Spatial mapping of landslide susceptibility in Jerash governorate of Jordan using genetic algorithm-based wrapper feature selection and bagging-based ensemble model

Abdel Rahman Al-Shabeeb, A’kif Al-Fugar, Khaled Mohamed Khedher, Ali Nouh Mabdeh, and Rida Al-Adamat

Department of GIS and Remote Sensing, Institute of Earth and Environmental Sciences, Al Al-Bayt University, Mafraq, Jordan; Department of Surveying Engineering, Faculty of Engineering, Al Al-Bayt University, Mafraq, Jordan; Department of Civil Engineering, College of Engineering, King Khalid University, Abha, Saudi Arabia; Department of Civil Engineering, High Institute of Technological Studies, Mrezgua University Campus, Nabeul, Tunisia

ABSTRACT
This study employs five genetic algorithm (GA)-based machine learning (ML) models, namely the Decision Tree (DT), k-Nearest Neighbors (kNN), Naïve Bayes (NB), Support Vector Machine (SVM), and Extreme Learning Machine (ELM), to build a novel ensemble algorithm that is founded on the Bagging method for landslide susceptibility mapping (LSM) in Jerash Governorate, north of Jordan. The GA-based wrapper feature selection (FS) was done based on the five individual models and in the initial stages of modeling, an inquiry for the best feature for each of the five models was made. Finally, five hybrid models, namely DT-GA, kNN-GA, NB-GA, SVM-GA, and ELM-GA were constructed and combined to create Bagging-based ensemble model. The FS outcomes uncovered that rainfall depth, distance to roads, the Stream Power Index, the Normalized Difference Vegetation Index, slope, geology, and aspect are the most influential determinants of landslides. After the significant variables were identified, they were selected as input predictors and entered into the models. GA-based Bagging ensemble model with the area under the receiver operating characteristic curve (AUROC) of 0.85 achieved the highest accuracy in the validation run, followed by SVM-GA (AUROC = 0.80), NB-GA (AUROC = 0.76), DT-GA (AUROC = 0.72), kNN-GA (AUROC = 0.70), and ELM-GA (AUROC = 0.48).

ARTICLE HISTORY
Received 19 May 2022
Accepted 6 August 2022

KEYWORDS
Ensemble models; hazard mapping; landslide susceptibility mapping; machine learning

1. Introduction
Natural disasters quite often result in injuries, loss of life, damage to infrastructure and sites of cultural heritage, and displacement of people, which bring about...
substantial economic losses. Landslide is a natural hazard leading to extensive loss of life and property across the world every year (Schuster and Fleming 1986; Petley 2012). As human population increases and cities are extended to steep regions, the risk of mass movement of earth materials increases (Smyth and Royle 2000; Sassa et al. 2004). In densely-populated regions, this mass movement can be highly disastrous (Knapen et al. 2006). Landslides, particularly those induced by rainfall, are frequently the most devastating natural disasters in the rugged and the mountainous areas. They usually lead to life losses, property damage, and economic losses (Youssef and Pourghasemi 2021). In the past, landslides were accompanied by high mortality rates such as the cases of landslide on 2014 in Afghanistan, which killed nearly 2,700 persons and landslide on 2014 in Shanxi, China, which killed 64 persons (Zhang et al. 2015; Lin and Wang 2018). Therefore, efforts to early identify landslide-susceptible lands and take them into account in land use planning can sharply reduce the landslide risk and associated damage and losses (Fell et al. 2008).

In Jordan, flash floods and landslides are more frequent than other natural hazards. The governorate of Jerash in Jordan is the particular governorate that experiences most of the natural disasters, especially landslides. Almost every year, landslides take place in this governorate, mostly triggered by meteorological factors (e.g., heavy rainfall) and human activities, e.g., construction, road excavation, and agricultural activities (Malkawi et al. 2000). The city of Jerash is prominent in historical and scenic natural sites. As such, it enjoys high potential for tourism. Bearing this in mind, analysis and mapping of landslide susceptibility in this governorate is badly needed in order to reduce the damage associated with landslides.

Landslide susceptibility mapping (LSM) is an action by which landslide-susceptible lands can be identified and, based on that, construction and irrelevant uses of land can be prohibited (Saha et al. 2005). Maps of land susceptibility to slides are quite vital sources of fundamental data that are frequently needed for determining the areas under risk of landslides and for reducing the probable losses associated with the landslides when they occur (Yi et al. 2020; Thi Ngo et al. 2021). In previous studies, there have been numerous attempts to identify landslide-susceptible zones using a variety of approaches (Zeybek and Şanlıoğlu 2015; Zeybek et al. 2015). However, the flexibility and accuracy of the approaches differ, depending on the geological and climatic characteristics of the studied regions.

A review of literature on this research theme uncloses that, in general, the task of LSM is achieved using either of two main approaches: knowledge-based and data-driven approaches (Arca et al. 2018; Mallick et al. 2018; Meena et al. 2019; Zhu et al. 2019; Mabdeh et al. 2022). The former approach encapsulates a wide range of techniques, the most famous of which is the Analytical Hierarchy Process (Pawluszek and Borkowski 2017; Abedini and Tulabi 2018; Mandal and Mandal 2018). In general, the outputs of knowledge-based approaches are dependent on experts’ opinion, where contradictory and dogmatic comments may affect the results. Many of these approaches for LSM fail to deal with predictors that have negative values (e.g., the normalized difference vegetation index (NDVI)). As well, some techniques like entropy (Zhang et al. 2018) have limited capabilities to handle such data. Nevertheless, these groups of approaches performed well in many applications (Turan...
et al. 2020; Ali et al. 2021). The data-driven approach uses historical data for modeling (Zhu et al. 2018). In other words, the data-driven models use data on previous landslides locations and consider their influential factors in the modeling process (He et al. 2019; Nhu et al. 2020). This approach encircles a wide range of models, including the artificial neural networks (e.g., Sevgen et al. 2019; Manohar and Anbazhagan 2021; Youssef and Pourghasemi 2021); decision tree (e.g., Bui et al. 2012; Pradhan 2013; Nhu et al. 2020; Biçici and Mustafa 2021); generalized linear model (e.g., Goetz et al. 2015; Pourghasemi and Rahmati 2018); logistic model tree (e.g., Chen et al. 2017; Pham et al. 2018); Naïve Bayes classifier (Hungr et al. 2014; Zhang et al. 2018; He et al. 2019; Sevgen et al. 2019; Nhu et al. 2020; Akinci et al. 2021); and the support vector machine (e.g., Akinci and Zeybek 2021).

In previous studies, researchers made attempts to enhance the accuracy of machine learning models in LSM. For achieving this purpose, numerous hybrid models have been proposed (Li and Chen 2019; Paryani et al. 2020; Chen and Chen 2021). Large number of these models employed heuristic algorithms in order to tune hyper-parameters of machine learning models, and, accordingly, they have the disadvantage of high computational complexity and running time. In recent years, the ensemble models, as a viable alternative, have attracted the attention of researchers in various fields (Tehrany et al. 2014; Kadavi et al. 2018; Roy et al. 2019; Arabameri et al. 2020). Unlike the hybrid models, the ensemble models are not computationally expensive. Furthermore, they combine the outputs of simple, weak algorithms to generate models with higher accuracies than those of the simple models. Bearing this in mind, this study aimed at developing an ensemble model for LSM in an effort to solve the problem of recurrent landslides in the governorate of Jerash, north of Jordan. In addition, the study took into consideration the problem of accurate feature selection in the pre-modeling phase and proposed a suitable solution for this problem. It should be underscored that despite the large effect of feature selection on performance of the machine learning algorithms, most of the earlier studies failed to properly address this step.

