

# Disaster and Pandemic Management Using Machine Learning: A Survey

Vinay Chamola<sup>1</sup>, Senior Member, IEEE, Vikas Hassija<sup>2</sup>, Sakshi Gupta<sup>3</sup>, Adit Goyal<sup>4</sup>, Mohsen Guizani<sup>5</sup>, Fellow, IEEE, and Biplab Sikdar<sup>6</sup>, Senior Member, IEEE

**Abstract**—This article provides a literature review of state-of-the-art machine learning (ML) algorithms for disaster and pandemic management. Most nations are concerned about disasters and pandemics, which, in general, are highly unlikely events. To date, various technologies, such as IoT, object sensing, UAV, 5G, and cellular networks, smartphone-based system, and satellite-based systems have been used for disaster and pandemic management. ML algorithms can handle multidimensional, large volumes of data that occur naturally in environments related to disaster and pandemic management and are particularly well suited for important related tasks, such as recognition and classification. ML algorithms are useful for predicting disasters and assisting in disaster management tasks, such as determining crowd evacuation routes, analyzing social media posts, and handling the post-disaster situation. ML algorithms also find great application in pandemic management scenarios, such as predicting pandemics, monitoring pandemic spread, disease diagnosis, etc. This article first presents a tutorial on ML algorithms. It then presents a detailed review of several ML algorithms and how we can combine these algorithms with other technologies to address disaster and pandemic management. It also discusses various challenges, open issues and, directions for future research.

**Index Terms**—Crowd evacuation, disaster management, healthcare, machine learning (ML), pandemic management, social distancing.

## I. INTRODUCTION

OVER the last decade, more than 2.6 billion humans have suffered from catastrophic disaster outbreaks, such as tsunamis, floods, earthquakes, cyclones and landslides, and various pandemics. Disasters have been the cause of several fatalities in the past, one of the deadliest disasters was an earthquake in New Guinea which left around 58 300 people displaced according to the displacement tracking matrix (DTM) [1]. The floods that took place in China in July 1931,

Manuscript received October 16, 2020; revised November 26, 2020; accepted December 10, 2020. Date of publication December 15, 2020; date of current version October 22, 2021. This work was supported in part by Singapore Ministry of Education Academic Research Fund Tier 1 under Grant R-263-000-D63-114. (Corresponding author: Vinay Chamola.)

Vinay Chamola is with the Department of Electrical and Electronics Engineering & APPCAIR, Birla Institute of Technology and Science at Pilani, Pilani 333031, India (e-mail: vinay.chamola@pilani.bits-pilani.ac.in).

Vikas Hassija, Sakshi Gupta, and Adit Goyal are with the Department of Computer Science and IT, Jaypee Institute of Information Technology, Noida 201304, India (e-mail: vikas.hassija@jiit.ac.in; sakshigupta5000@gmail.com; aditgoyal@hotmail.com).

Mohsen Guizani is with the Department of Computer Science and Engineering, Qatar University, Doha, Qatar (e-mail: mguizani@ieee.org).

Biplab Sikdar is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 119077 (e-mail: bsikdar@nus.edu.sg).

Digital Object Identifier 10.1109/JIOT.2020.3044966

caused 4 000 000 deaths that are yet the highest number of deaths from a natural disaster. Disasters are usually physical environmental changes, whereas pandemics refer to the rapid spread of a disease over a wide area. There also have been several pandemic outbreaks across the world. To name a few, there's the American plague (16th century), yellow fever in Philadelphia (1793), H1N1 swine flu (2009–2010), Ebola pandemic (2014–2016), and the recent COVID-19 (2019–Present). The deadliest one was Black death (14th century) that had spread from Asia to Europe, causing many deaths. Recently, disaster and pandemic management has become one of the hotspot areas for research. There are some recent and important works done for prevention and management of COVID-19. Samuel Lalmuanawma and Chhakchhuak [2] presented a detailed review of applications of machine learning (ML) and artificial intelligence (AI) for tackling this pandemic. This article focuses on screening, predicting, forecasting, contact tracing, and drug development for the COVID-19 pandemic. Disasters can be either natural or man made [3]. Such situations are spontaneous and complex, risking human lives, the environment, and the economy of a country. Therefore, any nation would like to opt for the most efficient and accurate algorithms to control such ordeals. The methodology used to predict the foreseen consequences of a disaster or pandemic plays a significant role in its management. With more accurate predictions and understanding, we can utilize our resources more efficiently. The advances in computer science have made available a large volume of data for disaster management authorities. Such data is often unstructured, making it challenging to clean and process such high volumes of data.

To date, many people suffered greatly because of the lack of a proper disaster and pandemic management system. A proper prediction of a disaster could not be done, and victims were not evacuated on time from the disaster outbreak area. People were not provided with mitigation measures post-disaster. Also, during the pandemic, efficient steps could not be followed to prevent further spread of the outbreak. To address such issues, this article provides a detailed review of all the existing procedures and techniques that can be employed during the post, and predisaster period to minimize the losses as much as possible.

The systems developed for assisting with disaster prediction need to be robust enough to handle the challenges that can affect a disaster management system [4]. For instance, in case of a sandstorm or a hurricane, the vision of a system can

be lowered due to the presence of dust particles, or, in hazy conditions, autonomous driving system should be safe [5]. Another challenge could be the loss of communication during a disaster. Furthermore, other challenging tasks include maximizing the number of people protected during a disaster or a pandemic, evacuating people at the right time, identifying the vulnerable areas for the spread of a pandemic, reaching the most affected people/areas and providing them with sufficient resources, evaluating the loss to the economy, and many more [6]. Decision-makers are often provided with large volumes of data and need to make predictions and decisions as quickly as possible [7], [8]. Deep learning techniques can be used for image classification and 3-D segmentation for medical purposes [9].

ML has recently emerged as one of the key computing technologies and is increasingly being applied in day-to-day life, and various industrial domains [10]. ML is an application of AI that uses algorithms that work on characteristics of available data to make further predictions. Nowadays, in the era of various other emerging technologies, such as unmanned aerial vehicles (UAV), Internet of Things (IoT), and satellite-based technology, the network is becoming more autonomous. Such systems require several local decisions to be made, such as bandwidth selection, data rate selection, power control, and user association to a base station. We can use ML algorithms to address these issues and lower human intervention in uncertain and stochastic environments. To summarize, ML algorithms have the following advantages over other existing technologies.

- 1) ML algorithms can easily process a high volume of data and can use it to identify trends. Moreover, the ML algorithms help in easily analyzing various types of data. The application of ML in day-to-day life, such as traffic predictions, video surveillance, online customer support, has also increased its popularity.
- 2) Rule-based technologies in ML can help in detecting fake messages. The use of ML algorithms reduces the need for human intervention and decision making. Such techniques help prevent rumours, especially in the case of man-made disasters.
- 3) The performance of ML algorithms tends to improve as the data increases. For example, in an earthquake prediction model, when the data increases, the algorithm's ability to predict also increases.
- 4) ML algorithms can handle multidimensional data and detect outliers in the data-set. In situations dealing with extreme hazards, outlier analysis is an important technique. Rather than removing all these outliers, special attention should be given to them when we are trying to predict highly unlikely events like disasters or pandemics.

Various ML algorithms can be employed in such cases to make fast and reliable decisions. The further subsections of this article discuss how these algorithms can be applied to make decisions with better accuracy. These algorithms have a wide variety of applications and can help in decreasing human intervention [11]. Moreover, certain trends can also be identified to make better predictions based on previous data. These

algorithms can also be applied to detect and break the chain of the spread of diseases.

Although there are several surveys related to the use of ML algorithms, very few focus on the applications of ML in disaster and pandemic management. Sughasiny and Rajeshwari [12] presented a survey of ML models and big data analytics in the healthcare field. Shirahatti *et al.* [13] presented the use of ML algorithms to detect plant disease. Nguyen [14] give an overview of the use of AI for management of the COVID-19 pandemic by using it for image processing, data analytics, etc. Similarly, Chakraborty *et al.* [15] discussed comprehensive studies that use IoT in smart healthcare, AI, and big data analytics with a prime focus on COVID-19 pandemic. To the best of our knowledge, there is no detailed survey that exclusively focuses on the use of ML algorithms for disaster and pandemic management. In this article, we present a basic tutorial of ML algorithms and a comprehensive literature review on the applications of ML algorithms for disaster and pandemic management. The main contributions of this work are as follows.

- 1) We discuss the various ML algorithms that can be used in different phases of disaster management to construct deployable models.
- 2) We present a detailed review of various technologies used in disaster and pandemic management and how ML algorithms can be used by them for this end.
- 3) We discuss how different ML models integrated with other technologies can be deployed at various phases of disaster and pandemic management.
- 4) We carry out an assessment of the open issues, challenges, and future research directions for ML-based disaster and pandemic management.

The remainder of this article is organized as follows. Section II presents a tutorial on ML algorithms. Section III reviews the applications of ML in predisaster scenarios. Section IV presents the applications of ML to determine crowd evacuation routes. Section V reviews ML related works in post-disaster management. Section VI presents the applications of ML in pandemic management. Challenges, open issues, and future research directions are discussed in Section VII. Finally, Section VIII presents the conclusion of this article.

## II. MACHINE LEARNING OVERVIEW

In this section, we present a basic overview of various ML algorithms that can be used for disaster and pandemic management.

### A. Supervised Learning

In supervised ML algorithms, the training data provided to the computer is labelled, and a set of expected output results are provided. We expect the machine to learn the pattern from this data and predict the output values for new data inputs.

1) *Classification Techniques*: Classification techniques are used in the estimation of membership of the community for data instances, and classify a data item into one of many predefined classes [16]. Some of the classification techniques are mentioned below:

- a) *K-Nearest Neighbors (KNN)*: K-nearest neighbours (KNN) - KNN is a location-based approach used for classification and sometimes regression. The authors of [17] have compared KNN with other algorithms to detect Acute Respiratory Infections (ISPA). Similarly, [18] has used KNN to detect Influenza in patients.
- b) *Support Vector Machine (SVM)*: Support Vector Machine (SVM) - This algorithm works by identifying a hyper-plane that classifies the data points. SVM aims to find a plan with the maximum range. To reduce costs, authors of [19] have used SVM to choose the best representative crowd reflecting pilgrims' behaviour. In fact, in the model in [17], SVM showed the highest accuracy in comparison to other algorithms for detecting whether a person is suffering from ISPA. Furthermore, the authors of [20] have proposed the use of SVM to classify aerial images into flood-affected and non-flood affected areas.
- c) *Naive Bayes*: Naive Bayes - Naive Bayes classifiers are a set of Bayes Theorem-based classification algorithms. The authors of [21] have proposed naive Bayes combined with UAV technology to assess post-disaster damage. Also, authors of [22] have combined Deep Belief Networks (DBN) with naive Bayes to detect the user's location. In models mentioned in [23] and [24], naive Bayes outperformed all other classifiers to classify tweets.

2) *Regression Techniques*: The regression technique is a predictive learning feature that maps a data object to a predictive variable with real meaning [25]. Following are a few common regression techniques:

- a) *Logistic Regression*: Logistic regression is based on probability and uses sigmoid as its cost function. Authors of [23] and [24] have used logistic regression to extract useful post-disaster information from tweets. Furthermore, authors of [26] have used logistic regression to determine the survival rate of people during a disaster.
- b) *Decision Trees*: In a decision tree, features are represented by internal nodes, decision rules by branches, and outcomes by leaf nodes. Authors of [27] have used decision trees for sandstorm detection. Furthermore, the authors of [28] have used decision trees to determine the location of the user during a pandemic.
- c) *Bayesian Ridge Regression*: A Bayesian approach is a probabilistic method for estimating statistical models. An application of this model is given in [29], where the authors have used Bayesian ridge regression to predict the number of people in an area.
- d) *Random Forest*: Random forest is an ensemble (multiple models combined) model technique in which multiple decision trees are trained together to produce one output. Authors of [30] have used a random forest to detect changes post-disaster. Similarly, authors of [31] have used them to detect damages to buildings. The random forest model in [32] outperformed the rest with 95% accuracy in predicting the number of people infected with influenza in public places.
- e) *Gradient Boosting*: Boosting is a method by which weak learners are converted into strong learners. Authors of [32] have used gradient boosting to predict the number of people infected by Influenza. Also, authors of [33] have used gradient boosting to patterns for people who left from the disaster-affected area.
- f) *Artificial Neural Networks (ANN)*: ANNs are fully-connected, multi-layer neural networks. The authors of [34] have proposed the use of ANN to identify or calculate a storm's intensity. Also, ANN, in integration with IoT technology, can be used to detect the user's location [35].
- g) *Deep NNs*: Neural networks lack creativity, whereas DNN is a better and adapted model that can be used for various applications that require creative conclusions. The authors of [36] have used DNN to determine a crowd evacuation route, and the model achieved 78% accuracy. Moreover, the authors of [37] propose a deep neural architecture which can be used for crowd evacuation. Also, authors of [38] and [39] have used DNN to predict the number of people in an area.
- h) *CNNs*: CNN can classify images with better accuracy and are better at capturing orientation than ANN. Authors of [40] have proposed CNN to highlight the risk in areas after a disaster. Authors of [41] use an approach based on CNN to identify the areas affected by disaster and achieve 81-90% accuracy. Furthermore, the authors of [42] have used a CNN-based model for internet of healthcare things (IoHT).returning

### B. Unsupervised Learning

Unlike supervised machine learning algorithms, the data fed to the computer in unsupervised learning is unlabeled. These algorithms try different techniques to detect patterns or explore the structure of information.

1) *Clustering*: In clustering, a common descriptive function [43] in which a finite set of clusters are used to classify the data is pursued. Few common clustering techniques are as given below:

- a) *K-Means*: Authors of [44] used K-means to detect which areas were damaged and affected by flood. Authors of [45] have used K-means for clustering crowd behavior. Similarly, it is used to classify the crowd into different classes by the authors of [46]. Authors of [47] used it to predict the spread of cholera disease.
- b) *K-Medoids*: It is another clustering technique that is based on dissimilarities between data points. It works well with outliers. Authors of [48] have used the K-medoids algorithm to plan an evacuation route.
- c) *Fuzzy C-Means*: This also begins by randomly assigning centroids and initializing the data points randomly to clusters. Authors of [49] have proposed a hybrid model of Fuzzy C-means with a neural network to classify Landsat images, that can be used in predicting disasters.

### C. Reinforcement Learning

Reinforcement learning is a type of machine learning that can be regarded as a self-sustainable system, learning iteratively. In Q-learning, decision making happens sequentially.

Authors of [48] and [49] have used reinforcement learning to determine an evacuation route.

### III. APPLICATIONS OF MACHINE LEARNING MODELS IN PREDISASTER MANAGEMENT

Disaster management aims to reduce the impact of a disaster and save as many lives as possible. This section presents a comprehensive review of ML models in association with other technologies, such as IoT, UAVs, geodesics, satellite, remote-sensing, and smartphone-based that are used for predicting a disaster, crowd evacuation, and post-disaster scenarios.

#### A. Predicting Disaster

If a disaster is predicted in time, warning signals can be floated to people, and they can take the necessary safety measures. The accurate prediction of a disaster lies in analyzing spatial and temporal data of an area and predicting the characteristics of a disaster, such as the water level of a flood or the magnitude of an earthquake. In the subsequent sections, we will discuss the use of different technologies in disaster prediction and how various ML-based models can help in enhancing the efficiency of other methods in predicting disasters.

1) *IoT and ML-Based Models for Disaster Prediction:* IoT is an arrangement of interconnected machines used to collect data over a network without human interference [60]. It has enabled the deployment of a large number of applications in various fields. Saha *et al.* [61] presented the use of sensors on trees for predicting fire outbreaks in forests. The sensors measure features, such as temperature, CO levels, greenhouse gases, and moisture. Similarly, microwave sensors can be used to analyze the earth's movements for the prediction of earthquakes.

ML algorithms can be used to process the data collected by IoT devices and provide better accuracy. Flood management raises a variety of obstacles for IoT-based strategies. It involves a complex set of parameters with multiple interdependencies, including rainfall, pressure, and rate of flow. The sensor network has to be configured according to the parameters selected. The number of sensors and their interconnection relies significantly on the river's measurements. Anbarasan *et al.* [62] presented an approach in which ML is used for detecting a flood. It uses Hadoop MapReduce for the removal of duplicate values. Then the rules are generated based on four attributes, namely, rain sensor, humidity, water flow sensor, and water level sensor. The rules are further used in a CNN. The proposed NN classifies the data into positive and negative for the occurrence of floods.

The model in [63] uses multiple sensors connected to analog to digital converter (ADC) like temperature, moisture, water level, and CO level. The data is sent to a raspberry-pi device in digital form. Then, this data is analyzed with the  $Q$ -learning algorithm based on penalty and reward.  $Q$ -values are used to analyze the risk of any activity in mines. Similarly, the model in [64] uses ANN and logistic regression for rainfall prediction. An integrated framework focused on IoT and ML is proposed in [65] to forecast flood risk in

a river basin. To capture data, the software uses a revamped mesh network interface via ZigBee to the WSN and a GPRS module to transfer data over the Internet. The data sets are then analyzed using an ANN. The findings of the study indicate a significant change in the approaches already in use. The selection of sensor network deployable areas is also a big challenge. Additionally, certain QoS parameters, such as latency, instability, severity thresholds, etc., should be checked before implementation. Hassija *et al.* [66] have explained how various ML algorithms can be implemented to address the security concerns of IoT models.

