IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 1 March 2023, accepted 6 March 2023, date of publication 17 March 2023, date of current version 22 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3257429

RESEARCH ARTICLE

A Data-Intelligent Scheme Toward Smart Rescue and Micro-Services

NAFEES ZAMAN[1], AHMAD ABU SAIID[2], MD ARAFATUR RAHMAN[3], (Senior Member, IEEE), SHAVAN ASKAR[4], AND JASNI MOHAMAD ZAIN[5]

[1]Faculty of Electrical Engineering and Computing, University of Zagreb, 10000 Zagreb, Croatia
[2]École Polytechnique de Bruxelles, Université Libre de Bruxelles, 1050 Brussels, Belgium
[3]School of Mathematics and Computer Science, University of Wolverhampton, WV1 1LY Wolverhampton, U.K.
[4]Erbil Technical Engineering College, Erbil Polytechnic University, Erbil 44001, Iraq
[5]Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Komplek Al-Khawarizmi, Universiti Teknologi MARA, Shah Alam, Selangor 40450, Malaysia

Corresponding authors: Nafees Zaman (mohammad.zaman@fer.hr), Md Arafatur Rahman (arafatur.rahman@ieee.org), and Jasni Mohamad Zain (jasni@fskm.uitm.edu.my)

This work was partly supported by the University of Wolverhampton, U.K.; by Institute for Big Data Analytics and Artificial Intelligence (IBDAAI); by Universiti Teknologi MARA, Malaysia; and in part by Erbil Polytechnic University, Iraq.

ABSTRACT A considerable portion of the world frequently experiences flooding during the monsoon season. As a result of this catastrophic event, hundreds of individuals have become homeless. In addition, rescuers are not usually effective enough to rescue the majority of victims. This is due to inadequate rescue operations infrastructure, a severe flaw in today's technologically advanced society. This manuscript proposes a microservice-dependent secure rescue framework that uses geographic information system mapping with a K-Means clustering algorithm to identify flood-prone regions. Numerous microservices, such as fleet management, cloud computing, and data security, integrate and execute the framework in pre- and post-flood situations. Labeling data from the proposed framework generates a support vector machine-based classifier for predicting flood risk. Furthermore, a hybrid A* algorithm is developed to find an optimal route for the rescue operation. Based on the K-means clustering results, which reduced the variance by 89.2 percent overall, dividing the data into six clusters was the best option for this study. The smoothness of the suggested hybrid algorithm is also used to verify its superiority.

INDEX TERMS Data-intelligent, micro-services, geographic information system, risk map, K-means clustering, PCA, support vector machine.

I. INTRODUCTION

Natural disasters that occur around the globe are adverse events which are caused by natural factors in our environment that can have a massive impact on human lives [1], economic losses, and the overall welfare of any given society. Some examples of natural disasters are hurricanes, landslides, tsunamis, earthquakes, floods, and tornadoes. From these various events, floods are one of the most common forms of natural disaster that is inflicted on many parts of the world. This regularly occurs due to heavy rainfall, causing rivers to get overwhelmed with water, resulting in the river banks breaking out, and water starts to flood the dry lands (floodplains) closest to the river [2].

Although located in a stable tropical region, several countries face floods as the most persistent natural disaster. This tends to recur annually due to the prolonged heavy monsoon rains that come about as a cycle during the end of the year. Apart from monsoon floods, flash floods are another type that tends to take place in areas that are undergoing heavy and rapid development. Flash floods are normally made up of high-velocity water that rises very quickly with a lot of additional debris. In contrast with monsoon floods which can take up to multiple weeks for the water flow to return back to normal, flash floods may only take less than a day to recover [3].

The associate editor coordinating the review of this manuscript and approving it for publication was Mehul S. Raval.

Monsoons and heavy rainfall, the hefty siltation of rivers are the major affecting factor of occurring floods. To add to that, deforestation, change in climate, and rapid urbanization contributes significantly as well [4]. Such natural disasters can destroy numerous resources [5]. A recent study shows annual flood damage costs could go up to USD100 million, affecting lives, properties, agriculture, and major road systems [6]. Mostly the coastal area receives the wave of loss most. To compensate for the loss of lives and properties, many organizations take steps to provide aid to the victims. Moreover, the rescue teams are now well-equipped to face extreme challenges while rescuing folks. However, there is a limited structured framework [7], [8], [9] for the rescue operations to take initiatives and prevent the loss rather than cure. Due to the unpredictability of flood regions, time and efficient rescue management, and the dearth of micro-services, it is very hard to minimize the loss and rescue victims quickly.

Recently, micro-services are gaining popularity in different domains such as Education [10] industry. Each micro-service serves sub-portions of a use case which makes the use case easier to solve. Thus, utilizing on-demand micro-services in rescue scenarios improves the efficiency, reliability, and stability of the rescue architecture. These micro-services include fleet management to data and cloud security.

In this paper, we consider one of the major states of Malaysia named Terengganu's annual flood situation as an instance for the case study. Terengganu is affected by floods every year because of its geographic location. From early November to March, every year, Terengganu experiences the northeast monsoon season, which is exposed to heavy rainfall [11]. Due to this phenomenon, Terengganu often faces floods somewhere in that period.

Sustainable cities and communities are one of the United Nations' Sustainable Development Goals - Agenda 2030. And of the component of the goal is Disaster Risk Management. It is an intrinsic aspect of social and economic growth and is essential for the future sustainability of development. This study aspires to improve the efficiency of flood rescue operations and contributes to the conceptualization of sustainable smart cities. In light of the limitations of existing frameworks, this research explores and ranks the most flood-prone locations in Terengganu using supervised and unsupervised machine learning methods such as K-means and Support Vector Machines. Our methodology generates heat maps via predictive modeling in order to identify risky areas. Hence, a hybrid shortest path method is presented to quickly rescue persons in the aftermath of a flood. Our contributions specifically include the following.

- Develop a generic micro-service dependent, and data-intelligent secured rescue two steps (pre-flood and post-flood) framework, especially for floods.
- Identifying most vulnerable areas utilizing **K-means** algorithm. Several features have been taken into account such as average annual rainfall, population, altitude, etc. in order to create clusters.
- After that, generated clusters are visualized, and are classified as the low, medium, and high calculating risk through a weighted system. A **Heat-map** and a **Risk-map** are produced from the classified clustering result as an output through **QGIS**.
- Furthermore, a support vector machine-based predictive model is developed by feeding classified data.
- Finally, a hybrid shortest path algorithm is constructed exploiting **relative coordinate conversion**, **euclidean distance**, and **traveling salesman problem**.

The whole ecosystem is regulated through stakeholders and to execute the framework properly various micro-services such as fleet management, medical services, etc. are provided. The main purpose of the flood risk map & shortest path algorithm is to aid flood-rescue organizations to identify vulnerable areas to set up rescue team bases during the monsoon season to ensure that the teams are as close as possible to the regions that are affected by flood and can reach out to the victims more conveniently.

The rest of the paper is arranged as follows. Section 2 provides relevant knowledge of components used in previous frameworks. Sections 3 and 4 discuss in detail our proposed framework along with the methodology of generating heat-map and algorithm construction. Section 5 assesses the overall performance of our proposed system and recapitulates explicit results as well. Section 6 summarizes this research work with key findings.

II. RELATED WORK

There are various research works have been done to encounter against floods. The majority of these works adhered to discuss challenges and opportunities regarding risk management. As [12] outlines risk management strategies, [13] illustrates the potential of big data in flood risk management, and [14] presents the causes and effects of the flood, a small number of studies have been undertaken on GIS mapping and smart rescue procedures. Here we review the existing literature on both GIS mapping and smart rescue.

A. GIS MAPPING

Geographic Information System (GIS) acquires, analyzes, and interprets geographical or spatial data. GIS collects geographic location data, such as latitude, longitude, elevation, and land-use data, and arranges it into layers that can be visualized and evaluated. These layers can be combined and superimposed to generate maps providing specific regions or location details. It assists stakeholders in comprehending the relationships between geographical elements, recognizing trends, and making informed choices.

In regards to flood risk mapping, Tehrany et al. [15] focuses on Geographic Information System (GIS) mapping for identifying areas that are prone to flood occurrences in different parts of Malaysia. This manuscript targets a specific part of the Terengganu State known as Kuala Terengganu city. The main techniques involved an ensemble weights-of-evidence

(Bayesian Model) and Support Vector Machine (SVM) which were applied to 10 flood conditioning factors.