As mentioned earlier, the models commonly used in ensemble models have been constructed without considering the feature selection, which is one of the basic preprocessing steps in modeling using machine learning methods (Kadavi et al. 2018; Pham et al. 2020; Akinci et al. 2021; Pham et al. 2022). In fact, the use of redundant and irrelevant factors in the models increases the running time of the algorithms, the volume of data used for modeling, and in some cases, reduces the accuracy of the modeling (Shafizadeh-Moghadam et al. 2017). Therefore, in the present study, first the considered factors are entered into the feature selection phase based on a wrapper feature selection approach. Then the optimal factors are selected and used in individual models. Finally, these hybrid models constructed by wrapper-based genetic algorithm (GA) and machine learning models are used to construct the ensemble model based on bagging. Therefore, this study contributes to the literature by proposing a novel ensemble model based on Bagging and wrapper-based feature selection using GA for LSM. In consequence, the aim of this study was to produce landslide susceptibility maps for the governorate of Jerash using a bagging ensemble model constructed by the Decision Tree (DT), k-Nearest Neighbors (kNN), NaïveBayes (NB), Support
Vector Machine (SVM), and Extreme Learning Machine (ELM) models and GA-based wrapper feature selection.

2. Study area

The area selected for this study is Jerash Governorate, which is characterized by recurrent landslides, especially in winter. It lies northwest of Jordan (32° 16’ 12.12” N, 35° 53’ 17.41” E) and has a land area of 420 km² (Figure 1). The total population of Jerash is 262,100 capita, which corresponds to a population density of 639.6 people/km². This governorate is characterized by hot weather in the summer and cold to moderately-cold weather in winter.

Jerash has varying altitudes that lie within the range of 2–1,229 m above sea level. Accordingly, this governorate is considered a mountainous area. However, there are parts of Jerash that lie under the level of sea surface. This suggests presence of steep slopes in this governorate. These steep slopes are actually the main culprit for land sliding in this governorate. The field visits and inspection works performed by the researchers in Jerash disclosed that frequent mass movements (that is, landslides and rock falls) take place in this governorate. The traces on topography of this area that had been left by natural disasters are readily seen by the naked eye. Indeed, these and other natural phenomena keep altering topography of land. In other respects, the

Figure 1. Map of the study area.
recurrent landslides result in substantial economic losses, due, for example, to damage to roads, electricity infrastructure, and telephone lines and because of sheeting of slopes and construction of retaining walls to protect people from direct exposure to landslides (Figure 2). Furthermore, rock falls that result from geological structure of Jerash are somewhat frequent. They produce considerable material damage and create steady fear in the nearby community. Recently, investment in Jerash and the infrastructure and urban developments which have been made by the government and local people brought about growth in tourism infrastructure, which made a study of this kind necessary for helping the urban planners and decision makers in knowing the locations which are particularly vulnerable to landsliding so as to manage human activities in, and around, them.

3. Methodology

This study aimed at developing a bagging-based ensemble model by using well-known machine learning algorithms and wrapper-based feature selection. To this end, first, the landslide inventory maps and conditioning criteria (i.e., determinants of landsliding) were prepared. Thereafter, the frequency ratio (FR) was employed to
determine the weight of each class of variables affecting landslide occurrence. Before using the models to solve the LSM problem, the proposed hybrid feature selection algorithms based on five machine learning models and GA were employed to determine the optimal determinants of land sliding. The identified determinants were then selected and employed as inputs to the single models and each of the models is constructed. Then the output of these models is combined to construct the bagging-based ensemble model. Finally, the models were all evaluated based on the area under the receiver operating characteristic curve (AUROC) criterion. Then, landslide susceptibility maps were produced. Figure 3 shows the overall process applied in this study.

### 3.1. Inventory landslide map

Generation of inventory landslide maps is the earliest and most vital step in the LSM. This study tracked and documented the landslides that took place in the study area through site visits and collection and analysis of local information, Google Earth images, and site maps, which were obtained from The Watershed Office of Jerash, Jordan. Based on this, sixty-nine landslide locations were spotted on the map of Jerash. Thereafter, data were split in a random manner into training sub-set (70.0% of the data) and validation sub-set (30.0%). For non-landslide locations, sites in the governorate of Jerash with no reported land sliding history have been chosen by the researchers, considering a number of criteria. One of these criteria is that the site has no history of landslides few decades, or centuries, earlier. Another criterion is that the site has no likelihood for landslides. Moreover, in the case when the site has soft crust (like that the bare lands and the sites that are employed for cultivation of...
seasonal crops) the sites to include in this study were selected in places with low slopes like flat lands or tops of the hills, provided they have cohesive soils. In sites with no steep slopes (namely, sites with low to moderate slopes), non-slipping locations were selected if they were covered with permanent green cover (e.g., woodlands and forests) or if they were having solid cover of rocks. Meanwhile, locations with urban activity established for decades that are entirely free from indications of landslides were too selected. On the other hand, locations with established landslide activity were selected, irrespective of their soil types and geological structures.

3.2. Landslide conditioning criteria

Land sliding in Jerash is influenced by various factors. The most important of which are: (i) extending the transportation networks without studying geological structures of the areas in which the new roads will be created and the areas wherein the roads are to be widened; (ii) extending tourist infrastructure and facilities such as restaurants and vehicle parks in scenic areas (e.g., edges of steep slopes and summits of mountains) without considering soil and geological structure types for the new constructions themselves and for the new roads leading to them; (iii) performing construction activities in slums prevailing on village outskirts and countryside, where monitoring by the local municipalities is normally less regular and efficient than monitoring in urban centers. Usually, the people implementing these works do not take precautions against geologic threats to construction; and (iv) the agricultural works, especially in private farms, which do not take into account factors leading, or contributing, to sliding of land such as ploughing and repeated, heavy irrigation. Based on a review of the literature, the available data, and opinions of experts, sixteen landslide conditioning criteria (that is, determinants or influential variables) were selected, being highly relevant for LSM. These are aspect, distance to drainage, distance to faults, distance to roads, elevation, geology, land cover and use (LULC), Normalized Difference Vegetation Index (NDVI), plan curvature, profile curvature, rainfall, Sediment Transport Index (STI), slope, soil texture, Stream Power Index (SPI), and Topographic Wetness Index (TWI). Maps for these 16 predictors were, then, prepared in the Geographic Information System (GIS) environment. The selected factors were not all used in the modeling. Rather, they were filtered for

| Layers                      | Scale      | Source                                      |
|-----------------------------|------------|---------------------------------------------|
| Rainfall                    | 1:250,000  | Higher Council for Science and Technology (2007) |
| Drainage (Wadi)             | 1:250,000  | Royal Jordanian Geographic Centre (1995)    |
| Roads                       | 1:250,000  |                                             |
| Geology                     | 1:250,000  |                                             |
| Fault                       | 1:250,000  |                                             |
| Soil                        | 1:750,000  | Jordan Ministry of Agriculture (1993)       |
| Wells                       | Well Data (Excel File) | Water Authority of Jordan (2012) |
| Slope, elevation and aspect based on ASTER DEM | 30 m | United States Geological Survey (USGS) |
| NDVI based on Landsat 8 images | 30 m | USGS |
| Land use based on Landsat TM. | 30 m | USGS |
| Data 30 m                   |            |                                             |
modeling after careful feature selection. Table 1 shows the source and resolution/scale of the spatial layers.

3.2.1. Elevation
Elevation is the height above or below a reference point. Usually, the values of elevation are extracted from Digital Elevation Model (DEM). As Figure 4a indicates, the elevation values range in the study area from 2 m to 1229 m.
3.2.2. Aspect
Aspect is the steepest downhill direction (Kadavi et al. 2018). It affects shear strength of the sliding material and influences rainfall distribution, being associated with the physiographic properties of the area (Jebur et al. 2014).

