (IEEE, 2010), pp. 98–105

29. F. Martínez-Álvarez, A.T. Lora, A. Morales-Esteban, J.C. Riquelme, in *Hybrid Artificial Intelligent Systems - 6th International Conference, HAIS 2011, Wroclaw, Poland, May 23-25, 2011, Proceedings, Part II*, *Lecture Notes in Computer Science*, vol. 6679, ed. by E. Corchado, M. Kurzynski, M. Wozniak (Springer, 2011), *Lecture Notes in Computer Science*, vol. 6679, pp. 287–294

30. X. Dong, D. Pi, Natural Hazards and Earth System Sciences **13**(12), 3211 (2013)

31. K. Ravikumar, A.R. Kannan, (2018)

32. M.S. Tehrany, B. Pradhan, M.N. Jebur, Journal of Hydrology **504**, 69 (2013)

33. W. Chen, X. Xie, J. Peng, J. Wang, Z. Duan, H. Hong, Geomatics, Natural Hazards and Risk **8**(2), 950 (2017)

34. Q. Kong, R.M. Allen, L. Schreier, Y.W. Kwon, Science advances **2**(2), e1501055 (2016)

35. M. Bahrepour, N. Meratnia, M. Poel, Z. Taghikhaki, P.J. Havinga, in *Proceedings of the 2nd International Conference on Intelligent Networking and Collaborative Systems (INCOS)* (IEEE, 2010), pp. 507–512

36. M. Bahrepour, N. Meratnia, M. Poel, Z. Taghikhaki, P.J. Havinga, International Journal of Space-Based and Situated Computing **2**(1), 58 (2012)

37. B. Revilla-Romero, J. Thielen, P. Salamon, T.D. Groeve, G. Brakenridge, Hydrology and Earth System Sciences **18**(11), 4467 (2014)

38. C.Y. Kao, C.H. Loh, Structural Control and Health Monitoring **20**(3), 282 (2013)

39. V.V. Krzhizhanovskaya, G. Shirshov, N. Melnikova, R.G. Belleman, F. Rusadi, B. Broekhuijsen, B. Gouldby, J. Lhomme, B. Balis, M. Bubak, et al., Procedia Computer Science **4**, 106 (2011)

40. B. Pengel, V. Krzhizhanovskaya, N. Melnikova, G. Shirshov, A. Koelewijn, A. Pyayt, I. Mokhov, et al., IAHS Red Book **357**, 445 (2013)

41. M. Avvenuti, S. Cresci, P.N. Mariantonietta, A. Marchetti, M. Tesconi, in *Proceedings of the 2014 IEEE international conference on Pervasive computing and communications workshops (PERCOM workshops)* (IEEE, 2014), pp. 587–592

42. E. Fersini, E. Messina, F.A. Pozzi, Journal of Ambient Intelligence and Humanized Computing **8**(1), 37 (2017)

43. K. Mori, T. Nakamura, J. Fujimura, K. Tsudaka, T. Wada, H. Okada, K. Ohtsuki, in *Proceedings of the 13th International Conference on ITS Telecommunications (ITST)* (IEEE, 2013), pp. 337–343

44. H. Higuchi, J. Fujimura, T. Nakamura, K. Kogo, K. Tsudaka, T. Wada, H. Okada, K. Ohtsuki, in *Proceedings of the 43th International Conference on Parallel Processing Workshops (ICCPW)* (IEEE, 2014), pp. 349–354

45. D. Pohl, A. Bouchachia, H. Hellwagner, Multimedia Tools and Applications **74**(11), 3901 (2015)

46. W. Xu, L. Liu, W. Shang, Online Information Review **41**(4), 487 (2017)

47. X. Song, Q. Zhang, Y. Sekimoto, R. Shibasaki, in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (ACM, 2014), pp. 5–14

48. M. Beyreuther, C. Hammer, J. Wassermann, M. Ohrnberger, T. Megies, Geophysical Journal International **189**(1), 602 (2012)

49. J. Ye, T. Kobayashi, X. Wang, H. Tsuda, M. Masahiro, IEEE Transactions on Emerging Topics in Computing (2017)

50. J.P. Singh, Y.K. Dwivedi, N.P. Rana, A. Kumar, K.K. Kapoor, Annals of Operations Research pp. 1–21 (2017)

51. C. Caragea, N. McNeese, A. Jaiswal, G. Traylor, H.W. Kim, P. Mitra, D. Wu, A.H. Tapia, L. Giles, B.J. Jansen, et al., in *Proceedings of the 8th international conference on information systems for crisis response and management (ISCRAM2011)* (Citeseer, 2011)

52. T. Sakaki, M. Okazaki, Y. Matsuo, in *Proceedings of the 19th international conference on World wide web* (ACM, 2010), pp. 851–860