Mojaddadi et al. [16] implements SVM and a frequency ratio (FR) approach for weighing each flood conditioning factor to be inputted into the SVM model to produce a flood risk map for the Damansara river catchment area.

Ouma et al. [17] constructed a public-based flood map for estimating flood risks. The model was integrated with an Analytical Hierarchy Process (AHP) and Geographic Information System (GIS) analysis technique. The flood risk vulnerability mapping followed a multi-parametric approach and integrates some of the flood-causing factors.

Kia et al. [18] developed a flood model from various flood causative factors using ANN techniques and GIS to determine flood-prone areas. The ANN was used to directly produce water levels and then the flood map was constructed in GIS.

Tehrany et al. [19] have targeted the state of Kelantan for developing spatial prediction of flood-susceptible areas using Rule-Based Decision Trees with a novel ensemble bivariate and multivariate statistical approach in GIS. The main purpose of the research was to compare the prediction performance of the Rule-Based Decision Tree against the combination of a Frequency Ratio (FR) and Logistic Regression (LR) model.

Jalayer et al. [20] identifies urban flood risk hotspots. The study includes three GIS-based frameworks for identifying the flood risk hotspots for residential buildings and urban corridors. The Maximum Likelihood (MLE) method was used for estimating the threshold used for identifying the flood risk areas.

B. SHORTEST-PATH ALGORITHM

Shortest path algorithms determine the shortest route between two locations. Given a destination, the shortest path algorithm determines the shortest and most efficient route based on several input factors. These algorithms are utilized for navigational efficiency, enhanced path planning, optimization, etc. In addition to its advantages, it has disadvantages such as computational complexity, lack of adaptability, etc. Thus, these algorithms are selected based on the nature of the problem.

As for the construction of a shortest-path algorithm, Takeda et al. [21] focuses on a path generation algorithm for search and rescue robots based on insect behavior (Ladybirds). This implementation focuses on reducing the range of a search area once a target is identified and returns to a larger range once another target is identified. The path generation algorithm is optimized with the help of the Genetic Algorithm.

Becker et al. [22] introduces a new multi-agent algorithm for the purpose of search and rescue scenarios involving unknown terrain. This method combines the concept of the flooding algorithm (for exploration) and path-optimizing features from the ant algorithm. For most of the scenarios, their proposed multi-agent flood algorithm performed a little slower compared to the Brick and Mortar algorithm.

Arif et al. [23] conducted a comparative study of different path-planning algorithms for a water-based rescue system. A total of four algorithms were tested, namely, the Graph Search algorithm, the Breadth-First algorithm, Dijkstra's algorithm, and the A-Star (A*) algorithm. From the results, it can be seen that each algorithm performs better in different situations. For maps with the same cost cells, with one starting-one goal cell and multi-goal cells, using the Breadth-first algorithm was the best if the computational time is the main priority, but if the size of memory was the priority, then A* would be a better alternative. For maps with different cost cells and with one starting-one goal cell A* was the best in terms of computational time and size of memory.

Each of the mentioned researches conducted for shortest-path algorithms has proven to be beneficial for the specified unique scenario which shows that different problems require different approaches for acquiring the shortest path.

C. MICRO-SERVICES

Chen et al. [24] proposed a microservice-based architecture considering dynamic edge clouds to solve the optimization problem of a microservice-based framework. Coulson et al. [25] fostered a model in view of a microservice for web applications via an auto-scaling process and assessed it for prediction regulating AI. Halabi and Bellaiche [26] proposed a broker-based architecture to deal with service-level agreements of cloud security utilizing CIA principles (confidentiality, integrity, availability) against threats in the cloud.

Unlike the aforementioned works, in this manuscript, we apply an unsupervised learning method as the historical data does not contain any target label. Given that matter, K-means clustering, and a weight system is applied to identify the initial flood risk. We then consider the traveling salesman problem as the base case as it considers all possible paths to every single node and returns to the starting point.

III. DATA-INTELLIGENT SMART RESCUE FRAMEWORK

This section demonstrates an overview of the proposed smart data-driven framework (as depicted in Figure 1), which demonstrates the pre-flood and the post-flood procedures to minimize the loss. The pre-flood steps, vastly emphasize machine learning and quantitative analysis to focus on cure than mitigate the loss, whereas the post-flood steps deal with computational geometry and networking concepts to curtail the rescue time and overcome the limitations of the existing rescue procedures. The details are as follows:

A. PRE-FLOOD STEPS

Geospatial data are initially collected and stored in a database. While unlabeled geographic data is a common occurrence, this framework follows suit and works with unlabeled data. Thus, unsupervised machine learning plays a vital role in this scheme. Data is visualized to gain insights, and data-cleaning processes are applied to create more improvements. After

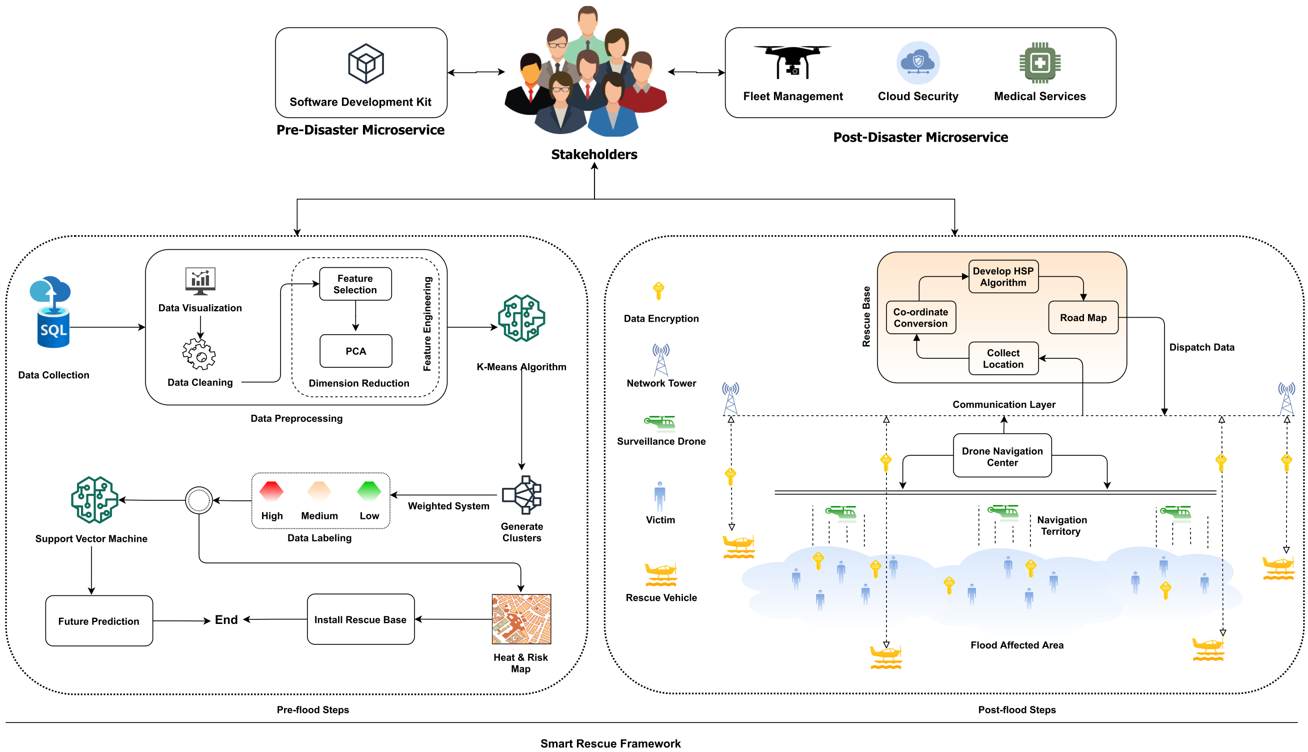


FIGURE 1. Proposed smart rescue framework.

the data is in an adequate state for analysis, less significant characteristics are eliminated which we call feature selection. Although the important features are picked, showing data points with a large number of attributes, which frequently refer to a higher dimension, makes the model complex and the behavior of the data trend difficult to comprehend. In order to minimize the dimension of data, principal component analysis (PCA) is offered.