3.2.3. Slope
Slope helps in controlling the moisture content and the level of pore pressure of the soil. An increase in the angle of inclination leads to instability of the slope and, consequently, to an increase in the likelihood for landslides to occur (Bahrami et al. 2020). The slope angle also controls the surface water runoff, slope stress distribution, slope sediment accumulation, and soil erosion (Fang et al. 2021). As shown in Figure 4.
the slope angles range in the study area (the governorate of Jerash) from 0° to 30°.

3.2.4. Land use

According to Souissi et al. (2020), land use influences hydrological processes such as runoff and infiltration, which are processes that are much affected by presence and distribution of bare lands, buildings, flow paths of streams, vegetation cover, roads, etc. Land cover and land use (LULC) may have negative and positive roles in the susceptibility of land to the slides. For example, built-up structures result in stabilization of slopes. However, non-expert building activities at slopes compromise slope stability and trigger landslides (Bahrami et al. 2020). As regards the study area, it is seen in Figure 4a that nine land covers and land use types prevail in Jerash.
3.2.5. Geology

Geology of the area influences the shear strength of the rock and indicates the probability of increase in neutral pressure in the sub-soil (Jebur et al. 2014). Figure 4b uncovers that there are five types of geology in the study area. During the Mesozoic and the Early Cenozoic era, sedimentation was actually controlled by the fluctuational movement of Tethys Ocean, northwest of Jordan, and the eustatic movement of the Arabian–Nubian Shield and, thus, its Palaeozoic rocks in the south of the country. The study governorate is characterized by exposures of sedimentary rocks that mainly include chalk, chert limestone, marl, and sandstone of the cretaceous and Jurassic ages (Table 2). Because of the pre-Jurassic tectonic movements, the Azab Group, which constitutes the oldest layers that are exposed in the area under study, overlies
the Triassic sequence. This group is covered by the Early Cretaceous Kurnub Sandstone (KS) Group that is overlain by thick sequence of marls and limestone of Ajlun (Cenomanian–Turonian) Group. This group of rock formations is divided into five major formations: Wadi As Sir Limestone (A7), Shuyab Limestone (A5/6), Hummar Limestone (A4), Fuheis Limestone (A3), and Na’ur Limestone (A1/2). Vast segments of the northern parts of Ajloun are covered with three formations of the Belqa Group, namely, Wadi Umm Ghudran Chalk (B1), Amman Silicified Limestone (B2a), and Al Hisa Phosphorite (B2b), which is a Coniacian–Campanian formation. Limited outcrops of Muwaqqar Chalk Marl formation (B3), which is Maastrichtian–Palaeocene formation, are found in the northwestern sections of Ajloun. In this governorate, Quaternary sediments are mainly plateau gravels, soil, wadi sediments, calcrete, and alluvium. On the other hand, two principal fault patterns, of the Late Tertiary age, presumably, characterize the geologic structure of this governorate: east–west and northeast–southwest faults. Accordingly, most of the folds in this governorate are either east–northeast folds or northeast to west–northwest folds.

3.2.6. Distance to roads
The values of distance to roads ranged in the current study area from 0 m to greater than 600 m (Figure 4b).

3.2.7. Rainfall
Rainfall influences the saturation level of the soil and the amount of runoff through it and raises the water pressure in the soil pores (Bahrami et al. 2020). With respect to Jerash, Figure 4b points out that the rainfall ranged in this governorate from 250 mm to 550 mm.

3.2.8. Topographic wetness index (TWI)
Since the TWI is hydrological variable, then it gives quantitative assessment of moisture condition of soil, which is an essential contributor to landslide incidence (Fang et al. 2021) that is widely used in representation of the influence of topography on them. Mathematically, the TWI is given by Eq. (1):

$$TWI = \ln \frac{A_S}{\tan \beta}$$

In this equation, the variable $A_S$ is specific catchment area whereas $\beta$ is slope gradient in degrees radian (Jebur et al. 2014). In this study, values of TWI were classified into three intervals: 2–9, 9–13, and 13–19 (Figure 4b).
### 3.2.9. Distance to faults
The distance to faults is the maximum distance from faults. In the case of the present study area, Jerash, the distances to faults fell in the range of 0–11,000 m (Figure 4c).

### 3.2.10. Plan curvature
Plan curvature is perpendicular to the slope direction (Fang et al. 2021). As can be noticed in Figure 4c, the plan curvature values in the governorate of Jerash have been grouped into three classes: $<-0.05$, $-0.05$ to 0.05, $>0.05$.

### 3.2.11. NDVI
The NDVI is an index commonly employed to outline vegetative cover (Kadavi et al. 2018). The NDVI is an indicator of vegetation intensity (and, therefore, biomass) and distribution. Yang et al. (2012) clarified that changes to vegetation in the vegetated areas can result in failure of slopes. They also indicated that drops in values of the NDVI, reflecting disturbance in vegetation, within short periods of time can be employed for quick identification of areas with high potential for landslides. In the current study, an NDVI map for the study area was created from Landsat 8 images in the ENVI 5.1 software environment. The resulting NDVI values were categorized into six groups: (i) $-0.46$ to $-0.015$, (ii) $-0.015$ to $-0.08$, (iii) $-0.08$ to $-0.03$, (iv) $-0.03$ to $0.01$, (v) $0.01$ to $0.05$, and (vi) $0.05$ to $0.28$ (Figure 3c). These values were obtained using Eq. (2):

$$\text{NDVI} = \frac{(\text{IR} - R)}{\text{IR} + R}$$

In this equation, the acronyms R and IR stand, respectively, for intensities of energies that are reflected in the red and infrared regions of the spectrum of electromagnetic radiation (Pradhan et al. 2010). Figure 4c unveils that the NDVI values varied in the study area from $-0.35$ to $0.20$. According to Yang et al. (2012), analyzing the disturbances on vegetation detected from the abnormal sudden drops of the normalized difference vegetation index (NDVI) within a short period can be used for the purpose of rapid landslide identification.

### 3.2.12. Profile curvature
The profile curvature normally lies parallel to the direction of slope. It represents the amount of deformation of the slope surface and contributes to generation of landslides (Fang et al. 2021). For the study area, the profile curvature values were divided into three categories: $<-1$, $-1$ to 0.2, and $>0.2$ as shown in Figure 4c.

### 3.2.13. Distance to drainage
Values of distance to drainage have the range of 0 to $>600$ m in the area of study (Figure 3d). These values have been categorized into five intervals as shown in Figure 4d.
3.2.14. Stream power index (SPI)

The SPI expresses stream erosive power (Fang et al. 2021). It can be calculated according to Eq. (3):

$$\text{SPI} = A_S \tan \beta$$

where \( A_S \) is the specific catchment area and \( \beta \) is the slope gradient (Jebur et al. 2014). As can be seen in Figure 4d, Jerash has SPI values ranging from 1.54 to higher than 11.08.

3.2.15. Soil texture

Figure 4d reveals that four major textures characterize soils of the study area: Silty clay loam, loam, silty loam, and clay loam.

3.2.16. Sediment transport index (STI)

The STI describes slope failure mechanism. It is measure of amount of sediment that is, or can, be transported (Fang et al. 2021) and is computed using Eq. (4):

$$\left( \frac{A_S}{22.13} \right)^{0.6} \left( \frac{\sin \beta}{0.0896} \right)^{1.3}$$

where \( \beta \) is the slope of each pixel and \( A_S \) is the upstream area (Jebur et al. 2014). The values of STI in Jerash had the range of 0 to > 80 (Figure 4d).

3.3. Correlations among landslides and their determinant factors

The frequency ratio (FR) was used to investigate probable relationships between landslide determinant factors (variables) and landslide incidence. The FR estimates the likelihood of the presence of a phenomenon in a class of a variable (Lee and Sambath 2006). Stated otherwise, the FR quantifies associations of every group of variables and the locations of landslides. Mathematically, it assumes the formula (Lee and Sambath 2006):

$$\text{FR} = \frac{A}{B} \frac{C}{D}$$

In this formula (Eq. 5), the acronym A describes number of pixels of landslides in the particular class, B denotes overall number of occurring landslides, C indicates area of class, and D expresses area of the studied district or province.