53. T. Sakaki, M. Okazaki, Y. Matsuo, IEEE Transactions on Knowledge and Data Engineering **25**(4), 919 (2013)

54. H.N. Teodorescu, Procedia Engineering **107**, 325 (2015)

55. S. Vieweg, A.L. Hughes, K. Starbird, L. Palen, in *Proceedings of the 28th International Conference on Human Factors in Computing Systems, CHI*, ed. by E.D. Mynatt, D. Schoner, G. Fitzpatrick, S.E. Hudson, W.K. Edwards, T. Rodden (ACM, Atlanta, Georgia, 2010), pp. 1079–1088

56. F. Alam, M. Imran, F. Ofli, in *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (ACM, 2017), pp. 601–604

57. D.T. Nguyen, F. Alam, F. Ofli, M. Imran, Computing Research Repository **abs/1704.02602** (2017)

58. F. Alam, F. Ofli, M. Imran, International Journal of Human–Computer Interaction **34**(4), 311 (2018)

59. J. Li, Z. He, J. Plaza, S. Li, J. Chen, H. Wu, Y. Wang, Y. Liu, Proceedings of the IEEE **105**(10), 1900 (2017)

60. S. Shen, N. Murzintcev, C. Song, C. Cheng, Information Discovery and Delivery **45**(4), 220 (2017)

61. J.R. Ragini, P.R. Anand, V. Bhaskar, International Journal of Disaster Risk Reduction **27**, 556 (2018)

62. H. Li, D. Caragea, C. Caragea, N. Herndon, Journal of Contingencies and Crisis Management **26**(1), 16 (2018)

63. S.D. Ramchurn, T.D. Huynh, Y. Ikuno, J. Flann, F. Wu, L. Moreau, N.R. Jennings, J.E. Fischer, W. Jiang, T. Rodden, et al., in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (International Foundation for Autonomous Agents and Multiagent Systems, 2015), pp. 533–541

64. M. Imran, P. Mitra, C. Castillo, in *Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016* (2016)

65. J. Yin, A. Lampert, M. Cameron, B. Robinson, R. Power, IEEE Intelligent Systems **27**(6), 52 (2012)

66. N. Pandey, S. Natarajan, REMOTE SENSING AND GIS (2012)

67. N. Pandey, S. Natarajan, in *Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (IEEE, 2016), pp. 1352–1356

68. V. Sadhu, G. Salles-Loustau, D. Pompili, S. Zonouz, V. Sritapan, in *Proceedings of the IEEE International Conference on Autonomic Computing (ICAC)* (IEEE, 2016), pp. 251–256

69. A. Trekin, G. Novikov, G. Potapov, V. Ignatiev, E. Burnaev, arXiv preprint arXiv:1803.00397 (2018)

70. Y. Bai, C. Gao, S. Singh, M. Koch, B. Adriano, E. Mas, S. Koshimura, IEEE Geoscience and Remote Sensing Letters **15**(1), 43 (2018)

71. A. Vetrivel, N. Kerle, M. Gerke, F. Nex, G. Vosselman, (2016)

72. D.T. Nguyen, F. Ofli, M. Imran, P. Mitra, in *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (ACM, 2017), pp. 569–576

73. N. Attari, F. Ofli, M. Awad, J. Lucas, S. Chawla, in *Proceedings of the 2017 IEEE International Conference on Data Science and Advanced Analytics, DSAA* (IEEE, 2017), pp. 50–59

74. D.M. Jiang, Z.B. Li, in *Applied Mechanics and Materials*, vol. 263 (Trans Tech Publ, 2013), vol. 263, pp. 3288–3291

75. G. Cervone, E. Sava, Q. Huang, E. Schnebele, J. Harrison, N. Waters, International Journal of Remote Sensing **37**(1), 100 (2016)

76. G. Cervone, E. Schnebele, N. Waters, M. Moccaldi, R. Sicignano, in *Seeing Cities Through Big Data* (Springer, 2017), pp. 443–457

77. Y. Zhang, H.V. Burton, H. Sun, M. Shokrabadi, Structural Safety **72**, 1 (2018)

78. B.T. Pham, K. Khosravi, I. Prakash, Environmental Processes **4**(3), 711 (2017)

79. Z. Wang, C. Lai, X. Chen, B. Yang, S. Zhao, X. Bai, Journal of Hydrology **527**, 1130 (2015)

80. A.R. Joshi, I. Tarte, S. Suresh, S.G. Koolagudi, in *IEEE transaction on Global Humanitarian Technology Conference (GHTC)* (IEEE, 2017), pp. 1–7

81. D.K. Yoon, S. Jeong, in *Quantitative Regional Economic and Environmental Analysis for Sustainability in Korea* (Springer, 2016), pp. 123–140

82. M.A. Zanini, F. Faleschini, P. Zampieri, C. Pellegrino, G. Gecchele, M. Gastaldi, R. Rossi, Structure and Infrastructure Engineering **13**(9), 1117 (2017)