Principal component analysis (PCA) is a well-known unsupervised machine learning technique that reduces the dimensionality of a dataset. PCA seeks to identify a new set of variables, known as principal components, that can account for the most significant amount of variance in the original dataset. PCA operates by identifying the direction of the data with the highest variance, then the direction orthogonal to the first direction with the highest variance, and so on until all principle components are discovered. Ordered by the amount of variance, each principal component is a linear combination of the original variables. By reducing the number of variables in a dataset, PCA can assist us in compressing the data while maintaining the vast majority of the pertinent information.

After finishing the feature selection part K-means, an unsupervised learning method is presented to build clusters and categorize them as high, medium, or low using a weighted 1 and 2) approach.

$$\sum_{i=1}^{k} r_i = r_1 + r_2 + \ldots\ldots + r_k \tag{1}$$

r_i refers to the final risk score, summed up the risk values(r_1, r_2,.. r_k) together of all features.

$$\alpha = \frac{\max\{f\ (r_i) : r_i \in z \wedge 0 < r_i \leq n\}}{d} \tag{2}$$

Here, α refers to the equal split for risk labeling, $f\ (r)$ is the final risk score of each cluster and d belongs to the number of classification labels.

Firstly, each attribute in each cluster is averaged to get its mean values. Next, for each attribute, the mean values of all clusters are compared with each other and given a score from 1 to the number of clusters. Here, 1 refers to the lowest risk, and the number of clusters refers to the highest risk. Once all attributes are compared, the risk values are summed up together to provide a final score for each cluster. Finally, the maximum value a cluster can acquire is divided by the number of classification labels (high, medium, and low) to split into equal portions where the size of each portion is α. The portions are multiplied with a coefficient corresponding to the increasing order. For instance, the value of the first portion should be α, the second portion should be 2α, etc. The higher value means more risk associated with it.

With this labeled data, risk maps are produced using QGIS (as depicted in Figure 10) to track down the most vulnerable territories and set up rescue bases. An add-on feature of this pre-flood step is to make a supervised predictive model from unlabelled data. From the weighted risk maps we labeled the dataset and the newly generated labeled dataset is fed into a supervised machine learning model to anticipate areas that are future flood inclined. We use support vector machine (SVM) [27] algorithm to develop the predictive model. According to the No Free Lunch theorem, there is no incentive to choose one machine learning algorithm over another until assumptions are made about the dataset [28]. This is related to inductive bias. Moreover, the support vector machine algorithm is effective in high-dimensional spaces and robust to outliers.

B. POST-FLOOD STEPS

Although several precautions are taken to reduce the risk of flooding, they do not adequately mitigate the collateral damage. As a result, we propose a series of collaborative tasks to be carried out after a flood has occurred in order to ensure the stability of our proposed framework in the face of any potential danger.

1) RESCUE BASE

One of the vital components of the post-flood step is the rescue base. A couple of sub-tasks are executed inside the control rescue base during the rescue operation. The workflow of the rescue base is detailed underneath.

- **Collect Location:** This section is responsible for gathering Geo-locations of victims through the communication layer.
- **Co-ordinate Conversion:** Geo-location (latitude and longitude) values are retrieved corresponding to the rescue base. However, rescue vehicles must need the locations of victims with respect to their location. In order to provide that, this phase transforms the original coordinate values to new values considering the assigned rescue vehicle as a center.
- **Construct SP Algorithm:** After transforming the values, the shortest path will be generated through hybrid A* (detail in Section IV-F). The blueprint of that hybrid shortest path will be dispatched to the communication layer.

2) COMMUNICATION LAYER

This layer works as a bridge to connect the rescue base with a rescue team. Telecommunication towers act as base stations to transmit data from the drone navigation centers to the rescue base. It is vital to have a reliable edge communication platform as noisy data may lead the rescue team in the wrong direction. Furthermore, considering the coverage area, to possess a feasible data transmission rate Wifi, Xbee technology might be considered. Although this architecture is fairly simple, to implement a more secure and stable communication layer other network architectures such as [29] might be useful to look at.

3) RESCUE TEAM

Surveillance drones search for victims utilizing the grid search manner. These drones are operated from the drone navigation center. If a drone detects any victim, the drone navigation center immediately captures the Geo-location of the victim. Once a single area is finished searching, those drones are assigned to search other territories that have not been visited yet. Meanwhile, the current data is sent to the rescue base for further processing through the communication layer. Afterward, the processed information is distributed to rescue vehicles to successfully execute the rescue operation within a curtailed period.

C. STAKEHOLDER

All the pre and post-flood steps are maintained by the parent entity otherwise known as stakeholders. In such scenarios, stakeholders are liable for offering various micro-services. Micro-services are as important as other components of the framework to operate the ecosystem smoothly. A few micro-services for rescue operations are discussed briefly in the accompanying subsections:

1) PRE-DISASTER MICRO-SERVICE

- Software Development Kit: During the clustering and prediction steps, one of the significant steps is to integrate all the outcomes appropriately and afterward interpret them to make a move accordingly. Establish a software development kit aids to perform these smoothly.

2) POST-DISASTER MICRO-SERVICES

- Fleet Management: Efficient rescue needs proper fleet management. It is admired that fleets must be proportionally distributed to all spots. This intends that there ought to be more vehicles where there are more casualties. In this way, rescue fleets can search areas optimally and rescue teams can accommodate all the victims in one go. Inefficient search and distribution of fleets would cost a great deal of time, capital, and in particular lives.
- Cloud and Data Security: Geographical information contains sensitive information and often has a threat of misusing. Thus, data would be encrypted while transmitting to avoid data breaching. Moreover, as personal information is needed to execute the proposed framework, micro-services such as cloud security are imposed in every data storage platform.
- Medical Services: Another important micro-service is to provide instant medical services to the victims in case of emergency. These services are available during the rescue operation inside the rescue fleets.

IV. METHODOLOGY

From section III we receive a clear synopsis of our proposed framework. However, in this manuscript, we split the workflow into several phases (as depicted in Figure 2]) which are demonstrated in the following subsections.

A. PROBLEM DEFINITION

As mentioned earlier in section II, conventional approaches focus on using the risk map for pre-flood preparations, of how to stop the floods from occurring by initially identifying the

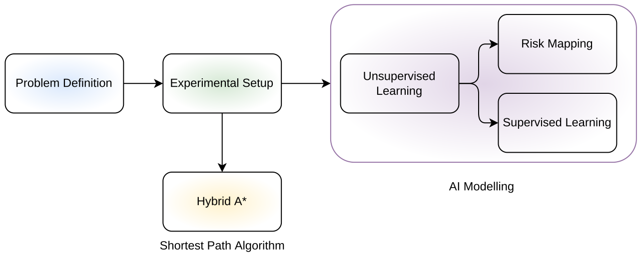


FIGURE 2. Methodology overview.

high-risk areas. Moreover, rescuing victims can take up to several weeks due to the absence of proper infrastructure for rescue methods. On the contrary, a risk map and shortest path algorithm both serve for pre-flood and post-flood preparations. Generating a risk map can identify the vulnerable areas and set the base stations and rescue teams in suitable positions before the hazard. Thus, they can rescue victims as soon as possible with the help of the shortest path. Post-flood risk management [30] is out of the scope of this manuscript.

B. EXPERIMENTAL SETUP

This research is mainly focused on the state of Terengganu in Malaysia. Terengganu is located on the east coast of Peninsular Malaysia and is the neighboring state to Kelantan and Pahang. It is situated between latitudes $05°51'06''N$ to $03°55'37''N$ and longitudes $102°21'11''E$ to $103°31'28''E$. Terengganu covers approximately 12,995 square kilometers and is comprised of eight districts and 33 towns. The place is normally hot and humid throughout most of the year, averaging at $28°C$ to $30°C$ during the daytime and a little cooler during the night. Its average annual rainfall is between 2575 and 2645 mm with heavy rain occurring mostly from December to January.

To generate results from our simulations, our data collection procedure involves two different stages. Firstly, the collection of spatial data on Terengganu is retrieved from a dataset that is recorded for the latest general election in Malaysia. This dataset needs to be cleaned and prepared before the next step. Next, the four main features are individually collected using Google Maps API, Malaysian state records, and weather forecast sources. The combined dataset is pre-processed before moving into the next stage.

C. APPLY UNSUPERVISED LEARNING

To begin with, Principal Component Analysis (PCA) [31] is applied for dimensionality reduction and scaling. From that, the first two principal components (composite combination of all attributes) are used for the K-Means model.