3.4. Feature selection using genetic algorithm (GA)

Since a better input than otherwise will lead to a less complex and more accurate model, selecting the appropriate features to be used as predictors has always been one of the main concerns of modeling algorithms (Bouhamed et al. 2015). Within this context, use of the correlations amongst the predictor variables and the target variable
is a common approach to feature selection. However, this method can only show the effect of a single predictor on the target variable, but a combination of multiple variables with high correlation coefficients may not always be as much effective as expected (Guyon and Elisseeff 2003; Liu et al. 2005). In this study, we used wrapper-based feature selection that combine the GA with the DT, kNN, NB, ELM, and SVM methods in an effort to identify, and, then, select the best features for each model from the pool of available input features. By doing so, the effects of multiple variables are considered simultaneously and, thereupon, the problem of feature selection is eliminated. However, in general, a complete search of the feature space that examines all likely feature combinations is time consuming and it can be infeasible once the number of input features exceeds a certain limit.

In the feature selection method proposed in this paper, which is known as the wrapper approach, first, a random initial population of solutions is generated (Rodrigues et al. 2014), where each chromosome represents a feature set, and each gene designates a single feature. In a single chromosome, if a gene has the value of 0, then the corresponding variable is not part of the final feature set. On the other hand, if a gene has the value of 1, then the corresponding feature is present in the final feature set. At the beginning of implementation of the algorithm, these genes are randomly initialized. Once the initial population is generated, the feature set concomitant with each chromosome is fed to the DT, NB, ELM, SVM, and kNN algorithms, and the modeling is carried out using the training dataset. Once the models are created, the model-predicted values are compared with the true values and the root mean-squared error (RMSE) is calculated. That is, the RMSE serves as fitness function in GA and its values are utilized as foundation, or criterion, for selection of the chromosomes by the roulette wheel method. The chromosomes with low RMSE values have a higher chance of being selected by the roulette wheel, and vice versa. When selection operator is run, the method of two-point crossover is employed to join all chosen chromosomes and generate the sequent generation of the solutions. This process will be repeated and, after a certain number of iterations, the best available feature set will be chosen for modeling.

3.5. The k-nearest neighbors (kNN)

The kNN algorithm represents non-parametric approach to supervised classification and stands as the simplest machine learning algorithm for this purpose (Kramer 2013). The foundation of this algorithm is the classification of any new data according to the most similar earlier data (Kramer 2013). For this purpose, the $k$ parameter, that is, number of nearest neighbors to take into account, is specified. After that, the $k$ samples of data which are most identical to the sample of interest are chosen on the basis of distances of data points from one the other (Kramer 2013). The Euclidean distance (Eq. 6), Manhattan distance, and Minkowski distance are among the distance metrics that are typically used for this purpose (Kramer 2013).
Euclidean Distance \[= \sqrt{\sum_{i=1}^{N} (p_i - q_i)^2} \] (6)

Once the nearest neighbors of the desired data sample are determined, the number of samples of each class in the neighborhood is determined and the class with the majority vote is considered as the class of the new data sample. The result of the algorithm is highly dependent on the value of \(k\) because the boundaries of the classes vary as the value of \(k\) varies. Hence, specification of the best \(k\) value is highly critical. Furthermore, the kNN algorithm is sensitive to the scale of the variables. Thus, data normalization can be helpful. However, the most prominent problem of the kNN algorithm is its high computation complexity and execution time, which is a result of the need for re-processing the entire training dataset whenever a new data sample is classified.

3.6. Naïve Bayes (NB) algorithm

The NB algorithm is a machine learning algorithm that employs statistical methods to classify data (Leung 2007). This algorithm is based on the Bayes theorem (Kelly and Kolstad 1999):

\[P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \] (7)

Given a test sample of data \(X = \{x_1, x_2, ..., x_n\}\) having \(n\) features. At first, we assume that the features are independent variables, i.e., each feature participates independently in determination of the probability of the data belonging to each class. This assumption simplifies the calculations. The NB algorithm determines the probability for the data sample to belong to each class in the set \(C = (C_1, C_2, ..., C_K)\) to predict its class. The algorithm computes the probability of occurrence or non-occurrence of any of the cases using the Bayes theorem (Kelly and Kolstad 1999):

\[P(C_k|X) = \frac{P(X|C_k) \cdot P(C_k)}{P(X)} \] (8)

However, the independence assumption leads to:

\[P(X|C_k) = \prod_{j=1}^{n} P(x_j|C_k) \] (9)

The data sample is, accordingly, classified into the class with the highest probability obtained using the training dataset.

\[\hat{y} = \arg\max_{k=1,...,K} P(C_k|X) \] (10)
This algorithm is suitable for handling multi-class problems and it does not require a lot of training data. Additionally, it has a high processing speed and it produces classifications of good accuracy (Kelly and Kolstad 1999).

3.7. Decision tree (DT)

The DT represents one of the machine learning methods which may be employed in classification and regression problems (Kotsiantis 2013). The DT comprises three parts: root node, internal nodes, and leaf nodes (Kotsiantis 2013). The data features are explored in the root node and each internal node. Then, the most discriminating data feature is selected for partitioning the data (Kotsiantis 2013). The Entropy (Eq. 11) or Gini (Eq. 12) criteria are often employed for partitioning data (Breiman et al. 1984). These two metrics measure the impurity of the node. The feature with the lowest Entropy or Gini value provides the highest amount of information for classification and it is the best choice for the purpose of data partitioning (Breiman et al. 1984).

\[
\text{Entropy} = -\sum_j p_j \log_2 p_j
\]  
\[\text{Gini} = 1 - \sum_j p_j^2\] (11) (12)

These two metrics are employed to disclose the optimal feature for partitioning the data. However, Entropy has somehow better accuracy performance than Gini. On the other hand, the Gini metric involves less computation and it is faster than Entropy (Breiman et al. 1984). Once the best feature is determined, the data is partitioned based on it. The DT grows as this process continues. In the end, each leaf node is labeled with a class. The chance of over-fitting is relatively high in the decision trees. Early stopping on the basis of such measures as maximal tree depth or lowest number of samples in the nodes of the leaf is an approach that is quite often employed for avoiding this sort of problem (Breiman et al. 1984). Interpretability represents one of the major favorable characteristics of this kind of model.

3.8. Extreme learning machine (ELM)

Relatively, the ELM is new machine learning model, which has been applied with success in differing fields and in providing solutions to a variety of problems (Huang et al. 2006). Initially, the ELM was proposed as a learning framework for the generalized single-hidden layer feed-forward neural network (SLFN) but was later extended to other generalized SFLN models (Huang et al. 2004). The output function of a generalized SFLN can be expressed as illustrated in Eq. 13 (Huang et al. 2006):

\[
f(x) = h(x)\beta = \sum_{i=1}^{l} \beta_i h_i(a_i, b_i, x)
\]  
\[x, \ a_i \in R^2, \ b_i \in R\] (13)
where \( h_i(x) \) maps the \( d \)-dimensional input space \((x)\) to an \( l \)-dimensional space. The acronyms \( a_i \) and \( b_i \) are actually randomly-initialized variables of the \( i \)th node while \( \beta_i \) stands for weights of connections between the output and hidden layers. In the meantime, the parameter \( L \) indicates the total number of functions that need to be estimated.