83. M. Izadi, A. Mohammadzadeh, A. Haghighattalab, Journal of the Indian Society of Remote Sensing **45**(6), 965 (2017)

84. B. Resch, F. Usländer, C. Havas, Cartography and Geographic Information Science pp. 1–15 (2017)

85. A. Nadi, A. Edrisi, International Journal of Disaster Risk Reduction **24**, 12 (2017)

86. M. Poblet, E. García-Cuesta, P. Casanovas, Information Systems Frontiers pp. 1–17 (2017)
87. P.M. Landwehr, K.M. Carley, in *Data mining and knowledge discovery for big data* (Springer, 2014), pp. 225–257
88. D. Yang, D. Zhang, K. Frank, P. Robertson, E. Jennings, M. Roddy, M. Lichtenstern, Personal and Ubiquitous Computing **18**(8), 2025 (2014)
89. R. Dubey, Z. Luo, M. Xu, S.F. Wamba, in *Proceedings of the IEEE International Conference on Data Science and Data Intensive Systems (DSDIS)* (IEEE, 2015), pp. 520–524
90. S. Murali, V. Krishnapriya, A. Thomas, in *Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), IEEE* (IEEE, 2016), pp. 264–268
91. T.H. Nazer, F. Morstatter, H. Dani, H. Liu, in *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (IEEE, 2016), pp. 1410–1413
92. H. Purohit, C. Castillo, F. Diaz, A. Sheth, P. Meier, First Monday **19**(1) (2013)
93. H. Bai, G. Yu, Natural Hazards **83**(2), 1177 (2016)
94. P. Harris, J. Anitha, in *Proceedings of the IEEE International Conference on Industrial and Information Systems (ICIIS)* (IEEE, 2017), pp. 1–6
95. M. Imran, C. Castillo, J. Lucas, P. Meier, J. Rogstadius, in *ISCRAM* (2014)
96. J. Liang, P. Jacobs, S. Parthasarathy, (2018)
97. T. Li, N. Xie, C. Zeng, W. Zhou, L. Zheng, Y. Jiang, Y. Yang, H. Ha, W. Xue, Y. Huang, S. Chen, J.K. Navlakha, S.S. Iyengar, ACM Computing Surveys **50**(1), 1:1 (2017)
98. K. Grolinger, E. Mezghani, M. Capretz, E. Exposito, Managing big data in cloud computing environments pp. 183–209 (2016)
99. A. Salfinger, S. Girtelschmid, B. Pröll, W. Retschitzegger, W. Schwinger, in *Proceedings of the 48th Hawaii International Conference on System Sciences, HICSS* (Kauai, Hawaii, USA, 2015), pp. 153–162
100. A.H. Tapia, K. Bajpai, B.J. Jansen, J. Yen, L. Giles, in *Proceedings of the 8th International ISCRAM Conference* (2011), pp. 1–10
101. C. Howard, D. Jones, S. Reece, A. Waldock, Procedia engineering **159**, 141 (2016)
102. A. Amaye, K. Neville, A. Pope, Journal of Decision Systems **25**(sup1), 76 (2016)
103. B. Fahimnia, J. Sarkis, H. Davarzani, International Journal of Production Economics **162**, 101 (2015)
104. S. Mehrotra, X. Qiu, Z. Cao, A. Tate, IEEE Intelligent Systems **28**(4), 5 (2013)
105. B.T. Hazen, C.A. Boone, J.D. Ezell, L.A. Jones-Farmer, International Journal of Production Economics **154**, 72 (2014)
106. V. Hristidis, S.C. Chen, T. Li, S. Luis, Y. Deng, Journal of Systems and Software **83**(10), 1701 (2010)
107. P. Boccardo, F.G. Tonolo, in *Engineering Geology for Society and Territory*, vol. 5 (Springer, 2015), pp. 17–24
108. R. Agarwal, V. Dhar, Information Systems Research **25**(3), 443 (2014)
109. M. Lenzerini, in *Proceedings of the Twenty-first ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, ed. by L. Popa, S. Abiteboul, P.G. Kolaitis (ACM, 2002), pp. 233–246
110. O. Okolloh, Participatory learning and action **59**(1), 65 (2009)
111. M. Imran, C. Castillo, J. Lucas, P. Meier, S. Vieweg, in *Proceedings of the 23rd International Conference on World Wide Web* (ACM, 2014), pp. 159–162
112. S. Kumar, G. Barbier, M.A. Abbasi, H. Liu, in *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*, ed. by L.A. Adamic, R.A. Baeza-Yates, S. Counts (The AAAI Press, 2011)
113. Q. Huang, G. Cervone, D. Jing, C. Chang, in *Proceedings of the 4th International ACM SIGSPATIAL Workshop on Analytics for Big Geospatial Data, BigSpatial@SIGSPATIAL 2015*, ed. by V. Chandola, R.R. Vatsavai (ACM, 2015), pp. 1–6
114. F. Alamdar, M. Kalantari, A. Rajabifard, Computers, Environment and Urban Systems **56**, 68 (2016)
115. G.C. Ohlmacher, J.C. Davis, Engineering geology **69**(3-4), 331 (2003)
116. B. Zmazek, M. Živčić, J. Vaupotič, M. Bidovec, M. Poljak, I. Kobal, Applied Radiation and Isotopes **56**(4), 649 (2002)
117. D. Yamazaki, F. O'Loughlin, M.A. Trigg, Z.F. Miller, T.M. Pavelsky, P.D. Bates, Water Resources Research **50**(4), 3467 (2014)
118. B. Lehner, P. Döll, Journal of Hydrology **296**(1-4), 1 (2004)