The recent development of UAVs has solved the accessibility issue for areas that are difficult to reach and monitor because of human and logistic constraints [67]. They amplify satellite images with observation gaps and are a scalable and portable solution that can provide high-resolution images. However, the cost of UAVs makes their deployment an obstacle for developing countries. Koh and Wich [68] and Hassija *et al.* [69], [70] presented cost-effective drones that have a longer range and bigger payload and can be effectively used for mapping real-time data and monitoring nearby activities. Also, conservationdrones.org, a nonprofit organization, is actively building low-cost UAVs without compromising any of its features. Yim *et al.* [71] presented a Smart-eye model wherein UAVs use LTE to communicate with Smart-eye centers. This model uses image stitching technology in which it compares the real-time images with the previous original images of an area. It uses a closed-circuit television (CCTV) camera, which makes it cost effective. The minimum altitude and velocity of UAVs need to be always regulated in this model. Furthermore, this model lacks a larger flight time. The model in [72] presents a solution to tackle these parameters using a Web-based platform in the case of Rotolan landslides. Furthermore, the model in [73] uses mobile nodes to provide resilience when a node fails during a disaster. A sensor node, connected to UAV and acting as a mobile node is deployed in [74] and is capable of communicating with WSN in the area. The limitation of this model is that it cannot survive during extreme weather conditions, and its low battery life is still a challenge. Though, after utilizing a mobile anchor node for localizing unknown nodes, WSN provides us with a better fault tolerance than other UAV models [75]. It focuses on predicting floods and deploys a pressure sensor that is used to easily estimate the height of the water level. Flooding is predicted based on observation of a sudden increase in water level. Also, this model needs less aerial space and can easily land on irregular grounds due to its rotary wings as compared to conventional fixed wings. With the use of UAVs, there is a threat to the privacy of the user's data because it is deployed in an open atmosphere. Alladi *et al.* [76] proposed that we use blockchain for ensuring the data security.

Remote sensing (RS) images are extensively utilized in diverse fields. After the collection of images from drones, they need to be classified efficiently. ML algorithms are well suited for this task. In the work proposed in [77], water features are distinguished from nonwater features to identify flooded areas from Landsat 7 (L7) images incorporated with digital elevation model (DEM) data.

Unlike [77], the model presented in [84] uses passive Landsat 7 and active RadarSat images to analyze floods and develop a flood hazard map. The classification of water and land areas are easy in RadarSat images, and results were compared with Landsat images. High-resolution Landsat images could be used to provide information about floods, except for the monsoon season due to the cloud cover. However, RadarSat images can analyze during the monsoon season also. A map highlighting all the information about the risk of floods was developed using the above and GIS data.

Akshya and Priyadarsini [24] used a hybrid of SVM and K-means to detect flooded areas. It classifies aerial images collected by drones into flood-affected and nonflood affected. This model achieved an accuracy of 92%. This model may also be deployed on drone sensors so that they can automatically detect the areas.

Singh *et al.* [56] aimed at forming a three-band image with higher resolution, also known as a fused image, using Brovey transform. Principal component analysis (PCA) is used to reduce this to a 1-D image. Then, a neuro-fuzzy approach that uses KCN and FLCIM algorithms are used. The FCM model depends on the value of  $m$  (fuzziness index) and leads to wrong results. The use of this hybrid approach, called fuzzy Kohonen local information C-means (FKLICMs), overcomes the limitations of both these models. This model also shows a higher accuracy than the FCM approach. The high accuracy and nonparametric nature of random forest classifier can be used for land cover classification [36]. The RF classifier, due to its advantages mentioned in Section II in this article, shows better accuracy. The key downside of the model is that because of the numerous classification trees produced by the resampling of the same data set, it becomes challenging to grasp the rules used to produce the final classification.

This has contributed to the advent of object-based image processing (OBIA) methods that have been developed to resolve these problems. The RS community has made significant efforts for almost 20 years now to encourage the usage of object-based technologies for land-cover mapping [88]. Supervised algorithms combined with object-based detection methods have always been an important part of land-cover mapping RS work since 2010 [89]. Many choices, such as segmentation system, accuracy evaluation, classification algorithm, training data, input characteristics, and objective groups must be chosen for object-based classification processes. To address these requirements, many authors have proposed supervised object-based classification methods particularly accommodated to each use case. Ma *et al.* [87] highlighted that Landsat series RS images are mostly used in controlled object-based detection due to their strong quality and usability. Moreover, the blurry rule-based classification strategy meets a plateau in object-based classification, while supervised object-based classification reaches a production height. RF shows the best results in an object-based grouping, followed by SVM. In fact, NNs remain unsuitable for more comprehensive usage in the description of objects. However, this approach still needs to be verified for urban areas.

The model presented in [37] uses a random forest combined with object-oriented image analysis (OOA) for landslide

detection tested on sample data sets of Haiti, Italy, China and France, which were recently affected by landslides. Integrated variable measurement and selection processes, and high-quality program applications that are freely available, allow random forests to integrate with OOA, as illustrated in Fig. 1. The authors found a clear over-prediction of landslide areas for all situations if a class-balanced sample was utilized in training. Over-prediction was more prominent for Barcelonnette (France) and Messina (Italy), where visual inspection of the photographs already indicated a large difference in class when compared to the other two regions. More work is required to refine the segmentation process, which is currently focused solely on spectral details. Therefore, a preliminary sample-based estimation of the variable value can be an important method for determining which additional layers would be included in the segmentation.

Demaria *et al.* [90] performed an analysis to observe intensity guidance accuracy of tropical cyclones over a period of two decades. The models proposed to forecast the strength of the hurricane up to 120 h advance, but forecasts of severity are less reliable relative to forecasting the path of the storm [90]. The weather researched forecast (WRF) is a regional predicting framework that monitors sandstorms dependent on the previous day's environmental conditions. Shaiba *et al.* [32] predicted whether, in the following hours, a sand storm would occur or not, by utilizing ML algorithms, such as classification and regression tree (CART) analysis, logistic regression, and Naive Bayes classifiers. The system classifies sandstorm or no-sandstorm events and sends alert signals to people. The results show that CART analysis performed better than Naive Bayes and regression-based classifiers.

CNNs are resilient to darkness, and capable of sufficiently and most critically obtaining the catastrophe attribute to resolve misleading by variables, which would impact the efficacy of disaster prediction. By trimming and scaling aerial imagery gathered from Google Earth Aerial Imagery, Amit and Aoki [40] generated training data patches for before and after disasters. Both patches are qualified using CNN to isolate the area which is recognized as a catastrophe zone. The accuracy was 80%–90% for this system. Tang *et al.* [85] presented an *a priori*-based association rule algorithm for detecting the intensity of tropical cyclones. According to the associativity algorithm, if items  $X$  and  $Y$  are selected at random, we can be certain that item  $Z$  was also selected. For the association rule, the user specifies minimum support. The issue with this model is that we need to decrease the minimum support and multiterm associations need to be selected. To solve this issue, Buranasing and Prayote [42] used symbolic aggregate approximation (SAX) combined with ANN to approximate the intensity of storms by using eight features. These features were reduced by piecewise aggregate approximation (PAA) and were given as inputs to NNs. The F-measure for this model was 0.93.

With the recent development of airborne-LiDAR, it is extensively used in diverse fields, such as medical, military, agriculture, pollution, and architecture [91]. It provides 3-D high-resolution samples of earth's topographical features and

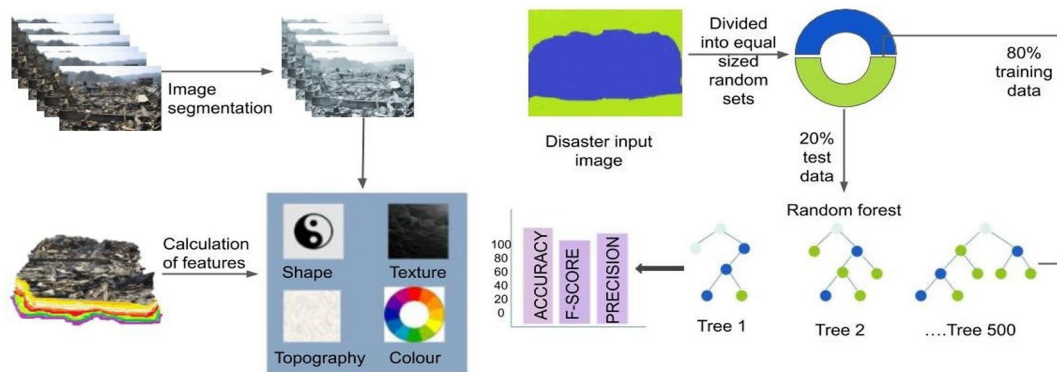


Fig. 1. Object oriented image analysis integrated with random forests [37], [87].

other objects. These detected features may be used to predict disasters, such as floods, landslides, earthquakes, etc. with better accuracy. Geodesic curves are used to estimate the mean curvature for each vertex [82]. It can detect terrain features according to a variety of surfaces: 1) pass; 2) channel; 3) pit; 4) ridge; and 5) peak. Then, the Gaussian-weighted average is computed for those curvatures. The currently used method in [82] can be sped up by using a multiresolution mesh. Also, more features could be taken into account, such as normals and principal curvatures.

Earth observational evidence in mapping disaster management is also used by group on earth observations (GEOs) and the committee on earth observation satellites (CEOSs). The application infrastructure provider (AIP) process of GEO has evolved and recently adopted the reference model of open distributed processing (RM-ODP) [92]. The global Earth observation system of systems (GEOSSs) (2005–2012) [93] aims at disaster reduction and management. By combining the earth's observations with other information, it aims at spreading early warning signals, better preparedness, and risk assessment for rebuilding houses post-disaster. CEOS is working on increasing the use and raising awareness of Earth's observation (EO) data for disaster management. It has undertaken a project to reduce the use of *ad hoc* arrangements in the management system that have become an obstacle in processing EO data. Authors of [94] presented us with a solution for these obstacles, in which research and improved EO to help tackle the issues encountered by remote agencies in processing satellite data. However, some things remain unclear, like the resources shared, their interdependencies, the effect of shape data policies, and how new users can access services [83]. Percivall *et al.* [83] presented a CEOS GA.4.D architecture with an emphasis on using ground controlled interception (GCI) to address issues, such as user authorization and combining socio-economic data with EO for better observations. It also identifies gaps in services and focuses on limited resources for victims. However, this architecture requires more unrestricted access to disaster data. Attempts to use this architecture at regional and local levels are in progress.

Emergency departments (EDs) are usually equipped with syndromic surveillance systems that generally use statistical detection algorithms like CUMulative SUM (CUSUM) [95],

auto-regressive integrated moving average (ARIMA), the Holt-Winters algorithm, and the early aberration reporting system (EARS). The performance of these algorithms depends on a tradeoff among false-positive rate, sensitivity, and timeliness. The work by authors of [96] shows that the models mentioned above require intensive parameter tuning, and simply using the preset parameters is not always optimal for a given algorithm or data set. Moreover, they showed that data needs to be examined regularly to fine-tune the parameters in these algorithms according to the data used. These observations limit the abilities of small syndromic systems to predict a widespread outbreak.

### B. Early Signals

If early signals are delivered to people in an area at the right time, numerous causalities can be prevented. Such systems can be installed in offices, homes, or other public places to warn people in time. Sensor networks are recommended for the systems to track disasters (e.g., the Building Administration System) and address this issue [97], [98]. Such devices need a wide range of separate sensors, such as smoke, heat, radiation, etc., installed in advance in houses, and such technologies may have significant costs to be incurred.

Yawut and Kilaso [99] have proposed a program to suggest a model that is specialized for weather detection. However, we can only utilize such systems in limited instances where the sensor grid has formerly been installed. Even then, if the system cannot operate due to power loss, sensor damage, broken transmission lines, and so on, then this sensor network cannot function properly. The model in [25] aims to improve a previous model emergency response support system (ERESS) based on mobile *ad hoc* network (MANET) using ML [100]. ERESS mobile terminal (EMT) accumulates data from its sensors. Every EMT uses its personal sensors, which allows it to operate everywhere. The disaster-recognition algorithm method is described as follows.

- 1) An EMT uses SVM for detecting some temporary risk.
- 2) If the EMT senses a temporary risk, it floods the details to those in the area.
- 3) Each EMT regularly measures temporary risk identification accumulation.

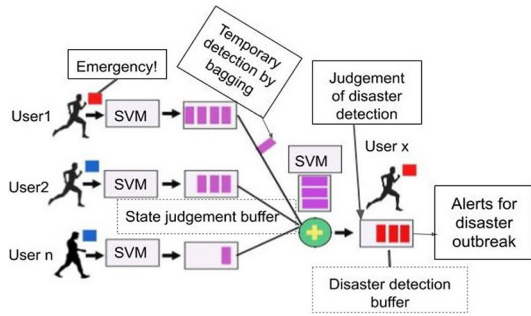


Fig. 2. Disaster prediction by BB-SVM [25].

- 4) Every EMT matches the logistic model with a regression curve.
- 5) A disaster is determined if the curve is convergent to the values of all EMTs within a building in a period of 1 min or less. This method of detection also makes the system autonomous.

The buffering and bagging SVM (BB-SVM) model's EMT [25] has two buffers: 1) state judgment buffer and 2) disaster detection buffer. The model comprises of two phases. The first phase is shown in Fig. 2. First, each EMT generates inputs from all sensors and transfers them to a SVM model. When the input data is determined to be that of a crisis, the EMT amasses the outcome in its state judgment buffer. The outcome stays in the buffer before time to live (TTL) runs out. If the reserve for the state judgment is reached to the limit, the EMT must transmit about it being in a crisis. Second, the EMT accumulates the temporary risk in its disaster detection buffer, ignoring other EMTs. Often, the temporary risk stays in the buffer before TTL runs out. If the buffer for disaster prediction accumulates to the amount of half the number of EMTs, a disaster outbreak is determined by the EMT. The TTL duration of a buffer may be adjusted through bagging learning. For this purpose, the BB-SVM changes the temporary warning or disaster outbreak parameters according to specific circumstances with less data than the traditional approach. The EMT in this model changes the judgment of the disaster outbreak repetitively. Also, there is the issue that EMT relies on the activity of just the individual to determine the occurrence of a catastrophe. Finally, ERESS cannot distinguish the incident caused by an unusual sound or light.

Sikder *et al.* [86] presented another approach that uses a mobile application to alert people. It uses a partition-based trajectory distance to find the nearest shelter place during evacuation. SQLite database is used, and this application processes the JavaScript object notation (JSON) file and delivers messages through notification via the app. This model lacks the ability to deliver early signals during immediate effect disasters, such as cyclones, floods, etc.

*Summary:* In this section, we reviewed the various applications of ML models in disaster management. The various algorithms and models used for disaster prediction and communication are presented. The approaches using ML algorithms for this are summarized in Table I.

#### IV. APPLICATIONS OF MACHINE LEARNING MODELS IN CROWD EVACUATION

Delays in the evacuation during a disaster often lead to increased casualties. This often happens because the evacuees do not recognize the routes of catastrophe occurrence and evacuation. To minimize the number of casualties caused by such incidents, it is necessary to quickly identify catastrophe locations and to figure out as soon as possible any appropriate routes for evacuation. However, it is challenging to identify the catastrophe occurrence and in very limited time, e.g., 30–40 s during earthquakes, to direct individuals to suitable escape routes. Moreover, traditional approaches have not been able to assist evacuation immediately following the onset of a catastrophe. More rapid evacuation is needed for sudden disasters. For example, carbon monoxide can spread quickly during fires in houses [101].

Artificial potential field (APF) can be used to develop a crowd lives oriented track and help optimization system (CLOTHO) for crowd evacuation [102]. This app uses a mobile terminal (IoT side) for collecting data and cloud for storing data. This model has four main modules: data collection, network transmission, cloud, and user platform. Attraction and repulsion potential field from a disaster point are calculated. The APF calculated is further controlled by the shelter distance threshold. The resultant force is used to determine the direction of evacuation. Unlike [102], the model mentioned in [103] involves two layers: sensor layer, which is used to collect data and uses the MTS400 CC sensor, and the IoT layer, in which real-time data is processed using resource scheduling and banker's algorithm. The limitation of this model is that developments are still needed in a heterogeneous scenario. ML algorithms can be used to process a large amount of data in these models.

Mobile phones are an important component used to communicate with victims during a disaster [104]. The development of GPS has made it easier to track the location of victims. During a disaster, the network is often congested due to a lot of SMS messages, and thus, cell broadcasting service is used to contact victims [64]. This further reduces panic. For evacuation of people, Rossnagel and Scherner [105] presented a global system for mobile communications (GSM) alarm device to be deployed in the nearby area for early warnings. This device collects its data from the weather office, but this does not provide the residents with a quick way for evacuation. An area mail disaster information system for tsunami warning and evacuation system to assist fishermen is presented by Jayasinghe *et al.* [106] and Torii *et al.* [107]. This system tracks fishing boat locations using GPS and provides them with warning signals through mobile phones. This is still a faster approach than the previous one.