For a binary classification problem, the decision function can be defined as shown in Eq. (14) (Huang et al. 2006):

\[
f(x) = \text{sign} \ (h(x)\beta)
\]

Unlike the conventional learning algorithms, the ELM not only aims at minimizing the error on the training set but also tries to minimize the norm of the output weights (Huang et al. 2006). Based on Bartlett’s Theory, minimizing the norm of weights in the SFLNs, in addition to minimizing the training error, results in better generalization (Huang et al. 2004). Thus, the objective function of ELM is formulated according to Eq. (15) to reduce the training error and the norm of the output weights (Huang et al. 2004).

\[
\text{Minimize : } \|H\beta - T\|^2 + ||\beta||
\]

where \( T \) is the target matrix \([t_1, t_2, \ldots, t_N]^T\), \( ||\cdot|| \) indicates the standard norm, and \( H \) is the matrix of the hidden layer, which is defined according to Eq. (16) (Huang et al. 2004):

\[
H = \begin{bmatrix}
h(x_1) \\
\vdots \\
h(x_N)
\end{bmatrix} = \begin{bmatrix}
G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_L) \\
\vdots & \ddots & \vdots \\
G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N)
\end{bmatrix}
\]

where \( G(a_i, b_i, x_i) \) expresses the relationship between the \( i \)th variable and its weights.

As can be observed, minimizing \( \|\beta\| \) is the same as maximizing \( 2\|\beta\|^{-1} \), which is equivalent to maximizing the margin of the two classes in a binary classification problem. To improve the learning model and extend it and to avoid projection, which is not always obvious for the designer, the kernel-based model is used. In this case, the ELM kernel is defined based on Eqs. (17) and (18) (Huang et al. 2004):

\[
\Omega_{ELM} = HH^T : \Omega_{ELM} \ i,j = h(x_i)h(x_j) = K(x_i, x_j)
\]

\[
f(x) = h(x)H^T\left(\frac{I}{C} + HH^T\right)^{-1}T = \begin{bmatrix}
K(x, x_1) \\
\vdots \\
K(x, x_N)
\end{bmatrix}^T \left(\frac{I}{C} + \Omega_{ELM}\right)^{-1}T
\]

where \( T \) denotes the transpose of a vector and \( K(x_i, x_j) \) represents the relationship of the \( x_i \) and \( x_j \) variables in terms of a function \( K \). Meanwhile, the parameter \( I \) is the identity matrix with appropriate dimensions and \( C \in \mathbb{R}^+ \) is a tuning parameter.
3.9. Support vector machine (SVM)

Based on the statistical learning theory, the SVM is one of the widely-used machine learning methods. In this method, the features or factors are used along with an optimization algorithm to identify the samples that constitute class boundaries. They are also employed to create an optimal, linear decision-making boundary to separate classes. These samples are called support vectors (Vapnik et al. 1995). The optimal margin method is adopted to determine the decision-making boundary between two completely separate classes (Vapnik et al. 1995). Using kernel functions, this model transfers data to a space with higher dimensions, where data can be separated through a linear discriminating function (Patle and Chouhan 2013). The model can be defined using Eq. (19) by considering datasets \( \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\} \) in which \( x_i \) refers to the samples whereas \( y_i \) is the label or class of each sample (Patle and Chouhan 2013).

\[
\max \Omega(x) = \sum_{i=1}^{n} x_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j y_i y_j K(x_i, x_j) \\
\text{s.t.} \sum_{i=1}^{n} x_i y_i = 0, \ x_i \in [0, C], \ i = 1, 2, \ldots, n
\]

(19)

where \( x_i \), \( K \), and \( C \) refer to the Lagrangian coefficients, a kernel function, and a constant, respectively. The value of \( x_i \) can be obtained by solving Eq. (19). The support vectors (s) are the internal values \( (0, C) \) and the constant that is called the bias \( (b) \) is determined by using these vectors. After these parameters are calculated, the decision-making function is obtained from Eq. (20) (Patle and Chouhan 2013):

\[
f(x) = \text{sgn} \left( \sum_{x_i \in S} x_i y_i K(x_i, x) + b \right)
\]

(20)

Equation (20) shows the kernel function \( K \) to be having different types of functions such as the linear, sigmoid, polynomial, and radial basis functions (RBF). This study used the kernel of the RBF due to its good performance.

3.10. Suggested bagging ensemble model

Bagging stands for Bootstrap Aggregation. It is quite robust ensemble machine learning technique that may be employed in classification and regression problems (Breiman 1996). It is usually used in combination with non-robust and overfit-prone models such as the DT (Breiman 1996). Using this method reduces the variance, increases the robustness, and enhances the overall prediction accuracy of the model.

In this algorithm, once the base learning model is determined, bootstrapping and aggregation are performed. In the bootstrapping phase, multiple sub-sets of the training dataset are created using random sampling with-replacement sampling. With-replacement sampling ensures that the probability for a data sample to be selected for a new sub-set is equal for all data samples and that creation of each new sub-set is
Table 3. Relations among landslide determinants and locations in Jerash.

| Criterion                  | Classes                  | No. of pixels | No. of landslides | FR |
|----------------------------|--------------------------|---------------|-------------------|----|
| Profile curvature          | < -1                     | 17,440        | 4                 | 1.49 |
|                            | -1–0.20                  | 305,967       | 44                | 0.93 |
|                            | > 0.20                   | 124,264       | 21                | 1.10 |
| Distance to faults (m)     | 0–2,000                  | 43,088        | 10                | 1.50 |
|                            | 2,000–5,000              | 141,056       | 24                | 1.10 |
|                            | 5,000–8,000              | 105,579       | 10                | 0.61 |
|                            | 8,000–11,000             | 81,222        | 12                | 0.96 |
|                            | > 110,000                | 76,726        | 13                | 1.10 |
| Distance to drainage (m)   | 0–150                    | 13,588        | 27                | 1.30 |
|                            | 150–300                  | 122,197       | 24                | 1.27 |
|                            | 300–450                  | 97,285        | 13                | 0.87 |
|                            | 450–600                  | 63,439        | 4                 | 0.41 |
|                            | > 600                    | 28,869        | 1                 | 0.22 |
| Land cover and use (LULC)  | Desert                   | 162,075       | 14                | 0.56 |
|                            | Vegetation               | 258,858       | 47                | 1.18 |
|                            | Urban                    | 1,617         | 0                 | 0   |
|                            | Rural                    | 19,869        | 7                 | 2.29 |
|                            | Soil rocks               | 3,558         | 1                 | 1.85 |
|                            | Rocks                    | 642           | 0                 | 0   |
|                            | Valley                   | 884           | 0                 | 0   |
|                            | Sand                     | 730           | 0                 | 0   |
| Rainfall depth (mm)        | 250–300                  | 790           | 0                 | 0   |
|                            | 300–350                  | 113,431       | 10                | 0.57 |
|                            | 350–400                  | 154,134       | 34                | 1.43 |
|                            | 400–450                  | 145,346       | 23                | 1.03 |
|                            | 450–500                  | 33,970        | 2                 | 0.38 |
| TWI                        | < 9                      | 121,014       | 21                | 1.12 |
|                            | 9–13                     | 244,814       | 36                | 0.95 |
|                            | > 13                     | 81,843        | 12                | 0.95 |
| SPI                        | < 100                    | 121,014       | 21                | 1.13 |
|                            | 100–200                  | 643           | 0                 | 0   |
|                            | 200–300                  | 15,181        | 3                 | 1.28 |
|                            | 300–400                  | 27,571        | 3                 | 0.71 |
|                            | > 400                    | 283,262       | 42                | 0.96 |
| STI                        | 0–565                    | 123,515       | 21                | 1.10 |
|                            | 565–3108                 | 31            | 0                 | 0   |
|                            | > 3108                   | 324,125       | 48                | 0.96 |
| Aspect                     | Flat                     | 3,188         | 0                 | 0   |
|                            | North                    | 45,797        | 14                | 1.98 |
|                            | North east               | 40,583        | 3                 | 0.48 |
|                            | East                     | 57,843        | 5                 | 0.56 |
|                            | Southeast                | 68,758        | 10                | 0.94 |
|                            | South                    | 86,440        | 6                 | 0.45 |
|                            | Southeast                | 55,624        | 10                | 1.17 |
|                            | West                     | 52,370        | 15                | 1.86 |
|                            | Northwest                | 37,368        | 6                 | 1.04 |
| Slope                      | 0–5                      | 81,220        | 12                | 0.96 |
|                            | 5–15                     | 267,458       | 40                | 0.97 |
|                            | 15–30                    | 92,320        | 16                | 1.13 |
|                            | > 30                     | 6,973         | 1                 | 0.93 |
| Plan curvature             | < -0.05                  | 171,499       | 24                | 0.91 |
|                            | -0.05 to 0.05            | 96,498        | 13                | 0.87 |
|                            | > 0.05                   | 179,674       | 32                | 1.15 |
| NDVI                       | -0.47 to -0.15           | 5,061         | 0                 | 0   |
|                            | -0.15 to -0.08           | 13,580        | 0                 | 0   |
|                            | -0.08 to -0.03           | 35,929        | 0                 | 0   |
|                            | -0.03 to -0.01           | 91,890        | 6                 | 0.42 |
|                            | 0.01–0.05                | 158,552       | 30                | 1.23 |
|                            | 0.05–0.28                | 142,959       | 33                | 1.50 |
| Geology                    | Sandstone                | 96,729        | 24                | 1.61 |
|                            | Limestone, dolomitic limestone | 114,624 | 27                | 1.53 |