119. F. Pappenberger, P. Matgen, K.J. Beven, J.B. Henry, L. Pfister, et al., Advances in Water Resources **29**(10), 1430 (2006)
120. S. Bontemps, P. Defourny, E.V. Bogaert, O. Arino, V. Kalogirou, J.R. Perez, (2011)
121. M.C. Peel, B.L. Finlayson, T.A. McMahon, Hydrology and earth system sciences discussions **4**(2), 439 (2007)
122. J. Brown, O.F. Jr, J. Heginbottom, E. Melnikov, *Circum-Arctic map of permafrost and ground-ice conditions* (US Geological Survey Reston, 1997)
123. B. Lehner, C.R. Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, et al., Frontiers in Ecology and the Environment **9**(9), 494 (2011)
124. R.C. Hoetzlein, Leonardo **45**(2), 113 (2012)
125. J. Salamon, C. Jacoby, J.P. Bello, in *Proceedings of the 22nd ACM international conference on Multimedia* (ACM, 2014), pp. 1041–1044
126. K.J. Piczak, in *Proceedings of the 23rd ACM international conference on Multimedia* (ACM, 2015), pp. 1015–1018
127. M. Imran, Aidr: Artificial intelligence for disaster. Ph.D. thesis, Qatar Computing Research Institute (2014)
128. A. Olteanu, C. Castillo, F. Diaz, S. Vieweg, in *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media (ICWSM)* (2014)
129. A. Olteanu, S. Vieweg, C. Castillo, in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (ACM, 2015), pp. 994–1009
130. C. Mariano, E. Morales, in *Advances in Artificial Intelligence* (Springer, 2000), pp. 290–299
131. B. Truong, C. Caragea, A. Squicciarini, A.H. Tapia, Proceedings of the Association for Information Science and Technology **51**(1), 1 (2014)
132. T. Spielhofer, R. Greenlaw, D. Markham, A. Hahne, in *Proceedings of the 3rd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)* (IEEE, 2016), pp. 1–6
133. A. Kurkcu, F. Zuo, J. Gao, E.F. Morgul, K. Ozbay, in *Proceedings of the 65th Annual Meeting of Transportation Research Board* (2017)
134. A. Zubiaga, H. Ji, Social Network Analysis and Mining **4**(1), 163 (2014)
135. B. Anbalagan, C. Valliyammai, in *Proceedings of the IEEE 23rd International Conference on High Performance Computing Workshops (HiPCW)* (IEEE, 2016), pp. 50–59
136. R. Dimitroff, L. Schmidt, T. Bond, Project Management Journal **36**(1), 28 (2005)
137. J. Howe, Wired magazine **14**(6), 1 (2006)
138. A. Doan, R. Ramakrishnan, A.Y. Halevy, Communications of the ACM **54**(4), 86 (2011)
139. S.B. Liu, Computer Supported Cooperative Work (CSCW) **23**(4-6), 389 (2014)
140. M. Imran, S. Elbassuoni, C. Castillo, F. Diaz, P. Meier, in *Iscram* (2013)
141. P. Balena, N. Amoroso, C.D. Lucia, in *International Conference on Computational Science and Its Applications* (Springer, 2017), pp. 675–686
142. A. Nagy, J. Stamberger, in *Proceedings of the 9th International ISCRAM Conference* (2012), pp. 1–9
143. S.H. Burton, K.W. Tanner, C.G. Giraud-Carrier, J.H. West, M.D. Barnes, Journal of medical Internet research **14**(6) (2012)
144. A. Gupta, H. Lamba, P. Kumaraguru, in *eCrime Researchers Summit (eCRS), 2013* (IEEE, 2013), pp. 1–12
145. A. Ghenai, Y. Mejova, arXiv preprint arXiv:1707.03778 (2017)
146. M. Bica, L. Palen, C. Bopp, in *Proceedings of the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)* (2017), pp. 1262–1276