Different from [105], Anzengruber *et al.* [108] have proposed a smartphone app that notifies people about the evacuation routes. They have used two parameters to evaluate the location of the user and proximity-relations using Bluetooth. The data is processed through a SVM model that outperforms the linear regression and Gaussian models. However, this model requires testing with real-world data during a disaster. Wang *et al.* [55] have proposed an improved reinforcement learning approach for crowd evacuation. The

TABLE I  
SUMMARY OF RELATED WORKS FOR DISASTER PREDICTION AND EARLY SIGNALS (ACRONYMS USED IN THE TABLE- *CEOSs: DEM, NOIR:*  
NETWORK OPTIMIZATION INDICATOR REPORT)

Category	Reference	Target issue	Technology used	Hardware/API used	Case studies
IoT and ML	[53], 2020	Floods prediction	Convolution neural network	Hadoop MapReduce	Surat, India
	[55], 2019	Rainfall prediction	ANN and Logistic regression	LoraWan	UEM Campus
	[56], 2016	Forecast flood risk	Artificial Neural Networks	ZigBee, WSN	None
UAV	[69], 2019	Prediction	Geofencing	UAVs Base station, Flight controller	Surat, India
	[70], 2017	Assessing areas	Adhoc network as aerial mesh network	Raspberry-pi with NOIR pi camera	UEM Campus
	[71], 2019	Disaster assessment	SWIFTERS	DJI UAV, ROS library, Map server	None
	[72], 2019	Disaster assessment	Star algorithm, Tabu search, Gradient descent	Multi-UAVs	Jiuzhai valley earthquakes
ML and UAV	[20], 2019	Detecting flood areas	SVM, K-means clustering and PCA	Drones	Aerial images
Geodesics based	[73], 2010	Better prediction	Centre surrounded filters with gaussian weighted mean centers	Air-borne LiDAR	2010 Haiti earthquake
	[74], 2013	Better prediction	GEOSS and CEOS approach	Spatio-temporal infrastructure	2008 earthquake in Sichuan, China and Namibia flood plot
Satellite	[68], 2002	Water and non-water features	Landsat 7 Thematic mapper and DEM data	None	Floods in Pitt County, North Carolina
Remote sensing	[75], 2011	Develop flood hazard maps	Supervised MCL classification	RADARSAT remote sensing data, GIS data and ground data	Maghna river basin
ML and Remote sensing	[76], 2012	Classification of land cover	Random forest classifier	Landsat-5 Thematic Mapper data	Province of Granada
Data mining	[77], 2005	Tropical cyclone intensity	Apriori-base	None	Atlantic basin
Android based	[78], 2017	Alert signals	Partition based trajectory distance	JSON file	Haiti
ML and Object sensing	[79], 2011	Prediction of floods	Random forest classifier	None	Momance River — Haiti and Wenchuan town — China
	[80], 2017	Classification of land cover	Random forest classifier and SVM	None	Scopus databases
ML	[27], 2018	Sandstorm detection	CART decision tree, Naïve Bayes and Logistic Regression	None	Riyadh, Dammam, and Jeddah
	[41], 2017	Flood and landslide detection	Convolutional Neural Network	None	Japan and Thailand
	[34], 2014	Storm intensity	Symbolic Aggregate Approximation (SAX) and Artificial Neural Network (ANN)	Satellite-image data	Typhoon and Tropical cyclones
	[81], 2016	Disaster recognition	BB-SVM	ERESS	Kansai University

main challenges which were covered were modifications in the model when there is an increase in people density and effective communication between authorities. For the initial clustering, the K-medoids algorithm is used. Deep learning algorithms could also be implemented for the same model to improve its accuracy further.

With the recent increase in the usage of smartphones, several applications have also been deployed for crowd evacuations. Rahman *et al.* [110] used an app employing OpenStreetMap (OSM), document management service (DMS), and Google C2DM server. The limitation of this app is that the data must be uploaded on OSM. This model failed when tested

in Bangladesh because of the lack of uploaded data. When a user logs into the application, he/she is provided with a unique ID that is then sent to the DMS. If the application detects a risk in the user's area, it notifies him/her. In case of risk, it also provides users with the shortest route to the shelter. Aldecimo and De Leon [112] also used OSM and store data in a PostGIS database. One of its functions, which is to calculate the minimum distance between two geographical places, is used. A parameter called building risk factor (BRF) is calculated, which denotes the danger that a building poses to the road during an earthquake. Then, the Dijkstra algorithm is used to find a safe route. This



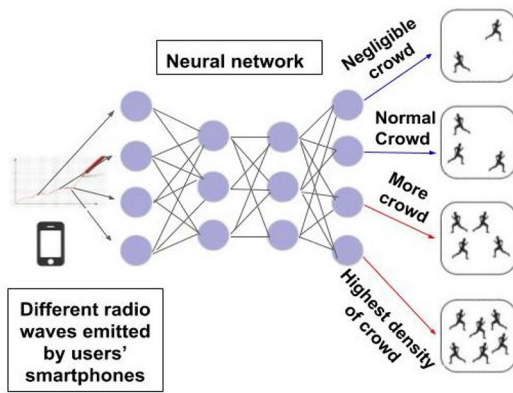


Fig. 3. DNN model to determine crowd density [44].

model only uses BRF as the parameter to decide the safest route. Furthermore, this model did not take into account the building's height. The taller a building, the more danger it would pose to the road. An application that covers all these limitations using Bayesian networks is proposed by Haddawy *et al.* [117].

The use of cameras for collecting data for crowd evacuation affects people's privacy. Thus, models are needed that work on a mechanism different from analyzing and processing camera images so that the public's privacy is not violated. Shibata and Yamamoto [44] have proposed a model that analyzes radio waves of a user's smartphone. The strength of waves is measured using sensors. Due to the complexity of the fluctuations of the device's waves, an ML-based model is necessary. The density of the crowd in an area is determined by using a deep NN (DNN), as illustrated in Fig. 3. This model achieved an accuracy of 78%. However, some other ML algorithms can be used that require a smaller training data set in comparison to this model.

Instead of Micropilot MP-Vision and other components, Ezequiel *et al.* [118] used a custom UAV designed with resources to achieve lower cost and maintain the quality level. This model's air-frame uses a Super SkySurfer fixed-wing extended prediction orbit (EPO) foam frame, ArduPilot Mega (APM) autopilot system, airspeed sensor, camera payload, and batteries. Radio telemetry is used to establish communication between the UAV and ground control station (GCS). The GCS uses APM Mission Planner that is employed to control and generate routes for the UAV during its flight. This UAV has several advantages over others: increased flight time (30–50 min), and large coverage of area (4 sq km). Also, the autopilot system is open source, which enables independent developers to be engaged in identifying issues and updating software. The disadvantage of this model is the loss of UAV due to GPS signal loss, and it lacks a safety mechanism during failure. Furthermore, it is more prone to hardware errors (assembly and wiring of parts) and bugs in code. A mechanism is also required to be added in the model to avoid risks during belly landings. An approach based on  $Q$ -learning is proposed to cover these limitations in [119]. This approach targets robust radio-network signals by not compromising on the UAV's features, such as its flight time, cost, and coverage area. Chhikara *et al.* [45] proposed a DNN architecture for

facilitating navigation of UAV in indoor environments, where GPS fails in precise localization and navigation.

The use of lane contraflow reversals is important for emergency evacuations. Cities or areas with emergency evacuation measures in effect may utilize contraflow lanes to increase the number of roads necessary for traffic to evacuate. Contraflow lane reversals reverse traffic paths along escape routes from their usual directions. Hurricane Katrina had devastated south-east Louisiana in August 2005. An evacuation plan which successfully evacuated an estimated 1.1 million citizens was implemented before landfall [114]. The reversal of the contraflow lanes was in effect for 25 h and resulted in significantly decreased traffic delays. However, ML algorithms can be implemented for locating new parts of the contraflow and establishing evacuation measures in other sensitive areas. Research in [120] explores a method focused on existing traffic patterns and path availability to enable contraflow lane reversals. To model the contraflow plan, the evacuation routes and local traffic situation graphs are used. A path assessment algorithm is used to evaluate if existing traffic patterns need contraflow lane activation. Decision trees are used to determine when the current state of the evacuation model should activate all contraflow lanes. The next steps in the proposed methodology is to expand the number of features that the training set takes into consideration, change the methodology to define specific contraflow segments, and use simulation tools to construct training sets and validate outcomes. Mori *et al.* [25] presented a BB-SVM-based model for predicting disasters. Then an evacuation path is chosen based on depth first search (DFS) and Dijkstra's algorithm.

Tracing safe routes can also be done using a multiobjective genetic algorithm (MOGA) [121]. This process involves three phases. The collecting phase involves collecting GPS and acceleration data from the user's smartphone to the cloud. The analyzing phase involves calculating the walking distance between nodes, walking speed (average), and pedestrian traffic to determine safety evaluations. With safety evaluations, time taken, and the distance between the initial and final nodes are calculated, and the maps are provided to the users. This marks the third phase. The user has a choice to choose maps according to their desired characteristics (shortest time or shortest distance or safety). The characteristics of the routes get upgraded continuously, which is a unique feature of this method compared to other works. Moreover, new maps are generated using selection and cross-over techniques that help in providing multiple evacuation route maps. The model has not yet been validated with real values of average speed and for pedestrian traffic (the authors here assumed it to be a set value in calculating safety evaluations). Also, alpha, beta, and gamma selection for the calculation of safety evaluation value have to be chosen optimistically, which remains a challenge for different types of places. Ofi *et al.* [122] proposed a hybrid of genetic algorithm and ML algorithms, such as SVM, Naive Bayes, and random forest to process UAV images for disaster management. This was successfully deployed on a wildlife reserve in Namibia. The model achieved 80% accuracy with SVM and random forest outperforming the rest.

Tracing safe routes can also be implemented through a similar approach based on Virtual Voting and Adaptive Pricing Algorithm [123]. Soni *et al.* [124] have used K-means clustering and KNN regression algorithms for the same. Data is first preprocessed and categorized into crime and accident data sets. The place is divided into clusters by latitude and longitude by K-means clustering based on crimes and accidents. “Direction service” is a class of Google API that gives all possible routes from source to destination, and way-points are assigned for every 2 km. The risk score of the way-points is calculated based on K-nearest clusters around it. A risk score is simply the summation of accident scores and crime scores determined by assigning scores to each type of crime and accident. The value of  $K$  is determined by the lowest root mean square error (RMSE). KNN scores for the way-points are analyzed using the  $R^2$  score, which is a simple measure of distance from the regression line. Experiments were conducted for Manhattan Borough in New York City, and it successfully showed alternate safe routes that are different from the default route shown by Google Maps based on the shortest distance. Previous works failed to categorize small areas and only did larger area classification with either SVM or other ML algorithms. However, this work has found safe routes covering small areas. KNN analysis showed high  $R^2$  scores of 0.910 (accident score) and 0.974 (crime score), which implies that the classification has been done well and is distant from the regression boundary. Moreover, previous works in this field were highly subjective, whereas this work is not. However, more factors have to be taken into account in calculating the safety of the route (risk score). Also, this method consumes more time in preprocessing, which can be improved (60%–80% of the entire time).

Traditional NNs encompass some constraints, such as the inability to vary the length and work on different training data. To target this issue, Jiang *et al.* [115] have proposed a long short-term memory (LSTM) model for planning crowd evacuation. The CNN layer can be utilized as a filter in this model that processes spatial data and leaves the LSTM with only temporal data. Since LSTM can be trained using complex real-time data, it achieved 95% accuracy in the reported experiments. Different approaches using supervised and unsupervised ML algorithms were presented in [21] and [116]. Radianti and Granmo [116] have used a Naive Bayes model. This model is deployed in a fire outbreak scenario. Using a Bayesian model, it determines the routes that are safe to use for evacuation. It assigns probability-based physical and emotional factors to every route which shows the safety probability of the route to be used.

Namoun *et al.* [21] have proposed a model that combines SVM with fuzzy logic. SVM works on choosing the best representative crowd group by integrating both the physical properties of an area and the characteristics of a crowd. The fuzzy rules are then created on the basis of this output of the SVM model. The authors reported accuracy of 98.77% for the model. However, the effect of psychological factors is not considered in this model.

One way to plan evacuation routes successfully can be pre-detection of the behavior of crowds so that optimal steps

can be taken. Amores *et al.* [52] have initially detected a cluster of crowds, built a feature matrix, and then used ML algorithms. The behavior of crowds is divided into three clusters: 1) the ones assembling in groups; 2) the ones approaching exit doors; and 3) the ones following leaders or responsible authority. K-means and hierarchical clustering are used for the same. Similarly, other behaviors can also be studied to plan evacuation routes.

During a fire outbreak, people generally rush toward the exit, which can create chaos and result in more casualties. To tackle this issue, Tian and Jiang [59] have proposed an evacuation plan based on reinforcement learning. On the basis of the previous movement of individuals, their next step is modified in every iteration. The objective function is defined as minimizing the evacuation time. However, situations where people wait for their friends or relatives or follow their paths, are not taken into account in this model. Also, the effects of harmful smoke generated from the fire are not taken into account.

Gupta *et al.* [39] have proposed a model which analyses video of the crowd in an area. It uses a deep CNN model to classify video features, such as the speed of the crowd and physical features of the environment to classify the crowd’s behavior as a normal situation or an alert situation. This achieves an accuracy of 77.82%. Then, three random forest classifiers were used to identify abnormal behavior in the motion of the crowd. Furthermore, multiple cameras can also be used to speed up the process. To improve the accuracy of the video-based model, Wei and Sheng [53] have proposed a new framework. The collected video data is classified in several situations by using a CNN model. This model also covers a mechanism to process distorted images by comparing them with the original image’s pixels. Furthermore, different crowd situations are determined based on entropy techniques.

*Summary:* This section has covered a survey of ML-based models to determine a crowd evacuation route. Apart from ML-based models, UAV, IoT, and android-based models are also discussed. All the models are summarized in Table II. The common requirement for all ML-based models was an initial clustering algorithm. K-means and CNNs were found to be used more commonly as compared to other ML algorithms.

## V. APPLICATIONS OF ML MODELS IN POST-DISASTER MANAGEMENT

### A. Detecting Changes

The changes post-disaster are important to analyze in order to calculate the loss to the economy, to plan specific measures for the recovery from a disaster, and rehabilitation measures. First, in order to perform rescue operations, cellular networks have to be regenerated in the affected area for communication [125], [126]. For this purpose, a framework for networks assisted by three different types of UAVs: 1) tethered Backhaul Drone (tBD); 2) untethered powering drone (PD); and 3) untethered communication drone base station (cDBS) is created. Furthermore, algorithms like ant-foraging are commonly used to drive UAVs to locations with a high density of survivors. GIS analysis and RS are then

TABLE II  
SUMMARY OF WORKS FOR CROWD EVACUATION (ACRONYMS USED IN THE TABLE- *CIoT*: CELLULAR IOT, *COCO*: COMMON OBJECTS IN CONTEXT, *LSTM*: LONG-SHORT-TERM-MEMORY, *RCNN*: RECURRENT CNN, *SSIM*: STRUCTURAL SIMILARITY INDEX)

Category	Reference	Target issue	Technology used	Hardware/API used	ML involved	Case studies
IoT	[96], 2018	Crowd Evacuation	CLOTHO (APA and APA-RF)	None	No	Chemical plant in Nanjing, China
	[103], 2018	CIoT	UAVs	None	No	Snow Avalanches
Android based	[104], 2012	Evacuation program	OpenStreetMap	DMS	No	Bangladesh, India
	[105], 2014	Rescue team tracking	Algorithm based on location	GPS-Receiver	No	Indonesia
	[106], 2016	Finding a safezone route	OpenStreetMap	PostGIS database	No	Makati, Phillipines (Earthquakes)
UAV	[107], 2019	Priority based route selection	Clustering (Bird flocking)	FANETS	No	None
	[72], 2019		Graph theory, Euler cycle and integer programming	None	No	Hospital Pavia Arcicibo
ML	[108], 2015	Activating Contraflows	Decision trees	Weka (Data mining software)	Yes	Hurricanes
	[102], 2013	Prevent crowd disasters	SVM, Linear regression and Gaussian model	GPS Sensors	Yes	San Francisco
	[81], 2016	Determine an evacuation route	BB-SVM, Dijkstra's algorithm and Depth first search (DFS)	ERESS	Yes	Kansai university
	[36], 2019		Deep Neural networks	Raspberry Pi-3 and spectrum analyzer	Yes	Ritsumeikan university
	[48], 2019		K-medoids and Reinforcement learning	None	Yes	Office scenario
	[109], 2017		LSTM model	Spatio and temporal features	Yes	Kumamoto earthquakes
	[110], 2014		Naive Bayes	None	Yes	Fire hazard
	[19], 2018		SVM and Fuzzy logic	None	Yes	Hajj (A Muslim pilgrimage event)
	[45], 2018		K-means and hierarchial clustering	Alarms	Yes	An office building
	[111], 2016		Classification of crowd situation	Deep CNN and Random Forest	None	Yes
	[46], 2019	CNN classifier and K-means		None	Yes	Video data
	[50], 2018	Planning evacuation route	Reinforcement learning	None	Yes	Hong Kong fire outbreak

used to plan and coordinate recovery. This is done by analyzing spatial differences and then providing the requisite information to those working on the ground. In the RS-based assessment, most of the developed methods focus on physical recovery. Their indicators are mostly at the building level, and thus collecting information manually becomes difficult and expensive. Furthermore, the majority of UAVs have very limited flight time, providing insufficient data for manual differentiation. Thus, incorporating techniques like ML through which machines can be trained become extremely important.

A vision-based disaster detection system using ML algorithms to detect the affected area from aerial images in real-time needs to be developed. Here, the image data would be collected by sensors mounted on UAVs, and then processed by deep learning algorithms, such as CNNs to detect the presence of a disaster such as fire, flood, and landslide in real time, and also the number of structural damages to the buildings [48]. A similar methodology using the algorithm of random forests was used in detecting damages to buildings during the 2010 Haiti earthquake [35].

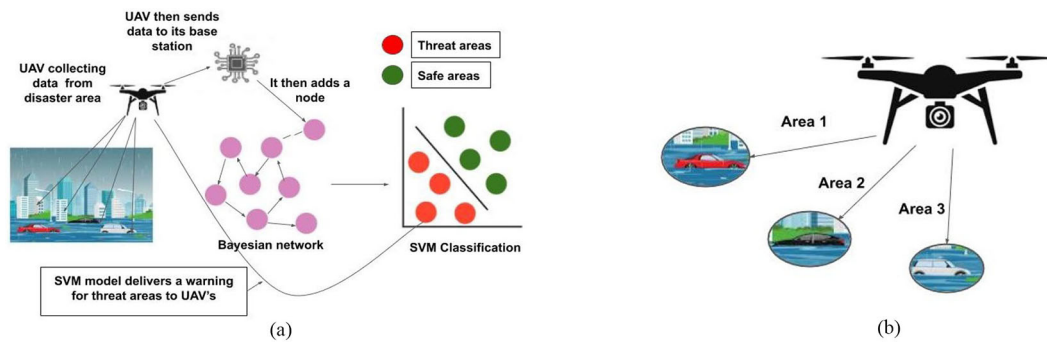


Fig. 4. UAV-ML model for disaster relief in a city [22]. (a) UAV-ML model for analyzing threats in an area. (b) When SVM model notifies UAV, it updates and divides its search based on the information provided.