(continued)
independent of selection of the others (Breiman 1996). The base model is not robust. Therefore, using the same training algorithm with different datasets results in different classifiers. Since sampling of each sub-set is independent of sampling of the others, then the models can be trained independently and in parallel. In the aggregation phase, the results of different classifiers are combined in order to provide the final prediction. In regression problems, the final prediction is determined by averaging the results of the trained models (Breiman 1996) while in the classification problems the majority voting approach is used to combine the results.

4. Results

The FR was employed in the present study to determine the weight of each class of variables affecting landslide occurrence. Table 3 displays the most and the least important classes of each factor. In elevation factor, the class 318–533 m had the highest weight in the FR model. In aspect factor, all classes almost contribute to the landslide occurrence and only class flat has no effect on the landslide occurrence. For the slope factor, all classes contribute to the landslide occurrence. Landslides were more common in rainfall classes 350–400 mm and 400–450 mm. Furthermore, the highest weights for other factors were achieved by SPI between 200 and 300, TWI < 9, STI between 0 and 565, distance to drainage between 0 and 150 m, distance to fault between 0 and 2000 m, NDVI between 0.05 and 0.28, Sandstone class from geology factor, and Soil rocks from land use factor. After determining the weights for each class, these weights were assigned to the classes.

To eliminate the co-linearity between the variables and remove the unnecessary factors, a combination of GA and five machine learning models based on wrapper were used for feature selection of the conditioning factors. When running the feature selection algorithms in the current study, the maximum iterations and the number of chromosomes in each generation were set to 200 and 20, respectively. Figures 5–9 show convergence of the fitness function in the feature selection algorithms. Figure 5 points out that mean RMSE of GA-DT model has started with 0.14, and after
running the algorithm for 200 generations reaches to 0.0093. It should be noted that the optimal value for this fitness function is zero. Meanwhile, Figure 6 indicates that the mean and best values of fitness for the GA-ELM were 0.1516 and 0.1096, respectively. For the other three remaining algorithms (Figures 7–9), the mean and the best fitness values were respectively 0.3361 and 0.3344 for the GA-kNN, 0.4449 and 0.4415 for the GA-NB, and 0.4039 and 0.4012 for the GA-SVM.

The results of feature selection using GA-kNN, GA-NB, GA-SVM, GA-ELM, and GA-DT algorithms are provided in Table 4. As this table reveals, the elevation was not among the selected features for any of the models. This is in line with the study of (Micheletti et al. 2014) in Plateau, Switzerland. The aspect and geology factors were among the influential variables for all models, and play a crucial role in LSM.
In addition, in four of the five basic models for feature selection, the slope, NDVI, SPI, distance to roads, and rainfall depth were among the relevant variables.

After choosing the optimal features for each of the five models, namely DT, SVM, kNN, ELM, and NB by wrapper-based feature selection based on GA, the modeling was done using these single models separately. Then the output of these models was combined by the Bagging approach and an ensemble model was created. In fact, this ensemble model is a combination of five machine learning models (formed by the factors selected by feature selection based on the wrapper approach) and bagging approach.

In the LSM task, the objective is identification of the areas with the high likelihood for incidence of landslide. Therefore, irrespective of the approach followed in order to generate landslide susceptibility map, what matters the most is, in effect, reliability.
of the approach. We used the AUROC metric to evaluate the various aforementioned models. Figures 10 and 11 present AUROC values for the six developed models in train and test phases. These figures show that Bagging-based ensemble model with the AUROC of 0.99 and 0.85 has the highest accuracy in training and validation phases, respectively. The results also indicate that the ELM model was over-fitted (Figure 10). For this model, the AUROC was 0.99 in training and 0.43 in validation.

When the models were constructed and evaluated, all cells of the study area along with their predictor factors were entered into the models to calculate likelihood of incidence of landslides. Then, these outputs were employed in creation of landslide susceptibility maps.

Figure 12 shows the landslide susceptibility maps generated by the five algorithms. These maps were classified into five classes namely, very low, low, moderate, high and very high susceptibility based on natural break method. The outputs of the
mapping reveal that high-risk locations are mostly found in the east, south, and north of Jerash (Figure 12); the western areas of this governorate prove to be less susceptible to landslides.

5. Discussion

Landslide susceptibility mapping is of great help for identification of instable areas. This study developed five models for LSM in Jerash; five individual models and novel Ensemble model. Those models were developed by using 16 predictor variables and five feature selection algorithms that are employed for the first time to specify the
most effective features for prediction of land susceptibility to sliding. Seven factors, namely, aspect, slope, geology, NDVI, SPI, distance to road, and rainfall depth were identified by most of the feature selection algorithms employed in this study as the landslide determinant factors. The outcomes of this mapping can be employed as foundation for planning a variety of activities that encircle slope stabilization and construction works.

The results of this study provide evidence on that using ensemble models can result in reliable modeling and produce robust results. This is partly due to employment of multiple single models for developing integrated ensemble model. In the ensemble model, the class label of a sample is decided based on the outputs of several single models. A number of previous studies conducted in various areas around the
world have reported similar results and shown that, in comparison with single models, the ensemble models often achieve better results. Naghibi et al. (2016) used an ensemble model based on an evidential belief function and a set of tree-based models for spatial modeling of groundwater potential. Their results indicate that the ensemble model achieves better performance than single models, which is a result with which our findings agree (Arabameri et al. 2020) used a combination of expert knowledge-based (AHP), bivariate (statistical index, SI) and multivariate (linear discriminant analysis) models as an ensemble model for landslide susceptibility mapping (LSM) in Mazandaran Province, Iran. They found that the AHP-SI ensemble model achieved the highest accuracy (Youssef et al. 2015) used an ensemble method of frequency ratio and logistic regression for LSM in Fayfa area which is located in the southwest of Saudi Arabia (SA) in Jazan region. They found that the proposed ensemble model outperformed other single models. (Di Napoli et al. 2020) used the ensemble of artificial neural network, generalized boosting model and maximum entropy for LSM in the Monterosso al Mare area, Cinque Terre National Park (Northern Italy). Their results show that proposed ensemble model has good performance for LSM.