D. RISK MAPPING

In order to produce a risk and heat map, equations 1 and 2 are used to label the data point. Detail of this procedure are discussed in section V-D

E. SUPERVISED PREDICTIVE MODELLING

Since the target label (risk feature) has been acquired, an SVM classifier with a linear kernel and 10-fold cross-validation is implemented for predicting the risk feature which could be used for future scenarios instead of going through the complicated method of K-Means clustering and risk weights. This can be seen as turning unsupervised learning into supervised learning.

F. PROPOSED HYBRID SHORTEST PATH ALGORITHM

Our proposed algorithm has two major components. The first component is the relative coordinate conversion, which changes the coordinates of each victim(s) relative to the rescue base coordinates (origin). The search algorithm for the quickest path to the victims is the other component. Each component is detailed separately.

1) RELATIVE COORDINATE CONVERSION

To identify the location of victims corresponding to the rescue vehicle, we convert the coordinate values using the change of basis method.

Let the original coordinate (control center) is $\boldsymbol{C}$ basis and another coordinate (rescue vehicle) $\boldsymbol{R}$ basis. Where $\boldsymbol{C} = \{\vec{c_1} + \vec{c_2} + \ldots. + \vec{c_n}\}$ and $\boldsymbol{R} = \{\vec{r_1} + \vec{r_2} + \ldots. + \vec{r_n}\}$. We indicate the victim's location as $\vec{v}$. Thus, with respect to $\boldsymbol{C}$ basis, we formulate a linear combination of all the $\vec{c}$ vectors.

$$\beta_1\vec{c_1} + \beta_2\vec{c_2} + \ldots. + \beta_n\vec{c_n} = \vec{v} \tag{3}$$

Here, coefficients β values are the instruction of how much to stretch each of the different basis vectors. Similarly, for the $\boldsymbol{R}$ basis we write,

$$\gamma_1\vec{r_1} + \gamma_2\vec{r_2} + \ldots. + \gamma_n\vec{r_n} = \vec{v} \tag{4}$$

From equation 3 and 4 we get,

$$\beta_1\vec{c_1} + \beta_2\vec{c_2} + \ldots. + \beta_n\vec{c_n} = \gamma_1\vec{r_1} + \gamma_2\vec{r_2} + \ldots. + \gamma_n\vec{r_n}$$

As linear combination is some sort of matrix multiplication, we can form this equation into a matrix multiplication format.

$$\rightarrow \begin{pmatrix}\vec{c_1} & \ldots & \vec{c_n}\end{pmatrix}\begin{pmatrix}\beta_1\\ \cdot\\ \cdot\\ \beta_n\end{pmatrix} = \begin{pmatrix}\vec{r_1} & \ldots & \vec{r_n}\end{pmatrix}\begin{pmatrix}\gamma_1\\ \cdot\\ \cdot\\ \gamma_n\end{pmatrix}$$

The $\vec{c}$ and $\vec{r}$ vectors are the coordinate values with respect to $\boldsymbol{C}$ and $\boldsymbol{R}$ basis. Substituting the coefficient values and vectors we get,

$$\begin{aligned}\rightarrow P_C(\vec{v})_C &= P_R(\vec{v})_R\\ \rightarrow \vec{v}_R &= P_R^{-1}P_C(\vec{v})_C\end{aligned} \tag{5}$$

From equation 5, we get the co-ordinate value ($\vec{v}_R$) of the victim's location with respect to the rescue vehicle's basis $\boldsymbol{R}$. Here, P_R^{-1} is the inverse of coefficient matrix, P_C is the standard basis and $\vec{v}_C$ is the victim's location written on $\boldsymbol{C}$ basis.

2) SHORTEST PATH DECIDER

To find the optimal path, we use the A* search algorithm. However, there are several limitations to implementing the A* search algorithm in real-world applications. As a result, we alter the algorithm and create a variant of the A* - **Hybrid A*** search method. To comprehend the dynamics and properties of these algorithms, both A* and Hybrid A* search methods are introduced. In addition, a synopsis of the workflow is provided.

a: A SEARCH ALGORITHM*

To understand the workflow of A* in grid maps, there are a few jargons must be addressed which are described in the following:

- **Node:** Nodes are basically states or cells in the grid map. All potential positions of an environment. Each node has a unique identification.
- **Search Space:** A collection of nodes or areas inside the grid map where a robot can be placed and moved.
- **Cost:** Often refers to the distance (e.g. euclidean, diagonal) value from one node to another.
- **g(n):** Represents the exact cost of the path from the starting node to any node.
- **h(n):** Represents the heuristic estimated cost from any node to the goal node.
- **f(n):** Summation of exact cost and heuristic estimated cost.

$$f(n) = g(n) + h(n)$$

Considering the grid map (as depicted in Figure 5), we see that in the traditional A* algorithm, the representation anchors always stay in the center and two anchors connect with each other through a straight line. Thus movements can be done in eight directions, up, down, left and right, and diagonally. Each time movement occurs or enters a node, it calculates the cost, *f(n)*(n being the neighboring node), to travel to all of the neighboring nodes, and then enters the node with the lowest value of *f(n)*. Although a graph representing the euclidean distance is used to measure the cost, in the case of grid maps, the Manhattan distance is used to calculate heuristics. The euclidean square runs into the scale problem for grid maps. The scale of g and h needs to match because they are together to form f. When A* computes $f(n) = g(n) + h(n)$, the square of distance will be much higher than the cost g and the algorithm will end up with an overestimating heuristic. For longer distances, this will approach the extreme of $g(n)$ not contributing to $f(n)$, and A* will degrade into Greedy Best-First-Search. On the other hand, using euclidean with a square root gives the optimal path but takes a much longer time to execute A*.

b: PROPOSED HYBRID A SEARCH ALGORITHM*

Consider the grid map (as depicted in Figure 4), where the starting node position is (0,0) and the goal position is (4,4). There are 3 red blocks which are considered obstacles and no movements are allowed there. The vehicle can move in eight directions, up, down, left and right, and diagonally. Each time robot changes state or enters a node, it calculates the cost, f(n)(n being the neighboring node), to travel to all of the neighboring nodes, and then enters the node with the lowest value of f(n).

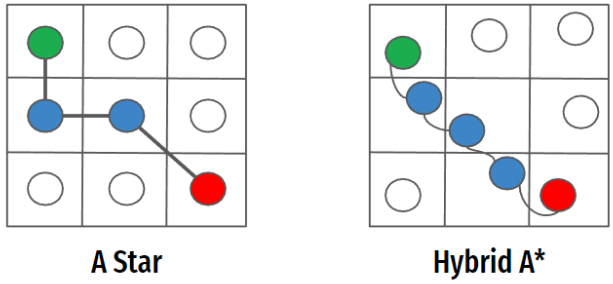


FIGURE 3. Path generation.

FIGURE 4. Grid map.

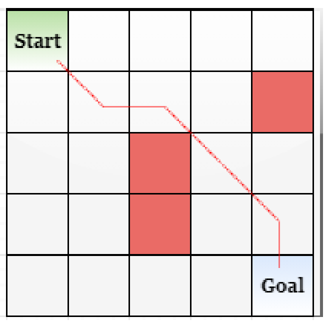


FIGURE 5. Shortest Path using Diagonal Heuristic.

As our map allows diagonal movement the heuristic of our algorithm is different and more efficient than the traditional A* algorithm in grid maps. Instead of moving first east and then north like Manhattan, we could simply move northeast directly. so the heuristic should be 4*D2, where D2 is the cost of moving diagonally. Here (as shown in Algorithm 1), we initialize the D1 and D2 as 1 where D1 is the cost of moving from one space to an adjacent space non-diagonally and D2 is the cost of moving diagonally. Then, we compute the number of steps taken if the algorithm can't take a diagonal **(line 3,4)**, then subtract the steps it saves ***(D2-2*D1)*** by using the diagonal **(line 5)**. There are ***min(dx, dy)*** diagonal steps, and each one costs D2 but saves 2*D1 non-diagonal steps. Hence, this heuristic is more effective than the conventional one.