Ren and Gao [22] have proposed a model in which UAVs collect the data, and it is further processed by a model built using structure-variable discrete dynamic Bayesian network and SVM. In this model, UAV collects data at a particular time interval from varying distances. As it collects new data, a new node is added in the Bayesian network with a particular probability, as shown in Fig. 4(a). This data is fed to a SVM classifier for identifying any threats in the area. If the SVM detects any threat, it signals a message to the UAV, and its time interval of collecting data is further divided into small time intervals so that the area can be evaluated accurately, as shown in Fig. 4(b). This leads to a better evaluation of an area.

ML techniques have the advantage of immediately filtering images, which would have required months to be sorted manually. Temporary settlements can also be detected, indicating areas of survivors in need. ML approaches can be used to combine data, remove unreliable data, and identify informative sources to finally generate a heat map, which identifies areas of urgent need. By application of CNNs, fire and smoke have been successfully detected from raw RGB images with an accuracy of 97.9% [48]. Also, due to their high accuracy, they prove to be a much better approach as compared to other techniques that are in popular use like SVM and random forests. CNNs have shown advantages over other ML methods. However, its hidden layer is a black box creating uncertainty and dependency on the training data.

de Alwis and So [127] used GIS analysis data to detect changes in an area post-disaster. It uses OpenStreetMap data and satellite images. Enhanced change detection index (ECDI) is used as a parameter to differentiate between pre-disaster and post-disaster images. This model is deployable in a large area. Since ECDI calculates edges, gradient, and texture, this can be further used to analyze effects in buildings. However, the issue lies in changes in the solar zenith that manipulates the data. To address these issues, we can incorporate ML to differentiate between the same. These models have a high reliance on real-time data sets and georeferencing of data. ML algorithms can work by only requiring post-disaster images and can result in better accuracy. The work in [51] shows how K-means can be used to detect areas that are affected by a disaster. It achieved an accuracy of 85% with respect to low-clarity and cloudy images data set. However, the K-means algorithm mixed the damaged and reconstructed areas during its initial clustering. The accuracy can be further improved if high-resolution images are used.

Unlike the K-means approach used in [51], Joshi *et al.* [34] used the simple linear iterative clustering algorithm (SLIC) for segmentation purposes. This method reduces the work further as the user only needs to specify the number of super-pixels. By applying the Random forest classifier, an accuracy of 90% can be obtained with the use of 400  $n\_trees$ . Since it takes the mean of all decision trees, it tends to show better accuracy. With the help of these results, rescue teams can prioritize their services to areas based on the level of damage.

Sheykhmousa *et al.* [128] used land cover information to evaluate post-disaster areas. This information has been widely used in remote-sensing technology [129]. SVM classifier is used to classify images into positive and negative recovery. Positive recovery refers to the areas which have recovered to their original state. SVM achieved 90.8% accuracy in this case. The CNN model, which is also competitive, is not used here because of the limited amount of training data (CNN models require a large amount of training data).

### B. Extracting Useful Information From Social Media Sites

Social media has become a powerful tool in today's world. Data generated on social media sites every second is huge, and this data can be employed to derive useful results. Incorporating this data can help post-disaster management teams to make better and more informed decisions. Studies have used Twitter data to concentrate on the role of development and exchange of knowledge on social networking sites to increase consciousness about the situation as well as to evaluate post-conditions. Technology is currently concentrating on using automatic approaches to obtain meaningful data during disaster times from Twitter data. Starbird and Palen [130] have differentiated data into initial and secondary from previous sources, where individuals not only shared information about their own situation but also the accessible and relevant knowledge that had been retweeted. Vieweg *et al.* [131] observed that the tweets contain contextual news that creates awareness of the disaster. The authors base their conclusions on a study of the 2009 wildfires in Texas.

Jamali *et al.* [132] used ML to evaluate people affected by a disaster, their information, and their comparison with those who were not affected. All disaster-related tweets are identified from the Twitter data set [Comment, Date, Location] using dynamic query expansion (DQE). The location from where a user tweeted more was estimated as his/her home location.

To analyze the trending topics, Dirichlet regression was used. The topics were classified according to factors, such as age, income, education, etc., and a correlation between them was observed. A similar study was conducted for people who were not affected to study the impacts on them. However, it is a bit difficult to extract reliable patterns from such a large data set.

Extracting useful information from tweets is also a helpful task. After the preprocessing of data by removing all the extras and representing it in a vector form using term frequency inverse document frequency (TF-IDF) and Bag of Words representation, classification was performed in [28]. SVM (73.66%) and logistic regression (74.58%) achieved the highest accuracy among all the classifiers. The voting classifier was also proposed, which showed a competitive accuracy (74.16%) However, a mechanism needs to be added that can verify tweets and cross-validate the location of the disaster so that there is no misguidance later for the rescuers. Assery *et al.* [29] proposed a different approach to classify tweets by checking whether they can be used as disaster-related information or not. Natural language processing (NLP) is used to clean data collected from Twitter. Count and TF-IDF vectorizer was used to convert tweets to a vector. Then, the data was fed to several classifiers. Logistic regression (99.7%) and Naive Bayes (98.8%) outperformed all other classifiers. This model can also be expanded to classify tweets on the basis of the nature of users.

Another area of research can be predicting returning patterns for the people who left their homes because of a disaster. Returning patterns need to be estimated so that further planning for resource allotment can be done. The government and disaster management authorities need to be ready to face a change in demand for vital public services, such as electric power, gas, and water. Yabe and Ukkusuri [41] used sentiments, such as the psychological state of users, the magnitude of the tweeted tragedy, and success in the recovery area to understand returning patterns. The assumption in this model is that sentiment, a manifestation of different dimensions and restoration levels, could be exploited to predict evacuee return behavior. Each tweet  $w$  is allotted a sentiment score  $S(w)$ , by connecting each word in the tweet to a dictionary and allocating a weight highlighting its emotional effect. Then, the time at which the evacuees left, the distance, and the destination they have traveled to are calculated. In addition to these parameters, it also extracts parameters from three other networks: 1) offline spatial network; 2) online personal network; and 3) online agency network. The Gradient Boost classifier is used to predict the returning behavior, which is defined as 0 for a negative return and 1 for a positive return. It outperformed all other classifiers with an accuracy of 86.4%. This is because gradient boosts can effectively model non-linear dependence between the parameters in comparison to SVM and logistic regression. However, to use this model on other disasters, certain parameters need to be changed.

### C. Minimizing Future Disaster Risk

Future disaster risk minimization is a crucial task that needs to be performed. A GIS-based emergency response database

server is used to help the emergency rescue authorities get prepared for life-saving operations in a small amount of time by predicting extreme climatic conditions [133]. This database server provides the climatic information to the relevant authorities via a satellite link. The satellite has two functions, first being RS and weather forecasting, and second is to transmit the extreme input climatic parameters received from GIS database servers to rescue and medical authorities to prepare them for life-saving procedures. However, the aforementioned technology is not well integrated with modern techniques of AI. The forwarded information is analyzed by humans, a tedious and error-prone process that can be replaced by more accurate machines.

Drought prediction systems using DBN, which can calculate and predict different drought indexes, such as the standardized precipitation index (SPI) with much higher precision is proposed in [134]. Similarly, ML techniques, such as pattern recognition NNs, recurrent NNs (RNNs), and random forest models can be used to determine relationships between calculated seismic parameters and future earthquake occurrences. Moreover, patchwise object detection techniques based on CNNs can be utilized for automatic disaster detection, such as floods and landslides [48]. Image classification based on the deep learning algorithms has been found to have much higher accuracy for landslide recognition than previously used methods. In addition, by integrating such processes, it is possible to respond to the evolving disaster situations [48]. The accuracy of landslide detection by CNN is calculated with a precision of 0.93, recall of 0.94, and F-measure of 0.93. While predicting the disaster, the major concern is to save human lives. This objective can be further accomplished by involving technologies, such as IoT in the processing. This could turn out to be a quick and alternative means of communication in the disaster-struck region, where IoT-enabled devices (battery-powered wireless devices) can be used to provide data network resilience during disaster situations [135].

*Summary:* This section covered different approaches for detecting changes post-disaster. In today's world, social media generates huge amounts of data every second and it can be extremely helpful if relevant information can be extracted from it. This section presented various models to make social media useful in disaster management. This section further reviewed ML algorithms for minimizing future disaster risks. Table III summarizes all the works for post-disaster and to minimize the future disaster risk.

## VI. APPLICATIONS OF MACHINE LEARNING MODELS IN PANDEMIC MANAGEMENT

### A. Prediction and Diagnosis of Pandemic

To prevent the outbreak of a pandemic, an early diagnosis is a must. Technologies for early diagnosis of ISPA (Acute Respiratory Infections) disease are required to reduce its adverse effects on babies and prevent it from turning fatal. Ginantra *et al.* [18] carried out a test to foretell and diagnose, whether the person is affected with ISPA or not, which is done by evaluating the evidence acquired with the usage of the

TABLE III  
SUMMARY OF WORKS FOR POST DISASTER AND TO MINIMIZE FUTURE DISASTER RISK (ACRONYMS USED IN THE TABLE-RS)

Category	Reference	Target issue	Technology used	Hardware/API used	ML involved	Case studies
RS and UAVs	[120], 2018	Collect information about survivors	Ant foraging	GIS analysis	No	800*800 area
ML and UAVs	[21], 2011	Assessing area for threats	Naive Bayes and SVM	None	Yes	Not mentioned
	[40], 2017	Highlighting areas with risk	CNN and Markov Random field to process data collected by UAVs	None	Yes	Earthquake, flood and landslides
RS	[122], 2017	Detect changes post-disaster	Enhanced change detection index (ECDI)	Openstreet map and satellite images	No	Van,Turkey (Earthquake)
ML	[31], 2018	Detecting damages to buildings	Random forest	None	Yes	Haiti (Earthquake)
	[40], 2018	Evaluating future disaster risk	CNN	None	Yes	Floods and Landslides
	[30], 2017	Detect changes post-disaster	Simple Linear Iterative Clustering algorithm (SLIC) and Random forest	Aerial images by GeoEye1	Yes	Japan (Earthquake and Tsunami)
	[123], 2019	Classify areas into positive and negative recovery	SVM	Land cover data	Yes	Tacloban,Phillipines (Typhoons)
	[44], 2019	Detect which areas were damaged and affected by flood	K-means clustering	Optical ASTER images	Yes	Tohoku (Tsunami)
Social media	[23], 2018	Extracting useful information	Logistic regression, SVM and Voting classifier	Twitter data	Yes	Chennai (Rainfall)
	[33], 2019	Predict returning pattern for people who left	Gradient Boosting		Yes	New Jersey (Hurricane Sandy)
	[24], 2019	Identifying which tweets could be used for useful information	Logistic regression and Naive bayes		Yes	Hurricane Florence and Hurricane Michael
	[127], 2019	Relation of tweets between disaster affected and not affected people	Dirichlet regression and Dynamic Query Expansion (DQE)		Yes	New York City (Hurricanes)

ML algorithms, such as SVM, NNs, KNN, and Naive Bayes. Various symptoms, such as fever, headache, flu, etc., were used by the classifiers. The confusion matrix was used to calculate the accuracy, and SVM outperformed all the algorithms. Yin *et al.* [19] predicted antigenic variants of H1N1 influenza structured on a stacking model. In the design of a stacking model based on epidemics and pandemics, three separate feature engineering methods were implemented to check its viability and universal adherence, namely, residue-based methods, ten regional band-based methods, and five epitope region-based methods. This stacking model employs a concept similar to k-fold cross-validation to build out-of-sample predictions that function for small to medium-size data sets. Accordingly, logistic regression, Naive Bayes, and NN methods were used to design the Level 2 models to get better results than single models. Also, random forest and gradient boosting were applied to create the Level 2 versions.

Such Level 2 models made projections for the outcome of all results, which were then put into the second stage of the

training data, identified as  $Y(n * n)$ . Parameter  $n$  accounts for the number of new functions in the data set. The Level 3 model was developed by the classifier of logistic regression and evaluated on  $Y(n * n)$  to generate the resulting effects of antigenic variant predictions. Because of the rapid antigenic shift or drift that cause antigenic variants to occur, the effects of specific mutations on antigenicity are not clearly known. Outputs based on various function vectors showed no difference, suggesting that producing new characteristics from residual sites was possibly not the key reason for affecting prediction. For the proper modelling of the H1N1 variation mechanism, several more factors, such as climate and the human immune system, need to be considered, which is out of focus of this model.

To solve this feature issue, Righetto *et al.* [54] used rainfall as a variable to predict the spread of cholera in Haiti. The rainfall pattern is treated as a marked Poisson process. The continuous rainfall events are processed as discrete events where the depth of rainfall is presented by the mark. K-means clustering was further used to predict the spread of disease.

The limitation of this model is the lack of ability to accurately predict the evolution of new cases. The improvement in the current situation, i.e., real-time data is also not taken into account.

Pioneering work in using Bayesian networks and NLP for influenza detection is presented in [145]. Data was collected from free-text ED medical reports comprising of 468 reports on polymerase chain reaction (PCR) positive influenza patients from the period January 1, 2008 to August 31, 2010 and 29,004 reports of patients not associated with a positive PCR test from the period July 1, 2010, to August 31, 2010. The classification architecture comprised of the NLP-finding-extraction component with Topaz used as NLP parser to encode findings into one of three values: acute, nonacute, and missing, followed by a classifier network. The authors studied seven ML classifiers and compared their diagnostic capabilities against an expert-built Bayesian classifier. Three configurations were used based on methods to deal with missing data for performance evaluation. The authors used AUROC and Brier Skill Score as a measure of classification performance while NLP-finding-extraction performance was measured using accuracy, recall, and precision.

The authors observed a tie between Naive Bayes, logistic regression, SVM, and ANN with the highest AUC being 0.93 in the configuration where all missing values were assigned "False." The NB showed superiority (BSS: 0.35, AUC: 0.93) in terms of least training time and the ability to treat missing values without preprocessing. The authors were able to show that ML classifiers performed better than Expert-made Bayesian classifiers, for a given NLP extraction system and a large amount of training data. Similar work was reported by Ye *et al.* [146] on the same data set with nine combinations of finding-extraction methods and Bayesian classifiers. The highest AUROC achieved by them was 0.79 for a combination of expert-finding with a BN-EM-topaz classifier. Moreover, they only considered AUC for classification evaluation which is susceptible to bias due to inclination of the test data with one class, potentially leading to an unskilled classifier problem. Ye *et al.* [146] only considered Bayesian networks as classifiers without a comparative study between other ML models.

Wynants *et al.* [147] have proposed a method of genome sequencing to identify SARS-CoV-2 (COVID-19). The method is based on the additional sequencing of viral complementary DNA. Sequencing data can be retrieved from the PCR of the original viral RNA which then can be used along with cDNA to identify COVID-19. Alignment methods like FASTA and BLAST are used for classification done from viral sequencing techniques. However, the major issue with this method is that it necessarily requires base sequences for the detection or classification. A detection test has been proposed that combines molecular testing with deep learning. This work applies state-of-the-art deep CNNs which can automatically create features starting with genome sequencing of the coronavirus. Samples are divided into 9:1 for training and testing using a ten-fold cross-validation scheme [148]. The data used was genome sequence data from the repository 2019 Novel Coronavirus Resource [149] and NCBI repository [150].

The results of the experiments show that the above approach correctly classifies SARS-CoV-2 and distinguishes it from other coronaviruses. The proposed method identifies COVID-19 with an accuracy of 98%, and can classify different coronaviruses with an accuracy of 98.75% [148]. There are several prediction models proposed for COVID-19 detection based on CT scan images of the chest using deep learning which have accuracy less than the above method (around 82%), and chest radiography may miss patients with pneumonia in early phase [147], [151], [152]. However, the above model has been trained on very few samples due to limited data available. Also, limited genome sequences are available and considered in this model. Santos *et al.* [153] also proposed a COVID-19 prediction model, for Brazil, using a time-sliding window algorithm.

Liang *et al.* [154] integrated current surveillance data with Internet search queries to forecast influenza epidemics. Data were obtained from the China National Scientific Data Center for Public Health along with a public Baidu search engine database based on Baidu Index for queries related to influenza. Regression was performed using SVMs, and hyperparameters were tuned by leave-one-out cross-validation. Performance metrics used for this task were RMSE, mean absolute percentage error (MAE), and root mean square percentage error (RMSPE). Correlation analysis was performed on data with different lag times and search queries with a significant correlation coefficient. Olson *et al.* [155] mention the features that were chosen for the SVM regression model. Application of Internet search query data for infectious diseases provides real-time surveillance of epidemics and overcomes the shortcomings posed by lag-time in conventional flu surveillance. A strong correlation between search terms and influenza cases was found in the study. Given the short incubation period of influenza, most of the search queries related to symptoms and medication were closely correlated with cases in the same month. The authors found that the models based on ensembled data performed better and were more robust than the models based on other single data sources. Similar work was done by Google flu trends (GFTs) [156]. However, it was concluded by Liang *et al.* [154] that GFT data could not replace conventional surveillance methods to be a reliable pandemic surveillance system. This is due to the negligence of epidemiological factors, such as geographical location, illness complaints, the age distribution of patients, or clinical manifestations in the Internet data. The correlation coefficient of the search keywords relies heavily on existing vocabulary data. However, due to the changes in the social media environment, many new search vocabularies are produced. The new vocabularies were not considered in the model.

