



Evaluation and comparison of the advanced metaheuristic and conventional machine learning methods for the prediction of landslide occurrence

Chao Yuan¹ · Hossein Moayedi^{2,3}

Received: 20 May 2019 / Accepted: 6 June 2019 / Published online: 8 July 2019
© Springer-Verlag London Ltd., part of Springer Nature 2019

Abstract

The present study aims to assess the superiority of the metaheuristic evolutionary when compared to the conventional machine learning classification techniques for landslide occurrence estimation. To evaluate and compare the applicability of these metaheuristic algorithms, a real-world problem of landslide assessment (i.e., including 266 records and fifteen landslide conditioning factors) is selected. In the first step, seven of the most common traditional classification techniques are applied. Then, after introducing the elite model, it is optimized using six state-of-the-art metaheuristic evolutionary techniques. The results show that applying the proposed evolutionary algorithms effectively increases the prediction accuracy from 81.6 to the range (87.8–98.3%) and the classification ratio from 58.3% to the range (60.1–85.0%).

Keywords Metaheuristic evolutionary · Classification · Landslide perdition

1 Introduction

Traditional approaches of natural slope failure analysis employed various engineering-designed tools [1, 2]. Presenting more progressive designed tools, such as the machine learning-based predictive algorithms, draw attention to a lot of researchers [3, 4]. Most studies have exposed that the machine learning-based techniques are dependable methods to approximate the engineering complex explanations and solutions [5]. The stability of the local slopes against failure is a critical matter that has to be investigated meticulously [6, 7], because of their high impacts on the adjacent engineering buildings (e.g., projects that include excavation and transmission roads, etc.). Also, slope failures cause a lot of damages (e.g., the loss of property and human life)

worldwide every year. There are many factors that need to be considered during the stability of such slopes. As an example, the saturation degree, along with other intrinsic of the soil properties, mostly affects the chances of slope failure [8, 9]. Up to now, many scientists intend to provide effective modeling for the stability of slopes [10, 11]. Some disadvantages of traditional approaches such as the necessity of utilizing laboratory equipment [12–14] along with the high level of complexity make them a difficult solution [15–18]. Additionally, they cannot be utilized as a certain solution, because of their limitation to investigate a specific slope condition (e.g., slope height, the angle of the slope, soil properties, depth of the groundwater level, etc.). Because of the criticality of slope steadiness evaluation, many types of research have been concentrating on tackling this problem of attention. At this time, machine learning, analytical methods, and expert assessment are usually employed to analyze slope conditions [19]. The very first approach is based on experts' knowledge and experiences [20–24]. Using slope stability specialists' judgments, considering the key factors that possibility have higher influences on the further slope failure could be recognized [25]. Though the main disadvantage of the expert evaluation approach is that it mostly relies on judging subjectively and it is infeasible to make sure the consistency of the prediction results [26]. Recently, complexity of landslide failures has caused plenty of losses

✉ Hossein Moayedi
hossein.moayedi@tdtu.edu.vn

¹ College of Architecture and Civil Engineering, Xi'an University of Science and Technology, Xi'an 710054, Shaanxi, China

² Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam

³ Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam

in both financial and psychological aspects around the world. Varnes and Radbruch-Hall [27] introduced the landslide as the whole sort of gravity-made downward mass (i.e., soil, natural cliffs, and artificial deposits) movements at slopes. Landslide may occur on these masses. Scholars [28] have stated that developing countries are more exposed to landslides and more than around 90% of the landslides occurred in these countries. Analyze of the landslide susceptibility of a zone is an effective task for reducing upper mentioned losses [29]. According to a set of geological and environmental states, landslide susceptibility is defined as the spatial possibility of landslide incidents. For various landslide-prone zones worldwide, there are different methods in the case of landslide hazard mapping [30, 31].