Algorithm 1 Heuristic Function

Input: Node

Output: Heuristic Cost

Initialisation: Chebyshev Distance

1: $D1 \leftarrow 1$

2: $D2 \leftarrow 1$

Horizontal and Vertical Distance, Manhattan

3: $dx \leftarrow$ abs(node.x - goal.x)

4: $dy \leftarrow$ abs(node.y - goal.y)

Heuristic Cost

5: $H \leftarrow D * (dx + dy) + (D2 - 2 * D1) * min(dx, dy)$

6: **return** H

Now considering vehicle dynamics, the path generated from the A* search algorithm is impossible to follow for non-holonomic vehicles. There is a minimum level of curvature that non-holonomic car tracks, and thus, such sharp turns would be inconvenient for vehicles. Unlike the A* search algorithm, the hybrid A* search algorithm takes into account continuous points in the grid map (as depicted in Figure 5). In hybrid A* we consider the bicycle kinematic model (see Equation 1). As a result, the representation anchors move in such a way that non-holonomic vehicles can follow a continuous and smooth path.

$$\begin{aligned} x_n &= x_o + v\cos(\theta) \\ y_n &= y_o + v\sin(\theta) \\ \theta_n &= wrap\_angle(\theta + v\tan(\gamma)) \end{aligned}$$

For our case, we consider the water vessel as a non-holonomic vehicle, and thus, we modify the kinematic model according to our vehicle's possible movement. Our vehicle can turn to any angle first and then follow a straight path. Hence, we update the value of θ first and provide the θ_n to the robot to move in a direction (see Equation 2).

$$\begin{aligned} \theta_n &= wrap\_angle(\theta + v\tan(\gamma)) \\ x_n &= x_o + v\cos(\theta_n) \\ y_n &= y_o + v\sin(\theta_n) \end{aligned}$$

We define ten possible directional movements for the vehicle. Seven directional movements to move forward and three in the backward direction (as depicted in Figure 6). The forward directions are $0°$, $15°$, $20°$, $25°$, $-15°$, $-20°$, $-25°$ and the backward directions are $0°$, $20°$, $-20°$. For all the forward directions we are providing velocity $v = 1.5$ and for backward directions the velocity $v = -2$.

We also put additional costs for each possible movement, such as [0.2, 0.1, 0.2, 0.05, 0, 0.05]. The reason for adding such cost is to encourage the vehicle to move forward rather than backward. If we examine the additional cost array,

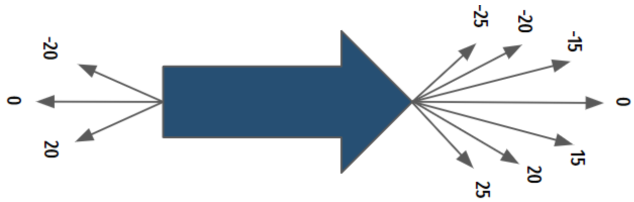


FIGURE 6. Vehicle's direction.

we can see the vehicle would be penalized more when it tries to move backward and the cost will be greater. So, the total cost for our hybrid A* search algorithm is:

$$f(n) = g(n) + h(n) + AdditionalCost$$

Here the heuristic is 3D Euclidean Distance:

$$h = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (\theta_1 - \theta_2)^2}$$

The rest of the finding path procedure is exactly the same as the A* search algorithm for grid maps.

V. VALIDATION, RESULT AND DISCUSSION

In the subsequent sections, we validate the proposed framework corresponds to different parameters. We find the optimal K value for the K-means algorithm verified by established methods. Furthermore, risk mapping and the performance of predictive modeling are also evaluated and discussed. The results are narrated below.

A. EXPLORATORY DATA ANALYSIS

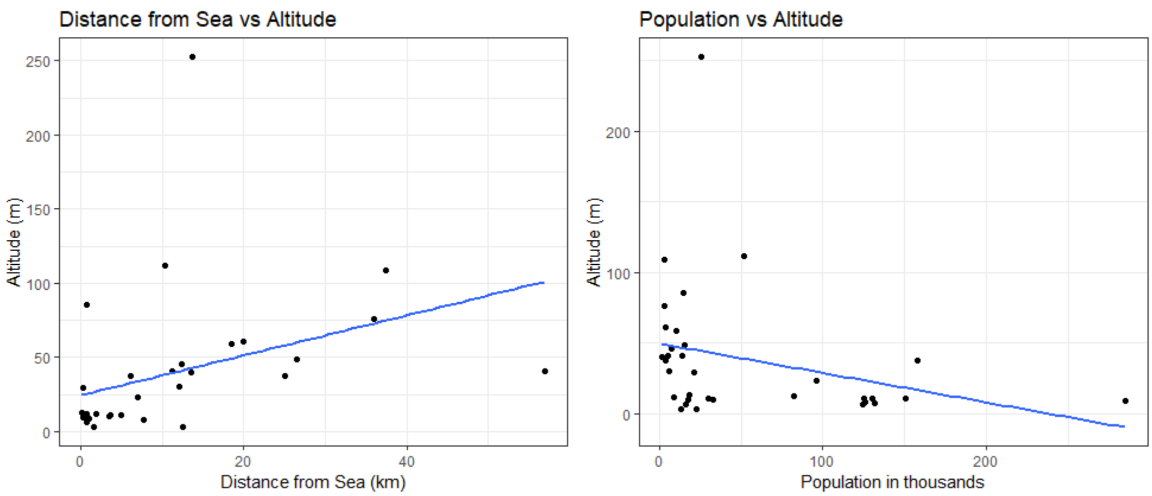
In the first stage, data is visualized using RStudio and QGIS to identify patterns and trends in the data. From Figure 7a, it can be understood that most of the populated towns were situated in areas close to the South China Sea and the land in that area had very low altitudes. In Figure 7b, it can be identified that the towns with high populations are located at low altitudes which means the majority of the population lives in the coastal area and is at risk of being affected during flood cases.

B. PRINCIPAL COMPONENT ANALYSIS

Next, in order to perform unsupervised machine learning algorithms data should be prepared in a way that reduces the intricacy of a model. To reduce the complexity of the machine learning model PCA is performed on the dataset.

Firstly, key features are selected using the Pearson correlation technique. From figure 8a it is evident that except for latitude and longitude remaining features are either positively or negatively correlated. For instance, there is a strong positive correlation between altitude and distance from the sea, while there is a negative correlation between population density and distance from the sea. Thus, those features namely population, distance_from_sea, altitude, and heavy_rainfall are selected for principal component analysis.

The drawback of PCA is that you risk losing information variability from the features that are combined. Hence,



(a) Distance vs Altitude

(b) Population vs Altitude

FIGURE 7. **Exploratory data analysis.**

a graph is plotted (as depicted in Figure 8b) to check the amount of explained variance that is captured by each component. The goal is to reduce the dimension of the data and keep as much variance from the original features. Figure 8b portrays, over 70% of variance is covered by components (dimension) 1 and 2. Thus, the initial two principal components are selected to be fed into the K-means algorithm as a trade-off between model intricacy and data originality.

The four main features, population, distance from the sea, altitude, and average annual rainfall were used to create clusters. To begin with, Principal Component Analysis (PCA) was applied for dimensionality reduction and scaling. From that, the first two principal components (composite combination of all attributes) were used for the K-Means model.

C. UNSUPERVISED ML MODEL: K-MEANS

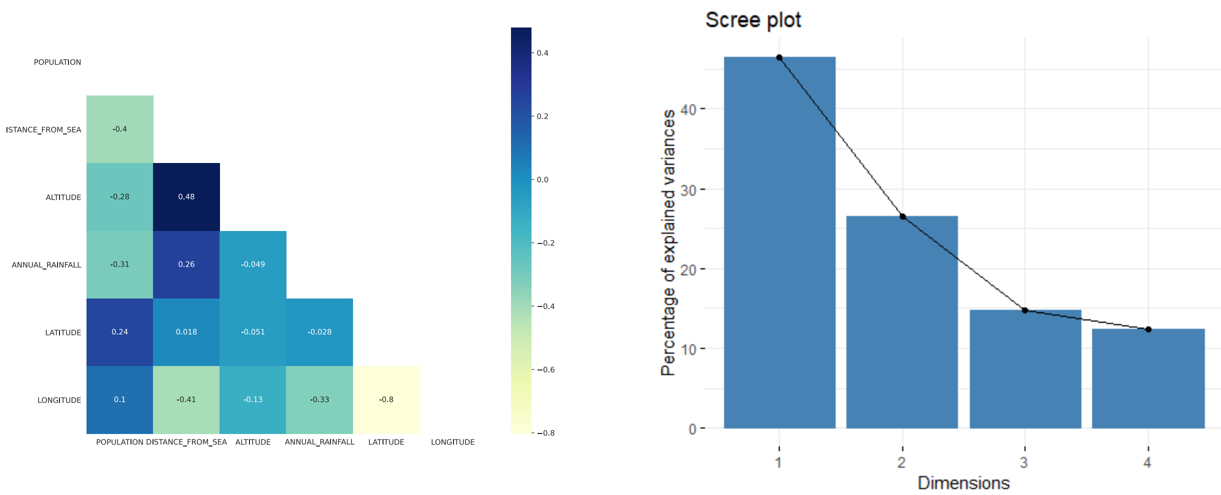
In order to identify the optimal number of clusters initially, the gap-static method [32] is implemented. However, Figure 11a, does not provide the exact optimal no. of clusters distinctly. Hence, the experiment is further run by applying the Elbow method (as depicted in Figure 11b) The Elbow method involves plotting the lowest within the sum of squared distance (WSS) values for different numbers of clusters and identifying the point where the curve starts to flatten out. Within sum of squared distance (WSS) measures the compactness of the clusters. It measures the sum of squared distances between each data point in a cluster and that cluster's centroid (mean). We note that the curve starts flattening at the sixth cluster. Thus six ideal clusters are picked up. We further verified the result using the Silhouette method shown in Figure 11c [33].