One of the most significant aspects of pandemic management is close monitoring of the people in the affected areas. Moreira *et al.* [157] provided a comprehensive survey that includes the use of several ML algorithms for smart decision support systems for healthcare. For determining the reoccurrence of a pandemic, traditional models utilize an ARIMA model for prediction [158]. The model construction trades off between parameter estimation, diagnostic checking, and identification. Examples of usage of IoT and big

data analytics through smart wearables to collect health-related information of citizens have also been prevalent [15], [159]. Conventionally deployed flu surveillance systems are based on hospital and laboratory data, thus limiting their real-time surveillance due to lag-time [160]. The IoT-based data collection techniques assume most citizens to possess a smart wearable or sensor-enabled device. However, in developing countries, that are generally the worst hit by an epidemic, this remains far-fetched.

### B. Social Distancing

COVID-19, a major pandemic outbreak, has caused large scale devastation in the world. Within only nine months (January to September, 2020), this has spread in 213 countries already with approximately 32 million cases [161] overall. Also, this figure is increasing at a high pace everyday. Chamola *et al.* [162] highlight the role of deep learning algorithms during COVID-19 in various phases, such as disease surveillance, risk prediction, medical diagnosis, screening, and virus modeling and analysis. Since the spread of influenza, SARS, and then COVID-19, social distancing has become an important tool to handle outbreaks [163], [164]. The work in [165] presents simple methods to maintain distance between people to prevent virus transfers. Social distancing is defined as living in such a way that physical contact with people is avoided. This can be divided into two categories: community and personal. Community social distancing involves avoiding any crowd gatherings in public, such as schools, colleges, and places of worship, minimizing traveling, and sealing the areas where people have already been affected. Personal social distancing involves quarantining individuals to their homes, leaving one's home only for important purposes, and motivating others also to follow the same.

Various technologies have been used for social distancing. GPS is a technology that can assist to detect the locations of people. Wang *et al.* [166] proposed indoor wireless positioning, which does not require the use of GPS and can thus work in areas where GPS cannot. Bluetooth is a technology that works well in such cases, as it allows multiple connections at the same time. This feature used by the model proposed in [136] to evaluate the dependency between distance and radio signal strength (RSS) of any two devices. If the distance detected is less than 1.4 m, it generates a warning message to enable the users to be aware and maintain a distance. However, this model tends to fail in some scenarios due to inaccuracies in the user's location, and there is no measure proposed yet to address it [167]. Also, for this model to work, Bluetooth should always be on. Thus, there are still challenges for this technology to work upon. Sattler *et al.* [168] have proposed a novel ML algorithm integrated with Bluetooth technology to calculate the risk of COVID-19 in nearby areas.

For directing people about public social distancing, Choi *et al.* [138] proposed a model that provides a count of people in an area. This model works in a complex network and is computationally effective, which various models using ZigBee and WiFi lack. The model works by picking up various clusters, and the one with the largest size is selected. Then,

a maximum likelihood equation is derived from a probability density function plotted. Finally, the number of people is estimated based on the equation. In the same manner, the work in [139] provides a mathematical design with lower computational complexity to count the number of people. This model uses a method based on detecting energy. These models can be combined with various ML algorithms to improve their accuracy.

Due to their varied applications, UAVs have proved fruitful in many situations. Challita *et al.* [142] have used the combination of recurrent and CNNs to process data collected from UAVs. These UAVs and their data are then used to predict the traffic density in an area. This information can be provided to users and can be used to signal them about the overcrowded areas, as shown in Fig. 5(c). Hassija *et al.* [170] have proposed a model for estimating traffic jams in an area based on blockchain and DNNs. Also, Hassija *et al.* [171] have proposed a model for intervehicle communication using a directed acyclic graph and game theory.

A portable device is proposed that collects data via a radar sensor and a camera and provides it as input to NNs in [140]. The ANN is trained constantly with data to identify whether the object near it is a vehicle, a stationary object, or a human being. This model's accuracy improves with use and gets trained when the user is near an object. This model can also keep a track of the users who are violating social distancing. However, different from [140], Jia *et al.* [27] proposed another approach for detecting a users' locations that is based on their most influential friends in a social network. Initially, a DBN was implemented by choosing  $M$  top influential friends and evaluating their social influence on a user. The model showed an approximately 16.87% gap in its theoretical and practical accuracy because it randomly selects some top friends which does not seem to be an effective way. Then, a Naive Bayes model was used, and it showed approximately 19.82% accuracy increase. Naive Bayes is comparatively more scalable, and its predictions are based on probability which makes it more reliable.

Cho [20] have combined ML with smartphones' locations to warn the users regarding their safety from infected neighbors in the area. KNN and decision trees are used for determining the location of a user, and a Markov model is used for location prediction. By training the model with the real-time data set of the area, the movement of infected people can be predicted. People can get alert notifications on their smartphones, as shown in Fig. 5(a).

ANN is combined with a robot-based model to increase location accuracy in [43]. With this integration, the model becomes more scalable, cost effective, and user friendly, and it showed an error of only 2.7m. Also, radial basis function networks are used with ZigBee which has an advantage over Wi-Fi systems. Wireless networks have low accuracy while verifying a user's location and require channel characteristics data to be available. Brighente *et al.* [23] have proposed a model using SVM and ANN which outperformed the traditional wireless networks. This model has a promising accuracy, and training the models is the only additional task required here. Yin *et al.* [143] presented a better model named the federated localization (FedLoc) framework. This model is



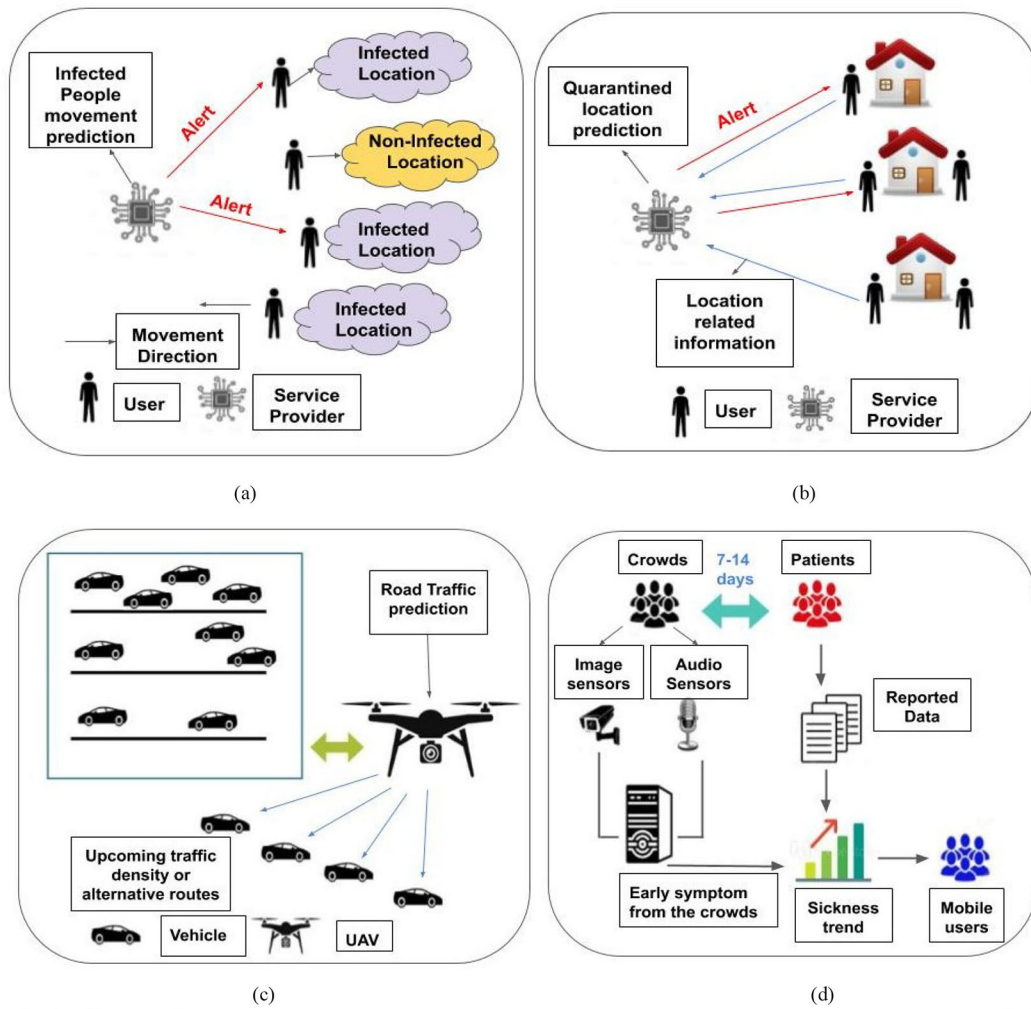


Fig. 5. Application of ML in social distancing. (a) Infected people movement prediction. (b) Quarantined people location prediction. (c) Road traffic prediction. (d) Sickness trend prediction [169].

built for verifying the user’s location, as shown in Fig. 5(b), without sacrificing their privacy, by using DNNs. This was compared with a model based on the Gaussian process, where DNN fared better because of their ability to generalize better. FedLoc is based on the integration of the DNN and Gaussian process models. To protect the user’s privacy, homomorphic encryption was used. However, to make the model more effective, the weights could be modified to a lower precision level from 64 b. Furthermore, Ahmed *et al.* [172] used a deep-learning-based framework for monitoring of the social distance protocols during COVID-19 pandemic. The model uses YOLOv3 object recognition paradigm for recognizing humans in a video footage. The use of transfer learning methodology ensures a higher accuracy for the model.

5G networks have attracted wide attention because of their low latency, ability to satisfy user’s potential needs, and requiring less energy to operate [173], [174]. Bui *et al.* [175] presented a detailed idea of how linear regression can be used in 5G networks. Polese *et al.* [33] presented a Bayesian ridge regression and random forest models to predict the number of people in an area in combination with 5G cellular networks. They also aim at using ANN to increase accuracy.

Furthermore, Alawe *et al.* [47] proposed to predict the number of people in an area using DNN and RNNs. RNN outperformed DNN as RNN showed 90% accuracy, whereas DNNs showed 80% accuracy. However, if more training data was available, then the accuracy of the DNN model could be improved further. Polese *et al.* [33] have used ML to predict the number of people in the user’s proximity. They began with clustering the base stations. Then, the prediction of the number of users is done by deploying the Bayesian model regressor, random forest regressor, and the Gaussian regressor. In order to predict the number of users who are in proximity to a defined cluster, we can use WiFi-hotspots. This architecture can also be deployed to predict the traffic density in an area so that the users can travel through areas with fewer vehicles so as to practice social distancing.

Wang *et al.* [144] presented the use of LSTM NNs integrated with 5G technology to determine the users’ path based on these historical paths. However, this model is very computation-intensive and requires a lot of training data for a single user, which is not practically feasible. Thus, a multiuser sequence-to-sequence model was proposed which is more user friendly and tackles all the former issues. This

sequence-to-sequence LSTM model outperformed SVM and linear regression. This study has open directions to combine its model with other information about the users. Bega *et al.* [46] presented another model using a mixture of DNN and CNN to build its several layers. A modified loss function is used to train this model. This is a cost-effective system and also tackles service-level agreement (SLA) violations.

Al Hossain *et al.* [38] have developed a model using ML to warn users in public places by informing them how many people near them are infected. This model was built on the basis of influenza symptoms as its detecting features. With a microphone array, it analyzes different cough sounds of users, and cameras are used to analyze the density of that area, as shown in Fig. 5(d). The random forest model showed the highest accuracy (95%). This model's accuracy can be further improved if its training data set is combined with some real-time data. This model is highly beneficial in informing people about the status of a public place during any phase of a pandemic.

### C. Extent to Which Economy Is Affected

Cho *et al.* [176] have tried to assess the economic impact of earthquakes through a hybrid model Southern California Planning Model (SCPM) which follows the Garin-Lowry model and is an integration of state-of-the-art input-output model and spatial allocation model. The work also encompasses the network model and structure performance model. The idea behind this model is to estimate economic losses by looking at the direct business and structure loss, lifeline network disruption, and spatial impacts that occurred due to the disaster. The key technology used for damage estimation is GIS and drives the state-of-the-art early post-earthquake damage assessment tool (EPEDAT) model [176]. Although these models estimate the cost of a disaster and its immediate impact on the economy, they do not explain the important long-term impact of disasters like earthquakes on the economy and growth.

Another method has been proposed which caters to the issues in [176] and estimates the sign and magnitude of both short and long-run effects of disaster on growth [177]. The study pursues a comparative event study approach, which compares the situation of affected countries post-disaster with the superficial situation of the country if the disaster would not have happened and estimates the loss. The superficial state of the country is generated by comparison from the built synthetic control group. The data used is the data on natural disasters from the emergency events database (EM-DAT), data from the Penn world tables, and real gross domestic product (GDP) per-capita at purchasing power parity (PPP). The above work's results show that only very large disasters display an impact on GDP growth in both long and short runs. Also, the results show that these impacts on the economy are driven by the event of radical political changes in the country post-earthquake, and the impacts are not significant in the absence of political changes. Results are informative about the long-run cost of disasters and the involvement of political structure in causing damage to the economy and can be used in further studies. This model works only when a set of regions is affected, and another set is totally unaffected by the disaster (i.e., control group), and

hence cannot be used in case of a global crisis. An approach using Naive Bayes can be used to calculate such an effect [178].

### D. Classifying Areas Into Red, Orange, and Green Zones According to Spread During Pandemic

Classification is generally done through ML algorithms. However, there are technologies like RS that capture sentinel-satellite images where data is collected across spectrums. However, those data are classified through ML algorithms [179]. Therefore, ML is inevitable.

Classification of areas into green, yellow, and red is based on certain features: number of infected patients, rate of infectivity, rate of recovery, death rate, and number of deaths. These features are extracted from the given statistical data of a given area and based upon these, the areas are classified. A comparison of KNN and SVM-based classification techniques is done on textual data to find the type of disease in [180]. First, the parameters are evaluated based on training data. The model is built using KNN and SVM, and the performance is evaluated using test data. Then,  $N$  cross-validations accuracy, model build time, search time, and memory used are compared between the two. Different sample sizes were chosen as training data, and results were compared. The results showed that KNN shows decreased accuracy when the size of the data increased, whereas SVM works better with a large amount of data. In all the varying sample set data, KNN consumed more build time and search time than SVM. The difference was significantly large as the sample size increased (from 50 to 1000). The memory used for both the algorithms was almost constant. Thus, if the sample size is small, KNN is preferable, and if the sample size is large, SVM is preferable. In our case, if we have to classify zones on a country basis, we would go for KNN else if it is based on states or cities, we go for SVM. However, the above algorithm comparison was done for textual data for medical purposes. The actual implementation of the classification of statistical data on COVID-19 has not been done yet.

*Summary:* This section reviewed the applications of ML models in pandemic management. First, it presented different ML models for early prediction and diagnosis of a pandemic which can be helpful in the prevention of an outbreak. Social distancing becomes essential if the spread of virus transfers needs to be controlled. This section also presented how ML algorithms integrated with various other technologies, like IoT, can be helpful in practising social distancing. KNN and SVM can be used for appropriate classification purposes, such as to classify the areas into red, orange, and green zones. Also, ML algorithms, such as Naive Bayes can be used to determine the extent to which the economy is affected. Table IV summarizes all the ML works for pandemic management.

## VII. CHALLENGES, OPEN ISSUES, AND FUTURE RESEARCH DIRECTIONS

The previous sections have discussed different ML algorithms and how they can be used in various phases during a disaster or a pandemic. However, some challenges, open issues, and research directions also exist, which are discussed as follows.

TABLE IV  
SUMMARY OF WORKS FOR PANDEMIC MANAGEMENT (ACRONYMS USED IN TABLE-COCO: COMMON OBJECTS IN CONTEXT, RSS)

Category	Reference	Target issue	Technology used	Hardware/API used	ML involved	Case studies
ML	[17], 2020	Detecting ISPA	SVM, KNN, Neural Networks and Naive Bayes	None	Yes	Indonesia
	[18], 2018	Predicting antigenic variants of H1N1 influenza virus	SVM, KNN, Neural Networks and Naive Bayes	Stacking model	Yes	WHO
	[47], 2013	Predicting spread of cholera disease	K-means	None	Yes	Haiti
Android-based	[131], 2013	To detect distance and signal the user	Bluetooth, RSS	Wireless positioning and GPS	No	Not mentioned
Radio-waves	[132], 2013	To detect distance and signal the user	Ultrasound	A Transmitter and A detector	No	A hospital
ML in android based app	[28], 2019	Movement of infected people	KNN and Decision Tree	Service Provider	Yes	Not mentioned
Sensor based	[133], 2017	Count people in an area	Maximum likelihood equation	Radar sensor	No	Novelda, Norway
Radar-based	[134], 2017	Count people in an area	Posteriori algorithm	None	No	Not mentioned
ML and Radar Sensors	[135], 2020	Detect distance between people in a crowded area	ANN and Kalman filters	Radar sensor and a camera	Yes	Not mentioned
Sensor based	[136], 2012	Detect human movements indoor	Improved heading estimation model	Smartphone sensors	No	iPhone 4s
ML and IoT	[35], 2014	Detect user's location	ANN and Radial functions	FT-6200 kit, Zig bee	Yes	Third floor, Chung Hua university
ML and UAV	[137], 2020	Predict traffic level	Recurrent Neural networks and Convolutional networks	UAV	Yes	Not mentioned
ML	[22], 2016	Detect user's location	Deep Belief Networks(DBN) and Naive Bayes'	None	Yes	Gowalla network and Brightkite network
	[138], 2019	Verify user's location	ANN and SVM	Wireless networks	Yes	Momentum project, Germany
	[32], 2020	Predict number of infected people	Random Forest and Gradient Boosting	Microphone and Thermal Cameras	Yes	Microsoft COCO dataset
	[139], 2020		DNN and Gaussian process	None	Yes	Indoor shopping mall
ML and 5G cellular networks	[39], 2018	Predict number of people in an area	RNN and DNN	Simpy simulator	Yes	Big data challenge, Italia
	[140], 2019		LSTM Neural Network	None	Yes	Geolife project
	[29], 2018		Bayesian regressor, Random Forest regressor and Gaussian regressor	None	Yes	San Francisco
	[38], 2020		DNN and CNN	GPRS	Yes	A mobile network data, Youtube, Snapchat and Facebook

### A. Challenges

1) *Requirement of Specific Data*: ML algorithms require data to be cleaned and refined for further processing. They sometimes also require the extraction of certain features from data, which is a costly process when the data is large. For example, when CNN is used to isolate the areas into catastrophe zones, it requires specific forecast features [40]. Similarly, the model in [120] requires certain other

features in its training data set to provide a better evacuation system.