6. Conclusion

Landslide susceptibility mapping has many applications in the infrastructural activities of establishing and developing structural, agricultural, and natural resource operations. This study developed a new ensemble model that is based on the bagging approach and which was constructed from five individual single models for mapping landslide susceptibility in Jerash, Jordan. The modeling was performed in two phases. In the first phase, we used five wrapper-based feature selection algorithms to select the most important predictor variables for each model. These algorithms combine the
GA with five basic machine learning models, that is, kNN, SVM, NB, DT, and ELM. The results of this phase have shown that aspect, geology, slope, NDVI, SPI, distance to roads, and rainfall depth are the seven most important determinants of landslides in the studied governorate. These seven variables were employed in LSM using the five aforementioned single models. Then, these models were used to create an ensemble model. The modeling results prove that the ensemble model is robust, both in the training and the validation phases. Thus, researchers and land use and urban development planners can use this model in their studies.

References

Abedini M, Tulabi S. 2018. Assessing LNRF, FR, and AHP models in landslide susceptibility mapping index: a comparative study of Nojian watershed in Lorestan province, Iran. Environ Earth Sci. 77(11):1–13.

Akinci H, Zeybek M, Dogan S. 2021. Evaluation of landslide susceptibility of Şavşat District of Artvin Province (Turkey) using machine learning techniques. In Landslides. IntechOpen.

Akinci H, Zeybek M. 2021. Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin.). Nat Hazards. 108(2):1515–1543.

Ali SA, Parvin F, Vojteková J, Costache R, Linh NTT, Pham QB, Vojtek M, Gigovic L, Ahmad A, Ghorbani MA. 2021. GIS-based landslide susceptibility modeling: a comparison between fuzzy multi-criteria and machine learning algorithms. Geosci Front. 12(2):857–876.

Arabameri A, Pradhan B, Rezaei K, Lee S, Sohrabi M. 2020. An ensemble model for landslide susceptibility mapping in a forested area. Geocarto Int. 35(15):1680–1705.

Arca D, Kutoğlu HŞ, Becek K. 2018. Landslide susceptibility mapping in an area of underground mining using the multicriteria decision analysis method. Environ Monit Assess. 190(12):1–14.

Bahrami Y, Hassani H, Maghsoudi A. 2020. Landslide susceptibility mapping using AHP and fuzzy methods in the Gilan province, Iran. Geol J. 86(4):1797–1816.

Bouhamed SA, Kallel IK, Masmoudi DS, Solaiman B. 2015. Feature selection in possibilistic modeling. Pattern Recognit. 48(11):3627–3640.

Breiman L. 1996. Bagging predictors. Mach Learn. 24(2):123–140.

Breiman L, Friedman J, Stone CJ, Olshen RA. 1984. Classification and regression trees. Boca Raton: CRC Press.

Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB. 2012. Landslide susceptibility assessment in the Hoa Binh province of Vietnam: a comparison of the Levenberg–Marquardt and Bayesian regularized neural networks. Geomorphology. 171:12–29.

Chen W, Xie X, Peng J, Wang J, Duan Z, Hong H. 2017. GIS-based landslide susceptibility modelling: a comparative assessment of kernel logistic regression, Naïve–Bayes tree, and alternating decision tree models. Geomat Nat Hazards Risk. 8(2):950–973.

Chen X, Chen W. 2021. GIS-based landslide susceptibility assessment using optimized hybrid machine learning methods. Catena. 196:104833.

Di Napoli M, Carotenuto F, Cevasco A, Confuorto P, Di Martire D, Firpo M, Pepe G, Raso E, Calcaterra D. 2020. Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. Landslides. 17(8):1897–1914.

Fang Z, Wang Y, Peng L, Hong H. 2021. A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. Int J Geogr Inform Sci. 35(2):321–347.

Fell R, Corominas J, Bonnard C, Cascini L, Leroi E, Savage WZ. 2008. Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning. Eng Geol. 102(3–4):99–111.

Goetz JN, Brenning A, Petschko H, Leopold P. 2015. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput Geosci. 81:1–11.
Guyon I, Elisseeff A. 2003. An introduction to variable and feature selection. J Mach Learn Res. 3(Mar):1157–1182.
He Q, Shahabi H, Shirzadi A, Li S, Chen W, Wang N, Chai H, Bian H, Ma J, Chen Y, et al. 2019. Landslide spatial modelling using novel bivariate statistical based Naïve Bayes, RBF classifier, and RBF network machine learning algorithms. Sci Tot Environ. 663:1–15.
Huang G-B, Zhu Q-Y, Siew C-K. 2004. Extreme learning machine: a new learning scheme of feedforward neural networks. Paper presented at the 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541).
Huang G-B, Zhu Q-Y, Siew C-K. 2006. Extreme learning machine: theory and applications. Neurocomputing. 70(1–3):489–501.
Hungr O, Leroueil S, Picarelli L. 2014. The Varnes classification of landslide types, an update. Landslides. 11(2):167–194.
Jebur MN, Pradhan B, Tehrany MS. 2014. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. Remote Sens Environ. 152:150–165.
Kadavi PR, Lee C-W, Lee S. 2018. Application of ensemble-based machine learning models to landslide susceptibility mapping. Remote Sens. 10(8):1252.
Kelly DL, Kolstad CD. 1999. Bayesian learning, growth, and pollution. J Econ Dyn Control. 23(4):491–518.
Knapen A, Kitutu MG, Poesen J, Breugelmans W, Deckers J, Muwanga A. 2006. Landslides in a densely populated county at the footslopes of Mount Elgon (Uganda): characteristics and causal factors. Geomorphology. 73(1–2):149–165.
Kotsiantis SB. 2013. Decision trees: a recent overview. Artif Intell Rev. 39(4):261–283.
Kramer O. 2013. K-nearest neighbors dimensionality reduction with unsupervised nearest neighbors. Berlin: Springer; pp. 13–23.
Lee S, Sambath T. 2006. Landslide susceptibility mapping in the DamreiRomel area, Cambodia using frequency ratio and logistic regression models. Environ Geol. 50(6):847–855.
Leung KM. 2007. Naive Bayesian classifier. Polytechnic University Department of Computer Science/Finance and Risk Engineering; p. 123–156.
Li Y, Chen W. 2019. Landslide susceptibility evaluation using hybrid integration of evidential belief function and machine learning techniques. Water. 12(1):113.
Lin Q, Wang Y. 2018. Spatial and temporal analysis of a fatal landslide inventory in China from 1950 to 2016. Landslides. 15(12):2357–2372.
Liu H, Dougherty ER, Dy JG, Torkkola K, Tuv E, Peng H, Ding C, Long F, Berens M, Parsons L, et al. 2005. Evolving feature selection. IEEE Intell Syst. 20(6):64–76.
Mabdeh AN, Ahmadiou M, Al-Adamat R, Al-Shabeeb AR. 2022. GIS-based landslide susceptibility assessment and mapping in Ajloun and Jerash governorates in Jordan using genetic algorithm-based ensemble models. ActaGeophysica. 70(3):1253–1267.
Mallick J, Singh RK, ALAwadh MA, Islam S, Khan RA, Qureshi MN. 2018. GIS-based landslide susceptibility evaluation using fuzzy-AHP multi-criteria decision-making techniques in the Abha Watershed, Saudi Arabia. Environ Earth Sci. 77(7):1–25.
Mandal B, Mandal S. 2018. Analytical hierarchy process (AHP) based landslide susceptibility mapping of Lish river basin of eastern Darjeeling Himalaya, India. Adv Space Res. 62(11):3114–3132.
Manohar DR, Anbazhagan P. 2021. Shear strength characteristics of geosynthetic reinforced rubber-sand mixtures. Geotext Geomembr. 49(4):910–920.
Malkawi AH, Saleh B, Al-Sheriadeh MS, Hamza MS. 2000. Mapping of landslide hazard zones in jordan using remote sensing and GIS. J Urban Plan Dev. 126(1):1–17.
Micheletti N, Foresti L, Robert S, Leuenberger M, Pedrazzini A, Jaboyedoff M, Kanevski M. 2014. Machine learning feature selection methods for landslide susceptibility mapping. Math Geosci. 46(1):33–57.
Meena S, Mishra B, Tavakkoli Piralilou S. 2019. A hybrid spatial multi-criteria evaluation method for mapping landslide susceptible areas in kullu valley, Himalayas. Geosciences. 9(4):156.
Naghibi SA, Pourghasemi HR, Dixon B. 2016. GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environ Monit Assess. 188(1):1–27.