Subsequent to portraying the clusters in Figure 9, we find that cluster number 2, and 5 include more than 7 towns each. The greater part of the towns are placed under the second cluster and the least towns are in the fourth cluster. The remaining clusters range between 2 to 7 towns. These 6 clusters have achieved a total within the sum of squared distance (WSS) of 10.04 and between the sum of squared distance (BSS) of 83.35. The WSS should always be low as it is the comparison of points within the same cluster, and a smaller WSS means more similarity which should be the goal for points inside a cluster. The BSS on other hand is for looking at the difference between each cluster so a high value is great as it indicates that each cluster is unique from the others. The total variance achieved is 89.2% which indicates that the model is effective. Furthermore, this is compared to other models that did not include scaling or PCA, and this model performed much better with the highest total variance.

D. WEIGHTED RISK MAP

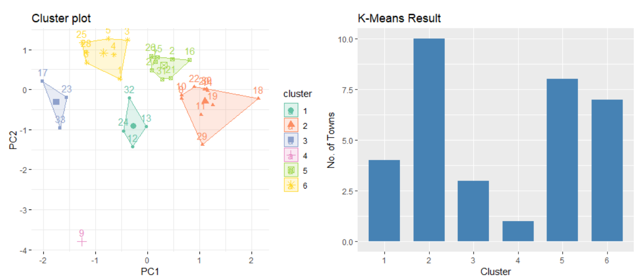
In order to identify the main flood-prone region we develop a weighted risk and heat map. According to the maps, rescue teams can set up their bases there during the early monsoon season. From Figure 10a, towns situated beside the south china sea are heavily flood-prone as they have a higher weighted value. As we move away from the coast, the risk decreases. As illustrated by the heat map in Figure 10b, the town of Hulu Terengganu is the most prone to flooding in all these locations.

To classify these risk clusters (shown in Figure 10) according to low, medium, and high risk, a weighing system (as referred to equation 1 and 2) is applied. Firstly, each attribute in each cluster s averaged to get its mean values. Next, for each attribute, the mean values of all six clusters are compared with each other and given a score from 1 to 6, where 1 refers to the lowest risk and 6 refers to the highest risk. Once all attributes are compared, the risk values are summed up together to provide a final score for each cluster. Given that the maximum number of points a cluster can acquire is 24, this value is split into three sections, low for a value close to 8,



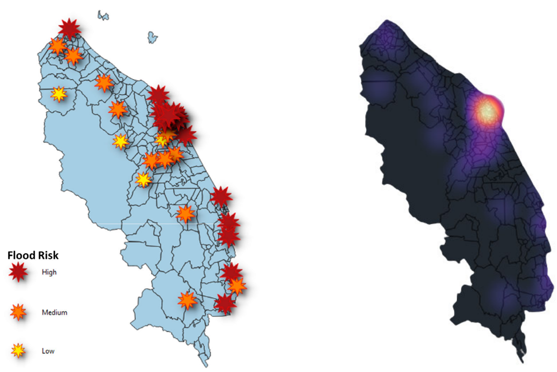
(a) Pearson Correlation Heatmap (b) Variance of Dimension

FIGURE 8. Principal component analysis.



(a) Clusters (b) Number of towns in each clusters

FIGURE 9. Size of clusters.



(a) Risk Map (b) Heat Map

FIGURE 10. Generated maps.

medium for a value close to 16, and high for a value close to 20. The outcome results in each cluster being assigned a low, medium, or high-risk classification. A flood risk map and heat map are then created using QGIS as depicted in Figure 10a and 10b.

E. SUPERVISED ML MODEL: SVM

For the SVM classifier, a linear kernel with a soft margin of two is used. The model is iterated through the ten-fold cross-validation method. From table 1, it is evident that the SVM classifier has a high True Positive Rate for all the classes.

TABLE 1. Confusion matrix.

		Predicted Label		
		High	Medium	Low
True Label	High	18	0	0
	Medium	2	9	0
	Low	0	2	2

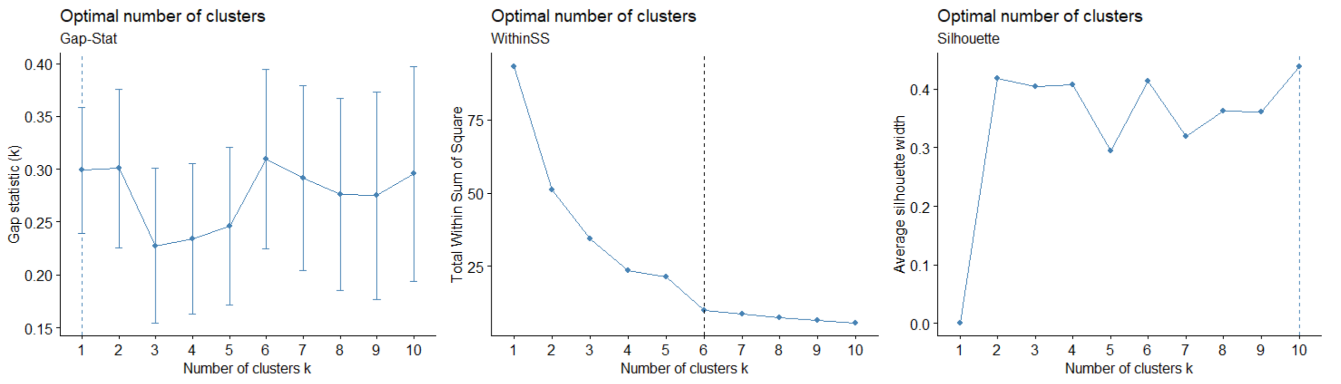
The accuracy of the prediction results is 90.17%. The prediction of high and medium risk is significantly high, while low risk appears satisfactory. The 'High' risk class had a precision and recall of 0.90 and 1.00 respectively, while the 'Medium' risk class achieved 0.82 for both precision and recall. The 'Low' risk class had a precision of 1.00 and a recall of 0.50. The F-1 score can also be seen through Table 2. To sum it up, the SVM model worked well in predicting accurately in most cases.

TABLE 2. Performance evaluation.

		Precision	Recall	F1-Score
Classification Report	High	0.90	1.00	0.95
	Medium	0.82	0.82	0.82
	Low	1.00	0.50	0.67

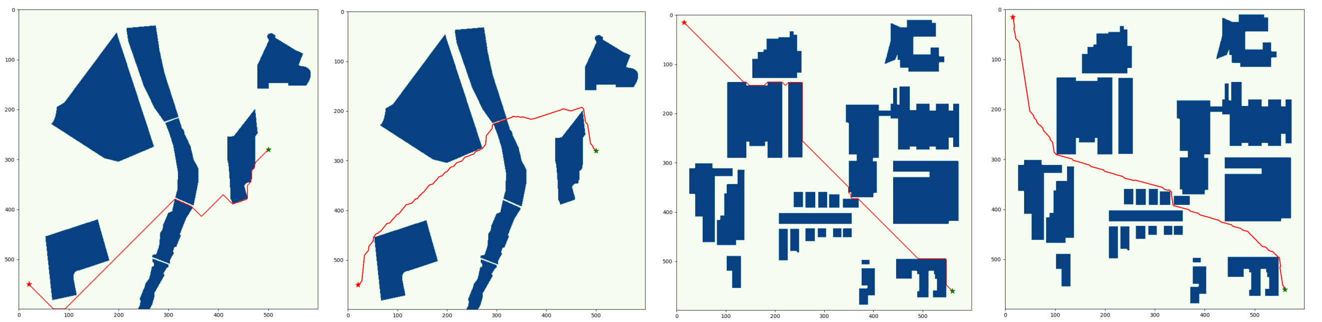
F. SHORTEST PATH FINDER: A AND HYBRID A**

We assume the area map is a grid map where each grid represents a vehicle waypoint. The blue cells in Figure 12 represent obstacles, while the white cells represent navigable spaces.