2) *Accuracy Measurement*: There exist several parameters for a model's accuracy, such as precision, recall, AUC, F-measure, and many more. To accurately measure a model, a nonbiased parameter must be chosen according to the scenario. Ye *et al.* [146] have only used the AUC measure, which is

susceptible to bias due to the inclination of the test data with one class.

Furthermore, certain parameters for many algorithms need to be set to increase accuracy. For example, the number of trees needs to be specified for better accuracy in random forest models [34]. Similarly, in the fuzzy model, parameter  $m$  needs to be set appropriately.

3) *User's Privacy*: The privacy of users must not be violated by any model [181], [182]. This is a factor which the developer must keep in mind before developing any model. There is a need for mechanisms to protect the user's privacy when their data is used by ML algorithms [48]. Some of the ML algorithms we discussed above also require the use of technologies like IoT, which can breach the privacy of its users, and hence, some mitigation strategies are required [183]. Moreover, Hassija *et al.* [184] presented a survey discussing the security threats and solution architectures in a supply chain using ML and other technologies. With an increase in the use of smart healthcare devices, user's health can be at risk because of the system's vulnerabilities [185]. Mobile-health system architecture also poses threat to user's data, and an encryption algorithm is needed to protect such vulnerabilities [186].

## B. Open Issues

1) *Neural Networks Are Not Deterministic*: NNs are stochastic in nature and not deterministic. A NN does not understand that smoke density cannot be negative, or that the amount of rainfall in an area cannot be negative. Some constraints to evaluate the inputs need to be added to these algorithms.

2) *Requirement of Large Amount of Data*: Kim *et al.* [151] have a limited amount of data to detect COVID-19. The CNN model requires a large number of samples for training, which are unavailable. Due to a lack of availability of large training data, it is often not possible to use the CNN model [47], [129]. Shibata and Yamamoto [44] need an ML algorithm that requires less amount of data.

3) *Inaccurate Data*: The data fed to AI systems can be inaccurate in some scenarios. Thus, the results produced will also be inaccurate. There is no mechanism to crosscheck such large amounts of data. Also, since the results are based on previous trend data, sometimes accurate predictions cannot be made. For instance, in case of an earthquake, consider a particular area that always experiences mild earthquakes. This area experiences a strong earthquake in a certain year. Based on the previous trends, ML algorithms will predict the results inaccurately.

## C. Future Research Directions

To predict a pandemic, ML algorithms, apart from the symptoms of the disease, need to encompass certain other features, such as climate and human immune systems. The model mentioned in [19] lacks both features. However, the rainfall pattern as a feature is used to predict the spread of cholera disease, but the human immune system is still not considered in this model [54]. Also, in OOA integrated with random forests, the segmentation process needs to focus on more details so that random forest's model accuracy can be increased. Currently,

the model is only focused on spectral details. Other features may be included to analyze the areas [37].

- 1) The preprocessing time for many models must be reduced to save time for their deployment in real scenarios. For example, in the model in [124], the preprocessing time comprises 60%–80% of the total time.
- 2) Models need to be adaptable to real-world changes in order to be successful in real disaster scenarios. There is no such mechanism for this when the crowd density of the area is increased [55].
- 3) Accuracy can be increased if high-resolution images are used in ML models [51]. Low-clarity and cloudy images can lead to the under-fitting of the model.
- 4) Some models which have proven to work well with virtual laboratory data sets need to be tested on real-time data, for example, the model presented in [160]. Also, we expect that many existing models that use different technologies can be combined with ML algorithms to increase their accuracy. Some examples of such models are [23] and [43].

## VIII. CONCLUSION

In the past few years, natural disasters and pandemics have become much more damaging and frequent. This surge in the number of disasters and pandemics has caused a strain on the emergency services, and this is where ML algorithms are required to work efficiently and make the best use of existing resources. This article presented a comprehensive survey of the applications of ML in disaster and pandemic management. This article first presented a detailed explanation of ML algorithms. Then, we discussed various phases of disasters and pandemics where ML algorithms can be used. The phases are predicting a disaster, detecting a pandemic, delivering early signals, determining crowd evacuation routes, minimizing future disaster risk, social distancing, and other miscellaneous issues. We have also discussed key issues in several technologies and how ML algorithms can overcome these issues. Finally, we sketched various challenges, open issues, and future research directions.

## REFERENCES

- [1] Gianluca Rampolla del Tindaro. (2018). *Resident/Humanitarian Coordinator Report on the Use of CERF Funds, Papua New Guinea Rapid Response Earthquake*. Accessed: Sep. 12, 2020. [Online]. Available: [https://cerf.un.org/sites/default/files/resources/18-RR-PNG-29464-NR01\\_Papua%20New%20Guinea\\_RCHC.Report.pdf](https://cerf.un.org/sites/default/files/resources/18-RR-PNG-29464-NR01_Papua%20New%20Guinea_RCHC.Report.pdf)
- [2] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for COVID-19 (SARS-COV-2) pandemic: A review," *Chaos Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110059.
- [3] Y. Jahir, M. Atiquzzaman, H. H. Refai, A. Paranjothi, and P. G. LoPresti, "Routing protocols and architecture for disaster area network: A survey," *Ad Hoc Netw.*, vol. 82, pp. 1–14, Jan. 2019. [Online]. Available: <https://doi.org/10.1016/j.adhoc.2018.08.005>
- [4] X. Du, Y. Xiao, M. Guizani, and H.-H. Chen, "An effective key management scheme for heterogeneous sensor networks," *Ad Hoc Netw.*, vol. 5, no. 1, pp. 24–34, 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1570870506000412>
- [5] A. Mehra, M. Mandal, P. Narang, and V. Chamola, "ReviewNet: A fast and resource optimized network for enabling safe autonomous driving in hazy weather conditions," *IEEE Trans. Intell. Transp. Syst.*, early access, Aug. 14, 2020, doi: [10.1109/TITS.2020.3013099](https://doi.org/10.1109/TITS.2020.3013099).

- [6] T. Yaqoob, H. Abbas, and M. Atiquzzaman, "Security vulnerabilities, attacks, countermeasures, and regulations of networked medical devices—A review," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3723–3768, 4th Quart., 2019. [Online]. Available: <https://doi.org/10.1109/COMST.2019.2914094>
- [7] L. He, Z. Yan, and M. Atiquzzaman, "LTE/LTE-A network security data collection and analysis for security measurement: A survey," *IEEE Access*, vol. 6, pp. 4220–4242, 2018. [Online]. Available: <https://doi.org/10.1109/ACCESS.2018.2792534>
- [8] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2923–2960, 4th Quart., 2018.
- [9] G. Bansal, V. Chamola, P. Narang, S. Kumar, and S. Raman, "Deep3DScan: Deep residual network and morphological descriptor based framework for lung cancer classification and 3D segmentation," *IET Image Process.*, vol. 14, no. 7, pp. 1240–1247, 2020.
- [10] Wikipedia Contributors. (2020). *Machine Learning—Wikipedia the Free Encyclopedia*. Accessed: Apr. 30, 2020. [Online]. Available: [https://en.wikipedia.org/w/index.php?title=Machine\\_learning&oldid=954000510](https://en.wikipedia.org/w/index.php?title=Machine_learning&oldid=954000510)
- [11] R. Xie, I. Khalil, S. Badsha, and M. Atiquzzaman, "Collaborative extreme learning machine with a confidence interval for P2P learning in healthcare," *Comput. Netw.*, vol. 149, pp. 127–143, Feb. 2019. [Online]. Available: <https://doi.org/10.1016/j.comnet.2018.11.002>
- [12] M. Sughasiny and J. Rajeshwari, "Application of machine learning techniques, big data analytics in health care sector—A literature survey," in *Proc. 2nd Int. Conf. I-SMAC (IoT Social Mobile Anal. Cloud) (I-SMAC)I-SMAC*, 2018, pp. 741–749.
- [13] J. Shirahatti, R. Patil, and P. Akulwar, "A survey paper on plant disease identification using machine learning approach," in *Proc. 3rd Int. Conf. Commun. Electron. Syst. (ICCES)*, 2018, pp. 1171–1174.
- [14] T. T. Nguyen, "Artificial intelligence in the battle against coronavirus (COVID-19): A survey and future research directions," 2020. [Online]. Available: <https://arxiv.org/abs/2008.07343>
- [15] C. Chakraborty, A. Banerjee, L. Garg, and J. J. P. C. Rodrigues, *Internet of Medical Things for Smart Healthcare: COVID-19 Pandemic (Series Studies in Big Data)*. Singapore: Springer, 2020.
- [16] A. Soofi and A. Awan, "Classification techniques in machine learning: Applications and issues," *J. Basic Appl. Sci.*, vol. 13, pp. 459–465, Aug. 2017.
- [17] M. Kantardzic. (2011). *Data Mining: Concepts, Models, Methods, and Algorithms*. [Online]. Available: [https://www.academia.edu/18807133/Data\\_Mining\\_Concepts\\_Models\\_Methods\\_and\\_Algorithms](https://www.academia.edu/18807133/Data_Mining_Concepts_Models_Methods_and_Algorithms)
- [18] N. L. W. S. R. Ginantra, I. G. A. D. Indradewi, and E. Hartono, "Machine learning approach for acute respiratory infections (ISPA) prediction: Case study indonesia," *J. Phys. Conf. Series*, vol. 1469, Feb. 2020, Art. no. 012044. [Online]. Available: <https://doi.org/10.1088>
- [19] R. Yin, V. H. Tran, X. Zhou, J. Zheng, and C. K. Kwok, "Predicting antigenic variants of H1N1 influenza virus based on epidemics and pandemics using a stacking model," *PLoS ONE*, vol. 13, no. 12, 2018, Art. no. e0207777.
- [20] S.-B. Cho, "Exploiting machine learning techniques for location recognition and prediction with smartphone logs," *Neurocomputing*, vol. 176, pp. 98–106, May 2015.
- [21] A. Namoun *et al.*, "A multi-agent architecture for evacuating pilgrims in panic and emergency situations: The HAJJ scenario," *J. Theor. Appl. Inf. Technol.*, vol. 96, p. 20, Oct. 2018.
- [22] J. Ren and X. G. Gao, "Situation assessment model for UAV disaster RELIEF in the city," in *Proc. Int. Workshop Multi Platform Multi Sensor Remote Sens. Map.*, 2011, pp. 1–6.
- [23] A. Brighente, F. Formaggio, G. M. Di Nunzio, and S. Tomasin, "Machine learning for in-region location verification in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2490–2502, Nov. 2019.
- [24] J. Akshya and P. L. K. Priyadarsini, "A hybrid machine learning approach for classifying aerial images of flood-hit areas," in *Proc. Int. Conf. Comput. Intell. Data Sci. (ICCCIDS)*, 2019, pp. 1–5.
- [25] K. Mori, T. Wada, and K. Ohtsuki, "A new disaster recognition algorithm based on SVM for ehrs: Buffering and bagging-SVM," in *Proc. 45th Int. Conf. Parallel Process. Workshops (ICPPW)*, 2016, pp. 22–30.
- [26] D.-N. Le, V. S. Parvathy, D. Gupta, A. Khanna, J. J. P. C. Rodrigues, and K. Shankar, "IoT enabled depthwise separable convolution neural network with deep support vector machine for COVID-19 diagnosis and classification," *Int. J. Mach. Learn. Cybern.*, to be published.
- [27] Y. Jia, Y. Wang, X. Jin, and X. Cheng, "Location prediction: A temporal-spatial Bayesian model," *ACM Trans. Intell. Syst. Technol.*, vol. 7, no. 3, p. 31, Feb. 2016. [Online]. Available: <https://doi.org/10.1145/2816824>
- [28] S. Sadhukhan, S. Banerjee, P. Das, and A. K. Sangaiah, "Producing better disaster management plan in post-disaster situation using social media mining," in *Computational Intelligence for Multimedia Big Data on the Cloud With Engineering Applications*. San Diego, CA, USA: Elsevier, 2018, pp. 171–183.
- [29] N. Assery, Y. Xiaohong, S. Almalki, R. Kaushik, and Q. Xiuli, "Comparing learning-based methods for identifying disaster-related tweets," in *Proc. 18th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, 2019, pp. 1829–1836.
- [30] P. A. Gutiérrez, M. Pérez-Ortiz, J. Sánchez-Monedero, F. Fernández-Navarro, and C. Hervás-Martínez, "Ordinal regression methods: Survey and experimental study," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 1, pp. 127–146, Jan. 2016.
- [31] V.-M. Le, "Machine learning methods for optimization in multi-agent decision support system: Application to sign placement for tsunami evacuation," Ph.D. dissertation, Laboratoire d'Informatique de Paris-Nord, Paris, France, Dec. 2016.
- [32] H. A. Shaiba, N. S. Alaashoub, and A. A. Alzahrani, "Applying machine learning methods for predicting sand storms," in *Proc. 1st Int. Conf. Comput. Appl. Inf. Security (ICCAIS)*, 2018, pp. 1–5.
- [33] M. Polese, R. Jana, V. Kounev, K. Zhang, S. Deb, and M. Zorzi, "Machine learning at the edge: A data-driven architecture with applications to 5G cellular networks," *IEEE Trans. Mobile Comput.*, early access, Jun. 3, 2020, doi: [10.1109/TMC.2020.2999852](https://doi.org/10.1109/TMC.2020.2999852).
- [34] A. R. Joshi, I. Tarte, S. Suresh, and S. G. Koolagudi, "Damage identification and assessment using image processing on post-disaster satellite imagery," in *Proc. IEEE Global Humanitarian Technol. Conf. (GHTC)*, 2017, pp. 1–7.
- [35] A. Cooner, Y. Shao, and J. Campbell, "Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 haiti earthquake," *Remote Sens.*, vol. 8, p. 868, Oct. 2016.
- [36] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogrammetry Remote Sens.*, vol. 67, pp. 93–104, Jan. 2012.
- [37] A. Stumpf and N. Kerle, "Object-oriented mapping of landslides using random forests," *Remote Sens. Environ.*, vol. 115, no. 10, pp. 2564–2577, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0034425711001969>
- [38] F. Al Hossain, A. Lover, G. Corey, N. Reich, and T. Rahman, "Flusense: A contactless syndromic surveillance platform for influenza-like illness in hospital waiting areas," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 4, pp. 1–28, Mar. 2020.
- [39] T. Gupta, V. Nunavath, and S. Roy, "CrowdVAS-Net: A deep-CNN based framework to detect abnormal crowd-motion behavior in videos for predicting crowd disaster," in *Proc. IEEE Int. Conf. Syst. Man Cybern. (SMC)*, 2019, pp. 2877–2882.
- [40] S. N. K. B. Amit and Y. Aoki, "Disaster detection from aerial imagery with convolutional neural network," in *Proc. Int. Electron. Symp. Knowl. Creation Intell. Comput. (IES-KCIC)*, 2017, pp. 239–245.
- [41] T. Yabe and S. Ukkusuri, "Integrating information from heterogeneous networks on social media to predict post-disaster returning behavior," *J. Comput. Sci.*, vol. 32, pp. 12–20, Mar. 2019.
- [42] A. Buranasing and A. Prayote, "Storm intensity estimation using symbolic aggregate approximation and artificial neural network," in *Proc. Int. Comput. Sci. Eng. Conf. (ICSEC)*, 2014, pp. 234–237.
- [43] L. Luoh, "ZigBee-based intelligent indoor positioning system soft computing," *Soft Comput.*, vol. 18, no. 3, pp. 443–456, Mar. 2014. [Online]. Available: <https://doi.org/10.1007/s00500-013-1067-x>
- [44] K. Shibata and H. Yamamoto, "People crowd density estimation system using deep learning for radio wave sensing of cellular communication," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIC)*, 2019, pp. 143–148.
- [45] P. Chhikara, R. Tekchandani, N. Kumar, V. Chamola, and M. Guizani, "DCNN-GA: A deep neural net architecture for navigation of UAV in indoor environment," *IEEE Internet Things J.*, early access, Sep. 28, 2020, doi: [10.1109/JIOT.2020.3027095](https://doi.org/10.1109/JIOT.2020.3027095).
- [46] D. Bega, M. Gramaglia, M. Fiore, A. Banchs, and X. Costa-Perez, "DeepCog: Cognitive network management in sliced 5G networks with deep learning," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, 2019, pp. 280–288.