Nhu V-H, Shirzadi A, Shahabi H, Singh SK, Al-Ansari N, Clague JJ, Jaafari A, Chen W, Miraki S, Dou J, et al. 2020. Shallow landslide susceptibility mapping: a comparison between logistic model tree, logistic regression, naïve bayes tree, artificial neural network, and support vector machine algorithms. IJERPH. 17(8):2749.

Paryani S, Neshat A, Javadi S, Pradhan B. 2020. Comparative performance of new hybrid ANFIS models in landslide susceptibility mapping. Nat Hazards. 103(2):1961–1988.

Patle A, Chouhan DS. 2013. SVM kernel functions for classification. Paper presented at the 2013 International Conference on Advances in Technology and Engineering (ICATE).

Pawluszek K, Borkowski A. 2017. Impact of DEM-derived factors and analytical hierarchy process on landslide susceptibility mapping in the region of Rożnów Lake, Poland. Nat Hazards. 86(2):919–952.

Petley D. 2012. Global patterns of loss of life from landslides. Geology. 40(10):927–930.

Pham BT, Phong TV, Nguyen-Thoi T, Parial K, K. Singh S, Ly H-B, Nguyen KT, Ho LS, Le HV, Prakash I. 2022. Ensemble modeling of landslide susceptibility using random subspace learner and different decision tree classifiers. Geocarto Int. 37(3):735–757.

Pham BT, Shirzadi A, Bui DT, Prakash I, Dholakia MB. 2018. A hybrid machine learning ensemble approach based on a radial basis function neural network and rotation forest for landslide susceptibility modeling: a case study in the Himalayan area, India. Int J Sediment Res. 33(2):157–170.

Pham BT, Phong TV, Nguyen-Thoi T, Trinh PT, Tran QC, Ho LS, Singh SK, Duyen TTT, Nguyen LT, Le HQ, et al. 2020. GIS-based ensemble soft computing models for landslide susceptibility mapping. Adv Space Res. 66(6):1303–1320.

Pourghasemi HR, Rahmati A. 2018. Rapid GIS-based spatial and regional modelling of landslide susceptibility using machine learning techniques in the R open source software. Catena. 162:177–192.

Pradhan B. 2013. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. Comput Geosci. 51:350–365.

Pradhan B, Sezer EA, Gokceoglu C, Buchroithner MF. 2010. Landslide susceptibility mapping by neuro-fuzzy approach in a landslide-prone area (Cameron Highlands, Malaysia). IEEE Trans Geosci Remote Sens. 48(12):4164–4177.

Rodrigues D, Pereira LA, Nakamura RY, Costa KA, Yang X-S, Souza AN, Papa JP. 2014. A wrapper approach for feature selection based on bat algorithm and optimum-plate forest. Expert Syst Appl. 41(5):2250–2258.

Roy J, Saha S, Arabameri A, Blaschke T, Bui DT. 2019. A novel ensemble approach for landslide susceptibility mapping (LSM) in Darjeeling and Kalimpong districts, West Bengal, India. Remote Sens. 11(23):2866.

Saha AK, Arora MK, Gupta RP, Virdi M, Csaplovics E. 2005. GIS-based route planning in landslide-prone areas. Int J Geogr Inform Sci. 19(10):1149–1175.

Sassa K, Wang G, Fukuoka H, Wang F, Ochiai T, Sugiyama M, Sekiguchi T. 2004. Landslide risk evaluation and hazard zoning for rapid and long-travel landslides in urban development areas. Landslides. 1(3):221–235.

Schuster RL, Fleming RW. 1986. Economic losses and fatalities due to landslides. Bull Assoc Eng Geol. 23(1):11–28.

Sevgen E, Kocaman S, Nefeslioglu HA, Gokceoglu C. 2019. A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression, ANN and random forest. Sensors. 19(18):3940.

Shafizadeh-Moghadam H, Tayyebi A, Ahmadlou M, Delavar MR, Haslanou M. 2017. Integration of genetic algorithm and multiple kernel support vector regression for modeling urban growth. Comput Environ Urban Syst. 65:28–40.
Smyth CG, Royle SA. 2000. Urban landslide hazards: incidence and causative factors in Niterói, Rio de Janeiro State, Brazil. Appl Geogr. 20(2):95–118.
Souissi D, Zouhri L, Hammami S, Msaddek MH, Zghibi A, Dlala M. 2020. GIS-based MCDM–AHP modeling for flood susceptibility mapping of arid areas, Southeastern Tunisia. Geocarto Int. 35(9):991–1017.
Tehrany MS, Pradhan B, Jebur MN. 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. J Hydrol. 512:332–343.
Thi Ngo PT, Panahi M, Khosravi K, Ghorbanzadeh O, Kariminejad N, Cerda A, Lee S. 2021. Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. Geosci Front. 12(2):505–519.
Turan İD, Özkan B, Türkeş M, Dengiz O. 2020. Landslide susceptibility mapping for the Black Sea Region with spatial fuzzy multi-criteria decision analysis under semi-humid and humid terrestrial ecosystems. Theor Appl Climatol. 140(3–4):1233–1246.
Vapnik V, Guyon I, Hastie T. 1995. Support vector machines. Mach Learn. 20(3):273–297.
Yang W, Wang M, Shi P. 2012. Using MODIS NDVI time series to identify geographic patterns of landslides in vegetated regions. IEEE Geosci Remote Sens Lett. 10(4):707–710.
Yi Y, Zhang Z, Zhang W, Jia H, Zhang J. 2020. Landslide susceptibility mapping using multi-scale sampling strategy and convolutional neural network: a case study in Jiuzhaigou region. Catena. 195:104851.
Youssef AM, Pradhan B, Jebur MN, El-Harbi HM. 2015. Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. Environ Earth Sci. 73(7):3745–3761.
Youssef AM, Pourghasemi HR. 2021. Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. Geosci Front. 12(2):639–655.
Zeybek M, Şanlıoǧlu İ. 2015. Accurate determination of the Taşkent (Konya, Turkey) landslide using a long-range terrestrial laser scanner. Bull Eng Geol Environ. 74(1):61–76.
Zeybek M, Şanlıoǧlu İ, Özdemir A. 2015. Monitoring landslides with geophysical and geodetic observations. Environ Earth Sci. 74(7):6247–6263.
Zhang J, Gurung DR, Liu R, Murthy MSR, Su F. 2015. Abe Barek landslide and landslide susceptibility assessment in Badakhshan Province, Afghanistan. Landslides. 12(3):597–609.
Zhang T, Han L, Chen W, Shahabi H. 2018. Hybrid integration approach of entropy with logistic regression and support vector machine for landslide susceptibility modeling. Entropy. 20(11):884.
Zhu A-X, Miao Y, Liu J, Bai S, Zeng C, Ma T, Hong H. 2019. A similarity-based approach to sampling absence data for landslide susceptibility mapping using data-driven methods. Catena. 183:104188.
Zhu A-X, Miao Y, Wang R, Zhu T, Deng Y, Liu J, Yang L, Qin C-Z, Hong H. 2018. A comparative study of an expert knowledge-based model and two data-driven models for landslide susceptibility mapping. Catena. 166:317–327.