(a) Gap-Stat Method (b) Elbow Method (c) Silhouette Method

FIGURE 11. **Optimal k-means clusters selection.**



(a) A* Search (b) Hybrid A* Search (c) A* Search (d) Hybrid A* Search

FIGURE 12. **Path generation comparison.**

In Figures 12a and 12c, the optimal shortest path is determined by the A* search method. Nevertheless, the path is discontinuous, has sharp turns, and passes near obstacles. This would impede the mobility of rescue vehicles and increase the likelihood of a collision. In contrast, our suggested Hybrid A* search method yields a suboptimal, continuous, and smooth path in Figures 12b and 12d. Although the route is not the shortest in terms of distance, the uninterrupted route allows the vehicle to operate without kinematic constraints. In this circumstance, our hybrid algorithm performs more efficiently in terms of qualitative aspects.

G. COMPLEXITY AND COST ANALYSIS

In order to deploy the aforementioned framework, there are a few perspectives that we need to consider such as (i) Data Collection and Data Storage Cost; (ii) System Design Cost; (iii) Network Infrastructural Cost iv) Computational Cost. In terms of data collection and data storage, as geographical data are available to Government, additional costs will not be executed. Next, for system design, we will rely on such design solutions that are affordable to the government. As this framework will exploit existing telecommunication networks, we don't need additional costing. Moreover, the network utilization cost would be relinquished considering the support from telecommunication companies to the Government for taking control of the natural hazard. Finally, the fanciest costing would be the computational cost which is assumed to take care of by the Government and Charity Organisations. As a result, the implementation cost will not be a significant matter and the complexity would be low as well.

VI. FUTURE WORKS

This framework is comprised of numerous components and can therefore be improved in a variety of ways. For instance, real-time image recognition for unmanned aerial vehicles could be implemented so that humans are not required to conduct drone surveillance manually. The number of microservices could be increased by acquiring additional data to enhance machine learning models. Lastly, the fairness path prioritizes victims whose locations require heightened attention due to higher water levels, collapsing structures, or other emergencies.

VII. CONCLUSION

Recent developments in big data and machine learning have expanded the scope of smart city development. Utilizing these ideas to mitigate the adverse effects of natural disasters would enhance the quality of smart cities. Consequently, the proposed method in this manuscript would be an excellent addition to a smart city. This strategy could also be applied to other natural disasters necessitating search-and-rescue operations. This study also paves the way for future work on optimization strategies to enhance the performance of shortest-path algorithms in rescue operations.

Micro-services that hold the ecosystem together and make it secure to use are one of the primary components of this research. A flood risk map is generated using machine learning and GIS to determine where rescue teams can establish bases. The machine learning models worked well with K-Means achieving a total variance of 89.2% and SVM achieving a prediction accuracy of 90.17%. Our hybrid A* search algorithm identifies a suboptimal, smooth route to rescue the victims.

REFERENCES

[1] B. Shi, T. Zeng, C. Tang, L. Zhang, Z. Xie, G. Lv, and Q. Wu, ''Landslide risk assessment using granular fuzzy rule-based modeling: A case study on earthquake-triggered landslides,'' *IEEE Access*, vol. 9, pp. 135790–135802, 2021.

[2] Z. A. Akasah and S. V. Doraisamy, ''2014 Malaysia flood: Impacts and factors contributing towards the restoration of damages,'' *J. Sci. Res. Develop.*, vol. 2, no. 14, pp. 53–59, 2015.

[3] F. Buslima, R. Omar, T. Jamaluddin, and H. Taha, ''Flood and flash flood geo-hazards in Malaysia,'' *Int. J. Eng. Technol.*, vol. 7, pp. 760–764, Nov. 2018.

[4] N. W. Chan, ''Increasing flood risk in Malaysia: Causes and solutions,'' *Disaster Prevention Manag., Int. J.*, vol. 6, no. 2, pp. 72–86, May 1997.

[5] R. Ghorani, S. Fattaheian-Dehkordi, M. Farrokhi, M. Fotuhi-Firuzabad, and M. Lehtonen, ''Modeling and quantification of power system resilience to natural hazards: A case of landslide,'' *IEEE Access*, vol. 9, pp. 80300–80309, 2021.

[6] B. Pradhan, ''Flood susceptible mapping and risk area delineation using logistic regression, gis and remote sensing,'' *J. Spatial Hydrol.*, vol. 9, no. 2, pp. 1–18, 2010.

[7] M. S. B. Khalid and S. B. Shafiai, ''Flood disaster management in Malaysia: An evaluation of the effectiveness flood delivery system,'' *Int. J. Social Sci. Humanity*, vol. 5, no. 4, pp. 398–402, 2015.

[8] N. S. Romali, Z. Yusop, M. Sulaiman, and Z. Ismail, ''Flood risk assessment: A review of flood damage estimation model for Malaysia,'' *Jurnal Teknologi*, vol. 80, no. 3, pp. 145–153, Feb. 2018.

[9] N. W. Chan, ''Flood disaster management in Malaysia: An evaluation of the effectiveness of government resettlement schemes,'' *Disaster Prevention Manag., Int. J.*, vol. 4, no. 4, pp. 22–29, Oct. 1995.

[10] M. A. Rahman, M. S. Abuludin, L. X. Yuan, M. S. Islam, and A. T. Asyhari, ''EduChain: CIA-compliant blockchain for intelligent cyber defense of microservices in education industry 4.0,'' *IEEE Trans. Ind. Informat.*, vol. 18, no. 3, pp. 1930–1938, Mar. 2022.

[11] N. A. Harun, M. Makhtar, A. A. Aziz, Z. A. Zakaria, F. S. Abdullah, and J. A. Jusoh, ''The application of apriori algorithm in predicting flood areas,'' *Int. J. Adv. Sci., Eng. Inf. Technol.*, vol. 7, no. 3, pp. 763–769, 2017.

[12] G. A. M. de Almeida, P. Bates, and H. Ozdemir, ''Modelling urban floods at submetre resolution: Challenges or opportunities for flood risk management?'' *J. Flood Risk Manage.*, vol. 11, pp. S855–S865, Feb. 2018.

[13] A. A. Monrat, R. U. Islam, M. S. Hossain, and K. Andersson, ''Challenges and opportunities of using big data for assessing flood risks,'' in *Applications of Big Data Analytics*. Cham, Switzerland: Springer, 2018, pp. 31–42.

[14] S. L. B. Sa'adin, S. Kaewunruen, and D. Jaroszweski, ''Heavy rainfall and flood vulnerability of Singapore-Malaysia high speed rail system,'' *Austral. J. Civil Eng.*, vol. 14, no. 2, pp. 123–131, Oct. 2016.

[15] M. S. Tehrany, B. Pradhan, and M. N. Jebur, ''Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS,'' *J. Hydrol.*, vol. 512, pp. 332–343, May 2014.

[16] H. Mojaddadi, B. Pradhan, H. Nampak, N. Ahmad, and A. H. B. Ghazali, ''Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS,'' *Geomatics, Natural Hazards Risk*, vol. 8, no. 2, pp. 1080–1102, Dec. 2017.

[17] Y. Ouma and R. Tateishi, ''Urban flood vulnerability and risk mapping using integrated multi-parametric AHP and GIS: Methodological overview and case study assessment,'' *Water*, vol. 6, no. 6, pp. 1515–1545, May 2014.

[18] M. B. Kia, S. Pirasteh, B. Pradhan, A. R. Mahmud, W. N. A. Sulaiman, and A. Moradi, ''An artificial neural network model for flood simulation using GIS: Johor river basin, Malaysia,'' *Environ. Earth Sci.*, vol. 67, no. 1, pp. 251–264, Sep. 2012.

[19] T. M. Shafapour, B. Pradhan, and M. N. Jebur, ''Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS,'' *J. Hydrol.*, vol. 504, no. 8, pp. 69–79, Nov. 2013.

[20] F. Jalayer, R. De Risi, F. De Paola, M. Giugni, G. Manfredi, P. Gasparini, and M. E. Topa, ''Probabilistic GIS-based method for delineation of urban flooding risk hotspots,'' *Natural Hazards*, vol. 73, no. 2, pp. 975–1001, 2014.