- [47] I. Alawe, A. Ksentini, Y. Hadjadj-Aoul, and P. Bertin, "Improving traffic forecasting for 5G core network scalability: A machine learning approach," *IEEE Netw.*, vol. 32, no. 6, pp. 42–49, Nov./Dec. 2018.
- [48] F. Alidoost and H. Arefi, "Application of deep learning for emergency response and disaster management," in *Proc. AGSE 8th Int. Summer School Conf.*, Feb. 2017, pp. 11–17.
- [49] S. More *et al.*, "Security assured cnn-based model for reconstruction of medical images on the Internet of healthcare things," *IEEE Access*, vol. 8, pp. 126333–126346, 2020.
- [50] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Trans. Neural Netw.*, vol. 16, no. 3, pp. 645–678, May 2005.
- [51] J. Sublime and E. Kalinicheva, "Automatic post-disaster damage mapping using deep-learning techniques for change detection: Case study of the tohoku tsunami," *Remote Sens.*, vol. 11, p. 1123, May 2019.
- [52] D. Amores, M. Vasardani, and E. Tanin, "Early detection of herding behaviour during emergency evacuations," in *Proc. 10th Int. Conf. Geograph. Inf. Sci. (GIScience)*, vol. 114, 2018, pp. 1–15.
- [53] G. Wei and Z. Sheng, "Image quality assessment for intelligent emergency application based on deep neural network," *J. Vis. Commun. Image Represent.*, vol. 63, Aug. 2019, Art. no. 102581. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1047320319301968>
- [54] L. Righetto *et al.*, "Rainfall mediations in the spreading of epidemic cholera," *Adv. Water Resources*, vol. 60, pp. 34–46, Oct. 2013.
- [55] Q. Wang, H. Liu, K. Gao, and L. Zhang, "Improved multi-agent reinforcement learning for path planning-based crowd simulation," *IEEE Access*, vol. 7, pp. 73841–73855, 2019.
- [56] K. K. Singh, M. J. Nigam, K. Pal, and A. Mehrotra, "A fuzzy kohonen local information c-means clustering for remote sensing imagery," *IETE Tech. Rev.*, vol. 31, pp. 75–81, Jan. 2014.
- [57] A. Jindal, A. Dua, N. Kumar, A. K. Das, A. V. Vasilakos, and J. J. P. C. Rodrigues, "Providing healthcare-as-a-service using fuzzy rule based big data analytics in cloud computing," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 5, pp. 1605–1618, Sep. 2018.
- [58] M. Buchholz. (Mar. 2019). *Deep Reinforcement Learning. Introduction. Deep Q Network (DQN) Algorithm.* [Online]. Available: <https://medium.com/@markus.x.buchholz/deep-reinforcement-learning-introduction-deep-q-network-dqn-algorithm-fb74bf4d6862>
- [59] K. Tian and S. Jiang, "Reinforcement learning for safe evacuation time of fire in Hong Kong–Zhuhai–Macau immersed tube tunnel," *Syst. Sci. Control Eng.*, vol. 6, pp. 45–56, Jul. 2018.
- [60] N. Chen, T. Qiu, X. Zhou, K. Li, and M. Atiquzzaman, "An intelligent robust networking mechanism for the Internet of Things," *IEEE Commun. Mag.*, vol. 57, no. 11, pp. 91–95, Nov. 2019. [Online]. Available: <https://doi.org/10.1109/MCOM.001.1900094>
- [61] H. N. Saha *et al.*, "Disaster management using Internet of Things," in *Proc. 8th Annu. Ind. Autom. Electromech. Eng. Conf. (IEMECON)*, 2017, pp. 81–85.
- [62] M. Anbarasan *et al.*, "Detection of flood disaster system based on IoT, big data and convolutional deep neural network," *Comput. Commun.*, vol. 150, pp. 150–157, Jan. 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0140366419310357>
- [63] Y. S. Lonkar, A. S. Bhagat, and S. A. S. Manjur, "Smart disaster management and prevention using reinforcement learning in IoT environment," in *Proc. 3rd Int. Conf. Trends Electron. Informat. (ICOEI)*, 2019, pp. 35–38.
- [64] B. MS and M. R. *Design of Disaster Management Based on Artificial Neural Network and Logistic Regression.* New York, NY, USA: Assoc. Comput. Mach., 2019.
- [65] I. J. of Science and S. A. B. Research (IJSR). *Review Paper on IoT Based Flood Prediction Model.* Accessed: Sep. 2, 2020. [Online]. Available: [https://www.ijrsr.net/get\\_abstract.php?paper\\_id=ART20174866](https://www.ijrsr.net/get_abstract.php?paper_id=ART20174866)
- [66] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal, and B. Sikdar, "A survey on IoT security: Application areas, security threats, and solution architectures," *IEEE Access*, vol. 7, pp. 82721–82743, 2019.
- [67] P. A. Newman, "Uninhabited aerial vehicles: Current and future use," in *Observing Systems for Atmospheric Composition.* New York, NY, USA: Springer, Jan. 2007, pp. 106–118.
- [68] L. Koh and S. Wich, "Dawn of drone ecology: Low-cost autonomous aerial vehicles for conservation," *Tropical Conservation Sci.*, vol. 5, pp. 121–132, Jun. 2012.
- [69] V. Hassija, V. Chamola, S. Garg, D. N. G. Krishna, G. Kaddoum, and D. N. K. Jayakody, "A blockchain-based framework for lightweight data sharing and energy trading in V2G network," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 5799–5812, Jun. 2020.
- [70] V. Hassija, V. Chamola, D. N. G. Krishna, and M. Guizani, "A distributed framework for energy trading between UAVs and charging stations for critical applications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5391–5402, May 2020.
- [71] J. Yim, J. Bang, S. Kim, S. Jeong, and Y. Lee, "UAV planning to optimize efficiency of image stitching in disaster monitoring using smart-eye platform," in *Proc. Int. Conf. Platform Technol. Service (PlatCon)*, 2017, pp. 1–4.
- [72] S. Frigerio *et al.*, "A Web-based platform for automatic and continuous landslide monitoring: The rotolon (Eastern Italian Alps) case study," *Comput. Geosci.*, vol. 63, pp. 96–105, Feb. 2014.
- [73] J. Ueyama *et al.*, "Exploiting the use of unmanned aerial vehicles to provide resilience in wireless sensor networks," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 81–87, Dec. 2014.
- [74] O. S. Oubbati, M. Atiquzzaman, T. A. A. Hanger, and A. Ibrahim, "Softwarization of UAV networks: A survey of applications and future trends," *IEEE Access*, vol. 8, pp. 98073–98125, 2020. [Online]. Available: <https://doi.org/10.1109/ACCESS.2020.2994494>
- [75] G. Han, J. Jiang, C. Zhang, T. Q. Duong, M. Guizani, and G. K. Karagiannidis, "A survey on mobile anchor node assisted localization in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 2220–2243, 3rd Quart., 2016.
- [76] N. S. T. Alladi, V. Chamola, and M. Guizani, "Applications of blockchain in unmanned aerial vehicles: A review," *Veh. Commun.*, vol. 23, Jun. 2020, Art. no. 100249.
- [77] Y. Wang, J. D. Colby, and K. A. Mulcahy, "An efficient method for mapping flood extent in a coastal floodplain using landsat TM and DEM data," *Int. J. Remote Sens.*, vol. 23, no. 18, pp. 3681–3696, 2002. [Online]. Available: <https://doi.org/10.1080/01431160110114484>
- [78] J. S. Kumar, S. K. Pandey, M. A. Zaveri, and M. Choksi, "Geo-fencing technique in unmanned aerial vehicles for post disaster management in the Internet of Things," in *Proc. 2nd Int. Conf. Adv. Comput. Commun. Paradigms (ICACCP)*, 2019, pp. 1–6.
- [79] S. Dey *et al.*, "Ad-hoc networked UAVs as aerial mesh network for disaster management application and remote sensing: An approach," in *Proc. 8th IEEE Annu. Inf. Technol. Electron. Mobile Commun. Conf. (IEMCON)*, 2017, pp. 301–304.
- [80] M. Terzi, A. Anastasiou, P. Kolios, C. Panayiotou, and T. Theocharides, "SWIFTERS: A multi-UAV platform for disaster management," in *Proc. Int. Conf. Inf. Commun. Technol. Disaster Manag. (ICT-DM)*, 2019, pp. 1–7.
- [81] Z. Zhang, J. Wu, and C. He, "Search method of disaster inspection coordinated by multi-UAV," in *Proc. Chin. Control Conf. (CCC)*, 2019, pp. 2144–2148.
- [82] Z. Wang, H. Li, and L. Wu, "Geodesics-based topographical feature extraction from airborne lidar data for disaster management," in *Proc. 18th Int. Conf. Geoinformat.*, 2010, pp. 1–5.
- [83] G. S. Percivall, N. S. Alameh, H. Caumont, K. L. Moe, and J. D. Evans, "Improving disaster management using earth observations—GEOSS and CEOS activities," *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, vol. 6, no. 3, pp. 1368–1375, Jun. 2013.
- [84] R. Hoque, D. Nakayama, H. Matsuyama, and J. Matsumoto, "Flood monitoring, mapping and assessing capabilities using radarsat remote sensing, GIS and ground data for bangladesh," *Nat. Hazards*, vol. 57, pp. 525–548, May 2011.
- [85] J. Tang, R. Yang, and M. Kafatos, "7.5 data mining for tropical cyclone intensity prediction," in *Proc. 6th Conf. Coastal Atmosph. Ocean. Prediction Process.*, Jan. 2005.
- [86] M. F. Sikder, S. Halder, T. Hasan, M. J. Uddin, and M. K. Baowaly, "Smart disaster notification system," in *Proc. 4th Int. Conf. Adv. Elect. Eng. (ICAEE)*, 2017, pp. 658–663.
- [87] L. Ma, M. Li, X. Ma, L. Cheng, P. Du, and Y. Liu, "A review of supervised object-based land-cover image classification," *ISPRS J. Photogrammetry Remote Sens.*, vol. 130, pp. 277–293, Aug. 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S092427161630661X>
- [88] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS J. Photogrammetry Remote Sens.*, vol. 65, no. 1, pp. 2–16, 2010. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924271609000884>
- [89] S. W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng, "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 1145–1161, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0034425711000034>

- [90] M. Demaria, C. Sampson, J. Knaff, and K. Musgrave, "Is tropical cyclone intensity guidance improving?" *Bull. Amer. Meteorol. Soc.*, vol. 95, pp. 387–398, Mar. 2014.
- [91] F. Ackermann, "Airborne laser scanning present status and future expectations," *ISPRS J. Photogrammetry Remote Sens.*, to be published, doi: [10.1016/S0924-2716\(99\)00009-X](https://doi.org/10.1016/S0924-2716(99)00009-X).
- [92] Wikipedia Contributors. (2019). *RM-ODP—Wikipedia, the Free Encyclopedia*. Accessed: Apr. 28, 2020. [Online]. Available: <https://en.wikipedia.org/w/index.php?title=RM-ODP&oldid=901830988>
- [93] OGC. (2005). *GEOSS AIP Architecture—GEOSS Architecture Implementation Pilot (AIP)*. Accessed: Feb. 28, 2012. [Online]. Available: <https://www.ogc.org/projects/initiatives/geoss-ogc>
- [94] IPCC *et al.* (2012). *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*. [Online]. Available: <https://www.ipcc.ch/report/managing-the-risks-of-extreme-events-and-disasters-to-advance-climate-change-adaptation/>
- [95] C. Polanco, J. Castañón-González, A. M. Hernondez, J. L. Samaniego, T. Buhse, and S. Martínez, "Detection of severe respiratory disease epidemic outbreaks by cusum-based overcrowd-severe-respiratory-disease-index model," *Comput. Math. Methods Med.*, vol. 2013, Aug. 2013, Art. no. 213206.
- [96] *BMC Public Health*. Accessed: Sep. 3, 2020. [Online]. Available: <https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-9-483>
- [97] N. Khan, N. Ahmad, N. Riaz, and M. Hussain, "Ad hoc wireless sensor network architecture for disaster survivor detection," *Int. J. Adv. Sci. Technol.*, vol. 34, pp. 9–16, Sep. 2011. [Online]. Available: <http://www.sersc.org/journals/ijast/vol34/2.pdf>
- [98] X. Du, M. Guizani, Y. Xiao, and H. Chen, "Transactions papers a routing-driven elliptic curve cryptography based key management scheme for heterogeneous sensor networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1223–1229, Mar. 2009.
- [99] C. Yawut and S. Kilaso, "A wireless sensor network for weather and disaster alarm systems," in *Proc. Int. Conf. Inf. Electron. Eng. (IPCSIT)*, vol. 6, 2011, pp. 155–159.
- [100] T. Nakamura *et al.*, "Development of emergency rescue evacuation support system (ERESS) in panic-type disasters: Disaster detection by positioning area of terminals," in *Proc. 42nd Int. Conf. Parallel Process.*, 2013, pp. 931–936.
- [101] W. PITTS and E. Johnsson, "Carbon monoxide formation in fires by high-temperature anaerobic wood pyrolysis," in *Proc. Symp. Int. Combustion*, vol. 25, Dec. 1994, pp. 1455–1462.
- [102] X. Xu, L. Zhang, S. Sotiriadis, E. Asimakopoulou, M. Li, and N. Bessis, "CLOTHO: A large-scale Internet of Things-based crowd evacuation planning system for disaster management," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3559–3568, Oct. 2018.
- [103] M. Choksi, H. Gamit, and M. A. Zaveri, "Data sensing and resource scheduling in clustered environment for post disaster management using Internet of Things," in *Proc. 3rd Int. Conf. Conver. Technol. (I2CT)*, 2018, pp. 1–5.
- [104] L. Abusalah, A. Khokhar, and M. Guizani, "A survey of secure mobile ad hoc routing protocols," *IEEE Commun. Surveys Tuts.*, vol. 10, no. 4, pp. 78–93, 4th Quart., 2008.
- [105] H. Rossnagel and T. Scherner, "Secure mobile notifications of civilians in case of a disaster," in *Proc. 10th IFIP TC-6 TC-11 Int. Conf. Commun. Multimedia Security (CMS)*, 2006, pp. 33–42. [Online]. Available: [https://doi.org/10.1007/11909033\\_4](https://doi.org/10.1007/11909033_4)
- [106] G. Jayasinghe, F. Fahmy, N. Gajaweera, and D. Dias, "A GSM alarm device for disaster early warning," in *Proc. 1st Int. Conf. Ind. Inf. Syst.*, 2006, pp. 383–387.
- [107] H. Torii, J. Sawamoto, N. Segawa, E. Sugino, and Y. Nomura, "Tsunami early alert and evacuation support system for fishery workers by mobile phones," in *Proc. IEEE 24th Int. Conf. Adv. Inf. Netw. Appl. Workshops*, 2010, pp. 704–709.
- [108] B. Anzengruber, D. Pianini, J. Nieminen, and A. Ferscha, *Predicting Social Density in Mass Events to Prevent Crowd Disasters*, vol. 8238. Cham, Switzerland: Springer, Nov. 2013, pp. 206–215.
- [109] F. Sharevski, "Leveraging cellular Internet-of-Things for resilient and robust disaster management," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–5.
- [110] K. M. Rahman, T. Alam, and M. Chowdhury, "Location based early disaster warning and evacuation system on mobile phones using openstreetmap," in *Proc. IEEE Conf. Open Syst.*, 2012, pp. 1–6.
- [111] T. Juhana, "Mobile application design of disaster rescue team tracking," in *Proc. 8th Int. Conf. Telecommun. Syst. Services Appl. (TSSA)*, 2014, pp. 1–3.
- [112] M. R. M. Aldecimo and M. M. De Leon, "Development of an OpenStreetMap based safe zone routing system for west valley fault earthquake disaster in the Makati central business district, Philippines," in *Proc. 3rd Int. Conf. Inf. Commun. Technol. Disaster Manag. (ICT-DM)*, 2016, pp. 1–7.
- [113] A. Khan, F. Aftab, and Z. Zhang, "UAPM: An urgency-aware packet management for disaster management using flying ad-hoc networks," *China Commun.*, vol. 16, no. 11, pp. 167–182, 2019.
- [114] B. Wolshon, *Contraflow for Evacuation Traffic Management*. Boston, MA, USA: Springer, 2008, pp. 165–170. [Online]. Available: [https://doi.org/10.1007/978-0-387-35973-1\\_210](https://doi.org/10.1007/978-0-387-35973-1_210)
- [115] F. Jiang *et al.*, "Supercharging crowd dynamics estimation in disasters via spatio-temporal deep neural network," in *Proc. IEEE Int. Conf. Data Sci. Adv. Anal. (DSAA)*, 2017, pp. 184–192.
- [116] J. Radianti and O.-C. Granmo, "A framework for assessing the condition of crowds exposed to a fire hazard using a probabilistic model," *Int. J. Mach. Learn. Comput.*, vol. 4, no. 1, pp. 14–20, 2014.
- [117] P. Haddawy *et al.*, "Situation awareness in crowdsensing for disease surveillance in crisis situations," in *Proc. ICTD*, May 2015, pp. 1–5.
- [118] C. A. F. Ezequiel *et al.*, "UAV aerial imaging applications for post-disaster assessment, environmental management and infrastructure development," in *Proc. Int. Conf. Unmanned Aircraft Syst. (ICUAS)*, 2014, pp. 274–283.
- [119] P. V. Klaine, J. Nadas, R. Souza, and M. Imran, "Distributed drone base station positioning for emergency cellular networks using reinforcement learning," *Cogn. Comput.*, vol. 10, pp. 790–804, May 2018.
- [120] J. W. Burris, R. Shrestha, B. Gautam, and B. Bista, "Machine learning for the activation of contraflows during hurricane evacuation," in *Proc. IEEE Global Humanitarian Technol. Conf. (GHTC)*, 2015, pp. 254–258.
- [121] Y. Ikeda and M. Inoue, "An evacuation route planning for safety route guidance system after natural disaster using multi-objective genetic algorithm," *Procedia Comput. Sci.*, vol. 96, pp. 1323–1331, Dec. 2016.
- [122] F. Ofli *et al.*, "Combining human computing and machine learning to make sense of big (aerial) data for disaster response," *Big Data*, vol. 4, no. 1, pp. 47–59, Mar. 2016.
- [123] V. Hassija, V. Saxena, V. Chamola, and F. R. Yu, "A parking slot allocation framework based on virtual voting and adaptive pricing algorithm," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 5945–5957, Jun. 2020.
- [124] S. Soni, V. G. Shankar, and C. Sandeep, "Route-the safe: A robust model for safest route prediction using crime and accidental data," *Int. J. Adv. Sci. Technol.*, vol. 28, no. 16, pp. 1415–1428, Dec. 2019.
- [125] M. Y. Selim and A. E. Kamal, "Post-disaster 4G/5G network rehabilitation using drones: Solving battery and backhaul issues," *Proc. IEEE Globecom Workshops (GC Wkshps)*, Abu Dhabi, UAE, 2018, pp. 1–6, doi: [10.1109/GLOCOMW.2018.8644135](https://doi.org/10.1109/GLOCOMW.2018.8644135).
- [126] N. Gondaliya and M. Atiquzzaman, "RTTMM: Role based 3-tier mobility model for evaluation of delay tolerant routing protocols in post disaster situation," in *Proc. Int. Conf. Internet Things Big Data*, vol. 2, 2016, pp. 11–20.
- [127] D. de Alwis and E. So, "Enhanced change detection index for disaster response, recovery assessment and monitoring of accessibility and open spaces (camp sites)," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 57, pp. 49–60, May 2017.
- [128] M. Sheykhmousa, N. Kerle, M. Kuffer, and S. Ghaffarian, "Post-disaster recovery assessment with machine learning-derived land cover and land use information," *Remote Sens.*, vol. 11, p. 24, May 2019.
- [129] C. J. Van Westen, "Where to build-back-better? Analyzing changing risk for post-disaster reconstruction planning," in *Proc. 5th Int. Symp. Mega Earthquake Induced Geo Disasters Long Term Effects*, Chengdu, China, 2018, pp. 11–16.
- [130] K. Starbird and L. Palen, "(How) will the revolution be retweeted? Information diffusion and the 2011 Egyptian uprising," in *Proc. CSCW*, Dec. 2012, pp. 7–16.
- [131] S. Vieweg, A. Hughes, K. Starbird, and L. Palen, "Microblogging during two natural hazards events: What twitter may contribute to situational awareness," in *Proc. CHI*, vol. 2, Jan. 2010, pp. 1079–1088.
- [132] M. Jamali, A. Nejat, S. Ghosh, and G. Cao, "Social media data and post-disaster recovery," *Int. J. Inf. Manag.*, vol. 44, pp. 25–37, Sep. 2018.
- [133] N. Khan, "Flood prediction and disaster risk analysis using GIS based wireless sensor networks—A review," *J. Basic Appl. Sci. Res.*, vol. 11, no. 8, pp. 438–446, Sep. 2013.
- [134] N. A. Agana and A. Homaifar, "EMD-based predictive deep belief network for time series prediction: An application to drought forecasting," *Hydrology*, vol. 5, no. 1, p. 18, 2018.

- [135] Y. Y. Li. (May 2018). *Big Data in Natural Disaster Management: A Review*. [Online]. Available: <https://www.mdpi.com/2076-3263/8/5/165#cite>
- [136] *Bluetooth Technology*. Accessed: Sep. 12, 2020. [Online]. Available: <https://www.bluetooth.com/>
- [137] S. Holm, "Airborne ultrasound data communications: The core of an indoor positioning system," in *Proc. IEEE Ultrasonics Symp.*, vol. 3, 2005, pp. 1801–1804.
- [138] J. W. Choi, D. H. Yim, and S. H. Cho, "People counting based on an IR-UWB radar sensor," *IEEE Sensors J.*, vol. 17, no. 17, pp. 5717–5727, Sep. 2017.
- [139] S. Bartoletti, A. Conti, and M. Z. Win, "Device-free counting via wideband signals," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1163–1174, May 2017.
- [140] S. Johnson, J. J. Hasan, and N. Ahmed. *Social-Distancing Monitoring Using Portable Electronic Devices*. Accessed: Apr. 17, 2020. [Online]. Available: [https://www.tdcommons.org/dpubs\\_series/3158/](https://www.tdcommons.org/dpubs_series/3158/)
- [141] W. Kang, S. Nam, Y. Han, and S. Lee, "Improved heading estimation for smartphone-based indoor positioning systems," in *Proc. IEEE 23rd Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, 2012, pp. 2449–2453.
- [142] U. Challita, A. Ferdowsi, M. Chen, and W. Saad, "Machine learning for wireless connectivity and security of cellular-connected UAVs," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 28–35, Feb. 2019.
- [143] F. Yin *et al.*, "FedLoc: Federated learning framework for data-driven cooperative localization and location data processing" *IEEE Open J. Signal Process.*, 2020.
- [144] C. Wang, L. Ma, R. Li, T. S. Durrani, and H. Zhang, "Exploring trajectory prediction through machine learning methods," *IEEE Access*, vol. 7, pp. 101441–101452, 2019.
- [145] A. L. Pineda, Y. Ye, S. Visweswaran, G. Cooper, M. Wagner, and F. R. Tsui, "Comparison of machine learning classifiers for influenza detection from emergency department free-text reports," *J. Biomed. Informat.*, vol. 58, pp. 60–69, Dec. 2015.
- [146] Y. Ye, F. R. Tsui, M. Wagner, J. U. Espino, and Q. Li, "Influenza detection from emergency department reports using natural language processing and Bayesian network classifiers," *J. Amer. Med. Inf. Assoc.*, vol. 21, no. 5, pp. 815–823, 2014.
- [147] L. Wynants *et al.*, "Prediction models for diagnosis and prognosis of COVID-19: Systematic review and critical appraisal," *BMJ*, vol. 369, Jul. 2020, Art. no. m1328. [Online]. Available: <https://www.bmj.com/content/369/bmj.m1328>
- [148] A. Lopez-Rincon, A. Tonda, L. Mendoza-Maldonado, E. Claassen, J. Garssen, and A. D. Kraneveld, "Accurate identification of Sars-Cov-2 from viral genome sequences using deep learning," *bioRxiv*, 2020.
- [149] Beijing Institute of Genomics, Chinese Academy of Science, China National Center for Bioinformation & National Genomics Data Center. *2019 Novel Coronavirus Resource (2019nCoV)*. Accessed: Oct. 15, 2020. [Online]. Available: <https://bigd.big.ac.cn/ncov/?lang=en>
- [150] S. T. Sherry *et al.*, "DBSNP: The NCBI database of genetic variation," *Nucl. Acids Res.*, vol. 29, no. 1, pp. 308–311, 2001.
- [151] J. Kim *et al.*, "A first case of 2019 novel coronavirus pneumonia imported into Korea from Wuhan, China: Implication for infection prevention and control measures," *J. Korean Med. Sci.*, vol. 35, no. 5, p. e61, Feb. 2020.
- [152] Y. Chaudhary *et al.*, "Efficient-CovidNet: Deep learning based COVID-19 detection from chest x-ray images," in *Proc. IEEE Healthcom*, Dec. 2020, pp. 1–6.
- [153] P. Santos *et al.*, "Prediction of COVID-19 using time-sliding window: The case of Piaui State—Brazil," in *Proc. IEEE Healthcom*, Dec. 2020, pp. 1–6.
- [154] F. Liang, P. Guan, W. Wu, and D. Huang, "Forecasting influenza epidemics by integrating Internet search queries and traditional surveillance data with the support vector machine regression model in liaoning, from 2011 to 2015," *PeerJ*, vol. 6, Jun. 2018, Art. no. e5134.
- [155] D. Olson, K. Konty, M. Paladini, C. Viboud, and L. Simonsen, "Reassessing google flu trends data for detection of seasonal and pandemic influenza: A comparative epidemiological study at three geographic scales," *PLoS Comput. Biol.*, vol. 9, Oct. 2013, Art. no. e1003256.
- [156] *Google Flu Trends and Google Dengue Trends*. Accessed: Oct. 15, 2020. [Online]. Available: <https://www.google.org/flutrends/about/>
- [157] M. W. L. Moreira, J. J. P. C. Rodrigues, V. Korotaev, J. Al-Muhtadi, and N. Kumar, "A comprehensive review on smart decision support systems for health care," *IEEE Syst. J.*, vol. 13, no. 3, pp. 3536–3545, Sep. 2019.
- [158] C. Wang *et al.*, "Epidemiological features and forecast model analysis for the morbidity of influenza in Ningbo, China, 2006–2014," *Int. J. Environ. Res. Public Health*, vol. 14, p. 559, May 2017.
- [159] S. G. Alonso, I. de la Torre Diez, J. J. Rodrigues, S. Hamrioui, and M. Lopez-Coronado, "A systematic review of techniques and sources of big data in the healthcare sector," *J. Med. Syst.*, vol. 41, no. 11, p. 183, 2017.
- [160] G. Manogaran, R. Varatharajan, D. Lopez, P. M. Kumar, R. Sundarasekar, and C. Thota, "A new architecture of Internet of Things and big data ecosystem for secured smart healthcare monitoring and alerting," *Future Gener. Comput. Syst.*, vol. 82, pp. 375–387, Nov. 2017.
- [161] (Jan. 2020). *COVID-19 Overview*. [Online]. Available: <https://www.coronatracker.com/analytics>
- [162] V. Chamola, V. Hassija, V. Gupta, and M. Guizani, "A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact," *IEEE Access*, vol. 8, pp. 90225–90265, 2020.
- [163] A. Handel, I. M. Longini, and R. Antia, "Intervention strategies for an influenza pandemic taking into account secondary bacterial infections," *Epidemics*, vol. 1, no. 3, pp. 185–195, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1755436509000413>
- [164] C. Fraser, S. Riley, R. Anderson, and N. Ferguson, "Factors that make an infectious disease outbreak controllable," *Proc. Nat. Acad. Sci. USA*, vol. 101, pp. 6146–51, May 2004.
- [165] J. Kelso, G. Milne, and H. Kelly, "Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza," *BMC Public Health*, vol. 9, p. 117, May 2009.
- [166] Y. Wang, X. Yang, Y. Zhao, Y. Liu, and L. Cuthbert, "Bluetooth positioning using RSSI and triangulation methods," in *Proc. IEEE 10th Consum. Commun. Netw. Conf. (CCNC)*, 2013, pp. 837–842.
- [167] S. S. Hassan, S. D. Bibon, M. S. Hossain, and M. Atiqzaman, "Security threats in Bluetooth technology," *Comput. Security*, vol. 74, pp. 308–322, May 2018. [Online]. Available: <https://doi.org/10.1016/j.cose.2017.03.008>
- [168] F. Sattler *et al.*, "Risk estimation of SARS-COV-2 transmission from Bluetooth low energy measurements," *NPJ Digit. Med.*, vol. 3, p. 129, Apr. 2020.
- [169] C. T. Nguyen *et al.*, "Enabling and emerging technologies for social distancing: A comprehensive survey," *IEEE Access*, vol. 8, pp. 153479–153507, 2020.
- [170] V. Hassija, V. Gupta, S. Garg, and V. Chamola, "Traffic jam probability estimation based on blockchain and deep neural networks," *IEEE Trans. Intell. Transp. Syst.*, early access, Jun. 3, 2020, doi: [10.1109/TITS.2020.2988040](https://doi.org/10.1109/TITS.2020.2988040).
- [171] V. Hassija, V. Chamola, G. Han, J. J. P. C. Rodrigues, and M. Guizani, "DAGIoV: A framework for vehicle to vehicle communication using directed acyclic graph and game theory," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4182–4191, Apr. 2020.
- [172] I. Ahmed, M. Ahmad, J. J. P. C. Rodrigues, G. Jeon, and S. Din, "A deep learning-based social distance monitoring framework for COVID-19," *Sustain. Cities Soc.*, to be published.
- [173] G. Praveen, V. Chamola, V. Hassija, and N. Kumar, "Blockchain for 5G: A prelude to future telecommunication," *IEEE Netw.*, vol. 34, no. 6, pp. 106–113, Nov./Dec. 2020.
- [174] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: Challenges and research advances," *IEEE Netw.*, vol. 28, no. 6, pp. 6–11, Nov./Dec. 2014.
- [175] N. Bui, M. Cesana, S. A. Hosseini, Q. Liao, I. Malanchini, and J. Widmer, "A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1790–1821, 3rd Quart., 2017.
- [176] S. Cho, P. Gordon, J. Moore, H. Richardson, M. Shinozuka, and S. Chang, "Integrating transportation network and regional economic models to estimate the costs of a large urban earthquake," *J. Reg. Sci.*, vol. 41, pp. 39–65, Feb. 2001.
- [177] E. Cavallo, S. Galiani, I. Noy, and J. Pantano, "Catastrophic natural disasters and economic growth," Dept. Econ., Univ. Hawaii, Manoa, Honolulu, HI, USA, Working Papers, Jan. 2010.
- [178] F. Li, "The information content of forward-looking statements in corporate filings—A Naive Bayesian machine learning approach," *J. Account. Res.*, vol. 48, pp. 1049–1102, Dec. 2010.



- [179] T.-N. Phan and M. Kappas, "Comparison of random forest,  $k$ -nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery," *Sensors*, vol. 18, p. 18, Dec. 2017.
- [180] J. Raikwal and K. Saxena, "Performance evaluation of SVM and  $k$ -nearest neighbor algorithm over medical data set," *Int. J. Comput. Appl.*, vol. 50, pp. 35–39, Jul. 2012.
- [181] M. A. P. Chamikara, P. Bertók, I. Khalil, D. Liu, S. Camtepe, and M. Atiqzaman, "Local differential privacy for deep learning," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5827–5842, Jul. 2020.
- [182] M. A. P. Chamikara, P. Bertók, I. Khalil, D. Liu, S. Camtepe, and M. Atiqzaman, "A trustworthy privacy preserving framework for machine learning in industrial IoT systems," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 6092–6102, Sep. 2020.
- [183] T. Alladi, V. Chamola, B. Sikdar, and K.-K. R. Choo, "Consumer IoT: Security vulnerability case studies and solutions," *IEEE Consum. Electron. Mag.*, vol. 9, no. 2, pp. 17–25, Mar. 2020.
- [184] V. Hassija, V. Chamola, V. Gupta, S. Jain, and N. Guizani, "A survey on supply chain security: Application areas, security threats, and solution architectures," *IEEE Internet Things J.*, early access, Sep. 22, 2020, doi: [10.1109/JIOT.2020.3025775](https://doi.org/10.1109/JIOT.2020.3025775).
- [185] V. Hassija, V. Chamola, B. C. Bajpai, Naren, and S. Zeadally, "Security issues in implantable medical devices: Fact or fiction?" *Sustain. Cities Soc.*, Nov. 2020, Art. no. 102552.
- [186] B. M. C. Silva, J. J. P. C. Rodrigues, F. Canelo, I. M. C. Lopes, and J. Lloret, "Towards a cooperative security system for mobile-health applications," *Electron. Commer. Res.*, vol. 19, pp. 629–654, Sep. 2019.



**Sakshi Gupta** is currently pursuing the B.Tech. degree with the Jaypee Institute of Information Technology, Noida, India.

She has completed a few projects in the field of evolutionary algorithms, machine learning, and data analytics. She is currently (the summer of 2020) pursuing her research internship with BITS-Pilani, Pilani, India. under Dr. V. Chamola. Her research interests include evolutionary computing, machine learning, and deep learning.



**Adit Goyal** is currently pursuing the B.Tech. degree with the Computer Science Department, Jaypee Institute of Information Technology, Noida, India.

He has completed a few projects in the field of data science, machine learning, and big data. He is currently pursuing his research internship with BITS-Pilani, Pilani, India, under Dr. V. Chamola. His research interests include machine learning, data science, and quantum computing.



**Vinay Chamola** (Senior Member, IEEE) received the B.E. degree in electrical and electronics engineering and master's degree in communication engineering from the Birla Institute of Technology and Science at Pilani, Pilani, India, in 2010 and 2013, respectively, and the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2016.

He is currently an Assistant Professor with the Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science at Pilani, where he heads the Internet of Things Research Group/Lab. He has over 50 publications in high ranked SCI Journals including more than 33 IEEE transaction and journal articles. His research interests include IoT security, blockchain, UAVs, VANETS, 5G, and Healthcare.

He serves as the Area Editor of *Ad Hoc Networks*, Elsevier. He serves as an Associate Editor for various reputed journals like *IET Networks*, *IET Quantum Communications* etc.



**Mohsen Guizani** (Fellow, IEEE) received the B.S. (with Distinction) and M.S. degrees in electrical engineering, and the M.S. and Ph.D. degrees in computer engineering from Syracuse University, Syracuse, NY, USA, in 1984, 1986, 1987, and 1990, respectively.

He served in different academic and administrative positions with the University of Idaho, Moscow, ID, USA; Western Michigan University, Kalamazoo, MI, USA; the University of West Florida, Pensacola, FL, USA; the University of Missouri–Kansas City, Kansas City, MO, USA; the University of Colorado Boulder, Boulder, CO, USA; and Syracuse University. He is currently a Professor with the Computer Science and Engineering Department, Qatar University, Doha, Qatar. He has authored nine books and more than 700 publications in refereed journals and conferences. His research interests include wireless communications and mobile computing, computer networks, mobile cloud computing, security, and smart grid.



**Vikas Hassija** received the B.Tech. degree from Maharshi Dayanand University, Rohtak, India, in 2010, and the M.S. degree in telecommunications and software engineering from the Birla Institute of Technology and Science at Pilani, Pilani, India, in 2014. He is currently pursuing the Ph.D. degree in IoT security and blockchain with the Jaypee Institute of Information and Technology (JIIT), Noida, India.

He is currently an Assistant Professor with JIIT. His research interests include the IoT security, network security, blockchain, and distributed computing.



**Biplab Sikdar** (Senior Member, IEEE) received the B.Tech. degree in electronics and communication engineering from North Eastern Hill University, Shillong, India, in 1996, the M.Tech. degree in electrical engineering from the Indian Institute of Technology Kanpur, Kanpur, India, in 1998, and the Ph.D. degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, USA, in 2001.

He was a faculty with the Rensselaer Polytechnic Institute from 2001 to 2013, first as an Assistant Professor and then as an Associate Professor. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, National University of Singapore (NUS), Singapore. He serves as the Vice Dean of Graduate Program and as the Area director of Communications and Networks Lab with NUS. His current research interests include wireless network, and security for Internet of Things and cyber-physical systems.