[21] T. Takeda, K. Ito, and F. Matsuno, ''Path generation algorithm for search and rescue robots based on insect behavior—Parameter optimization for a real robot,'' in *Proc. IEEE Int. Symp. Saf., Secur., Rescue Robot. (SSRR)*, Oct. 2016, pp. 270–271.

[22] M. Becker, F. Blatt, and H. Szczerbicka, ''A multi-agent flooding algorithm for search and rescue operations in unknown terrain,'' in *Proc. German Conf. Multiagent Syst. Technol.* Cham, Switzerland: Springer, 2013, pp. 19–28.

[23] S. M. MasudurRahmanAl-Arif, A. H. M. I. Ferdous, and S. H. Nijami, ''Comparative study of different path planning algorithms: A water based rescue system,'' *Int. J. Comput. Appl.*, vol. 39, no. 5, pp. 25–29, Feb. 2012.

[24] L. Chen, Y. Xu, Z. Lu, J. Wu, K. Gai, P. C. K. Hung, and M. Qiu, ''IoT microservice deployment in edge-cloud hybrid environment using reinforcement learning,'' *IEEE Internet Things J.*, vol. 8, no. 16, pp. 12610–12622, Aug. 2020.

[25] N. C. Coulson, S. Sotiriadis, and N. Bessis, ''Adaptive microservice scaling for elastic applications,'' *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4195–4202, May 2020.

[26] T. Halabi and M. Bellaiche, ''A broker-based framework for standardization and management of cloud security-SLAs,'' *Comput. Secur.*, vol. 75, pp. 59–71, Jun. 2018.

[27] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf, ''Support vector machines,'' *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul./Aug. 2008.

[28] M. A. Rahman, N. Zaman, A. T. Asyhari, S. M. N. Sadat, P. Pillai, and R. A. Arshah, ''SPY-BOT: Machine learning-enabled post filtering for social network-integrated industrial Internet of Things,'' *Ad Hoc Netw.*, vol. 121, Oct. 2021, Art. no. 102588.

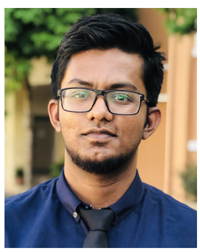
[29] M. A. Rahman, N. Zaman, A. T. Asyhari, F. Al-Turjman, M. Z. A. Bhuiyan, and M. F. Zolkipli, ''Data-driven dynamic clustering framework for mitigating the adverse economic impact of COVID-19 lockdown practices,'' *Sustain. Cities Soc.*, vol. 62, Nov. 2020, Art. no. 102372.

[30] J. M. Bodoque, A. Díez-Herrero, M. Amerigo, J. A. García, and J. Olcina, ''Enhancing flash flood risk perception and awareness of mitigation actions through risk communication: A pre-post survey design,'' *J. Hydrol.*, vol. 568, pp. 769–779, Jan. 2019.

[31] I. T. Jolliffe and J. Cadima, ''Principal component analysis: A review and recent developments,'' *Philos. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 374, Apr. 2016, Art. no. 20150202.

[32] R. Tibshirani, G. Walther, and T. Hastie, ''Estimating the number of clusters in a data set via the gap statistic,'' *Statist. Methodol.*, vol. 63, no. 2, pp. 411–423, 2004.

[33] F. Wang, H.-H. Franco-Penya, J. D. Kelleher, J. Pugh, and R. Ross, ''An analysis of the application of simplified silhouette to the evaluation of k-means clustering validity,'' in *Proc. Int. Conf. Mach. Learn. Data Mining Pattern Recognit.* New York, NY, USA: Springer, Jul. 2017, pp. 291–305.



NAFEES ZAMAN received the B.Sc. degree in computer science from International Islamic University Malaysia, in 2020. He is currently pursuing the joint master's degree in intelligent field robotic systems with the University of Girona, Spain, and the University of Zagreb, Croatia. He is currently a Visiting Researcher with the KTH Royal Institute of Technology, Stockholm, Sweden. Previously, he worked as a Machine Learning Research Intern with IBM CoE Malaysia and VICOROB, Spain. His research areas include intelligent systems (robotics), computer vision, 3-D deep learning, and computer graphics. He was a recipient of the Erasmus Mundus Scholarship 2021–2022 and other awards, including the Best Innovation Award in Malaysia Technology Expo Special Edition 2020, the Gold Medal in National Innovation and Invention Competition Malaysia 2019, and the second runner-up in Nasa Space App Challenge Malaysia 2019.

AHMAD ABU SAIID received the B.Sc. degree (Hons.) in computer science from International Islamic University Malaysia (IIUM), in 2020. He is currently pursuing the joint M.Sc. and master's degree in big data management and analytics (BDMA) with Université libre de Bruxelles, Belgium, as an Erasmus Mundus scholarship holder. He possesses industrial experience in data analytics, specifically in the fintech domain.

MD ARAFATUR RAHMAN (Senior Member, IEEE) received the Ph.D. degree in electronic and telecommunications engineering from the University of Naples Federico II, Naples, Italy, in 2013. He has around 15 years of research and teaching experience in the domain of computer and communications engineering. He was an Associate Professor with the Faculty of Computing, Universiti Malaysia Pahang. He worked as a Postdoctoral Research Fellow with the University of Naples Federico II, in 2014, and as a Visiting Researcher with the Sapienza University of Rome, in 2016. He is currently a Senior Lecturer with the School of Engineering, Computing and Mathematical Sciences, University of Wolverhampton, U.K. He has coauthored around 150 articles in prestigious IEEE and Elsevier journals, such as IEEE Transactions on Industrial Informatics, IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Green Communications and Networking, IEEE Transactions on Services Computing, *IEEE Communications Magazine*, *JNCA* (Elsevier), and *FGCS* (Elsevier), and in conference publications such as IEEE GLOBECOM and IEEE DASC. His research interests include the Internet of Things (IoT), wireless communication networks, cognitive radio networks, 5G, vehicular communication, cyber-physical systems, big data, cloud-fog-edge computing, machine learning, and security. He has served as a Specialty Chief Editor of the IoT Theory and Fundamental Research of IEEE Internet for Things journal (Frontiers), an advisory board member, an Editor of *Computers* (MDPI), a Lead Guest Editor of IEEE Access and *Computers*, an Associate Editor of IEEE Access, a Patron, the General Chair, an Organizing Committee Member, the Publicity Chair, the Session Chair, a Programme Committee Member, and a member of the Technical Programme Committee (TPC) of numerous leading conferences worldwide, such as IEEE GLOBECOM, IEEE DASC, IEEE iSCI, and IEEE ETCCE, and journals. His name was enlisted in the world's Top 2% scientists list released by Stanford University under the category of "Citation Impact in Single Calendar Year 2019–2021." He was awarded a Higher Education Academy (HEA) Fellowship from the U.K. He was endorsed by the Royal Academy of Engineering, U.K., as global talent under the category of "Exceptional Talent" in 2022.

SHAVAN ASKAR received the B.Sc. (Hons.) and M.Sc. degrees from the Control and Systems Engineering Department, Baghdad, in 2001 and 2003, respectively, and the Ph.D. degree in electronic systems engineering from the University of Essex, U.K., in 2012. He is currently an Associate Professor of Computer Networks at Erbil Polytechnic University, Iraq. He works in the field of networks that include the Internet of Things, software-defined networks, optical networks, and 5G. He has published more than 60 scientific research articles, some of which were published in very prestigious conferences such as OFC and ECOC, and high impact factor (3.58) journals such as *Optics Express*.

JASNI MOHAMAD ZAIN received the bachelor's degree in computer science from the University of Liverpool, U.K., in 1989, and the Ph.D. degree from Brunel University, West London, U.K., in 2005. She started her career as a Tutor with the University of Technology Malaysia (UTM), in 1997. She was the Deputy Director (Cybertechnology) of Research Nexus UiTM and the Head of Advanced Analytics and Engineering Centre (AAEC). She was the Dean of the Faculty of Computer Systems and Software Engineering, Universiti Malaysia Pahang, for seven years. She is currently the Director of the Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), UiTM Shah Alam, Malaysia. She has been actively presenting papers and keynote addresses at national and international conferences. Her research interests include data mining, digital watermarking, image processing, and network security. She has supervised 15 Ph.D. and six master's students and published more than 100 refereed articles. She has a patent file for digital watermarking (PI 2008047).