Researchers compared the performance of different landslide predictive evaluative methods. For predicting the landslide hazard in Longhai zone (China), He et al. [32] used Naïve Bayes (NB), RBF neural, radial basis function (RBF) and classifier and compared obtained results. They used FR and SVM approaches for predicting the performance of landslide conditioning factors and showed that RBF classifier has better performance in comparison to NB and RBF network with the precision of 88.1%. Chen et al. [33] used different new predictive approaches in the landslide probability prediction including generalized additive model (GAM), ANFIS-FR (adaptive neuro-fuzzy inference system synthesized with FR), and SVM for a specific zone in China. They showed that the SVM has better performance than ANFIS-FR as well as GAM with precisions around 87.5, 85.1, and 84.6%, respectively. In this regard, other scholars used LR and ANFIS predictive methods and planned the unstable rockfall of Firooz Abad-Kojour earthquake, which occurred in 2004, and indicated that ANFIS has better performance. In addition, many researchers have defined and made different hybrid evolutionary approaches with enhancing artificial neural network (ANN) and ANFIS in landslide susceptibility mapping [34–38] as well as flood susceptibility mapping [39–41] for enhancing the performance of usual approaches. To design the ensembles of Random Subspace-based Reduced Error Pruning Trees (RSREPT), MultiBoost according to Reduced Error Pruning Trees (MBREPT), Bagging-based Reduced Error Pruning Trees (BREPT), and Rotation Forest-based Reduced Error Pruning Trees (RFREPT) for landslide hazard prediction, Pham et al. [42] utilized a hybrid technique of Reduced Error Pruning Trees (REPT). The findings of this research revealed the superiority of the BREPT method. Also, Moayedi et al. [37] coupled an MLP neural network with particle swarm optimization (PSO) in the case of spatial landslide hazard modeling at Kermanshah province, western Iran. They concluded that using PSO can facilitate obtaining more precise outputs. Chen et al. [43] also studied the robustness of PSO differential evolution (DE) and genetic algorithm (GA) techniques

to enhance the ANFIS efficiency. Their results showed that ANFIS-DE method outperformed ANFIS-GA, and ANFIS-PSO with respect to the calculated area under the curve (AUCs) of 0.844, 0.821, and 0.780, respectively. For assessment of rainfall-triggered landslide risk, Tien Bui et al. [44] enhanced and used the model of Least-Squares Support Vector Machines (LSSVM) utilizing differential evolution (DE) method and proposed LSSVM model which has better performance than MLP, J48, and SVM algorithms with the accuracy of 82%. The required approximation is commonly performed using a spatial dataset that consists of various landslide conditioning parameters such as stream power index (SPI), altitude, soil, rainfall, climate, lithology, aspect, lithology, and the distance for linear phenomena. Many scholars have selected predictive techniques such as statistical index (SI), certainty factor (CF), index of entropy (IOE) and frequency ratio (FR), and also regression-based methods and evaluated the risk of a landslide by these approaches [45–48]. In this regard, Yang et al. [49] conducted a specific case study for investigating the efficiency of a spatial logistic regression (SLR) method for modeling the landslide hazard in Duwen Highway Basin at Sichuan Province located China. Moreover, many researchers have expanded and proposed a GeoDetector-based approach for selecting the landslide-related factors, properly. The occurred estimation of their proposed model was around 11.9% enhanced compared to the usual logistic regression (LR) model. Additionally, a case study has been conducted for analyzing the landslide occurrence risk of a specific location in China. In this way, IOE as well as certainty factor (CF) methods by utilizing conditioning factors of the slope angle, distance to rivers, plan curvature, profile curvature, distance to faults, geomorphology, distance to roads, topographic wetness index (TWI), slope aspect, general curvature, rainfall, lithology and altitude and also the sediment transport index (STI) as well as the stream power index (SPI) [50]. Based on the respective accuracy of 82.32% and 80.88%. It has been determined that CF calculated the landslide susceptibility map with more validity in comparison to IOE. In terms of making an efficient, fast, and inexpensive prediction of landslide hazard, soft computing (SC) approaches have been highly suggested by scholars because of its computational advances [36, 48, 51–54]. In this way, researchers have performed different studies. Lee et al. [55] have used a support vector machine (SVM) method for specific zones in Korea (Pyeong Chang and Inje). They have employed the SVM method as a reliable tool for analyzing landslide hazard and found that this method had proper results with accuracy by around 81.36% and 77.49% for the Pyeong Chang and Inje zones, respectively. Pradhan and Lee [56] utilized a back-propagation neural network in producing the landslide hazard map in a specific area in Malaysia with 83% precision.

This paper addresses a comprehensive optimization for landslide hazard analysis using six state-of-the-art metaheuristic algorithms. To this end, we first evaluate the capability of seven traditional classification techniques. Various statistical indices are used to distinguish the most capable model. Then the proposed elite model is coupled with six evolutionary algorithms to enhance its performance. As well as the statistical indices, area under the AUROC and the classification rate are considered to compare the efficiency of the used ensembles.

2 Methodology and data collection

2.1 Data collection

Referring to the previous researches in the context of landslide hazard assessment, fifteen key factors including altitude, slope, total curvature, profile curvature, plan curvature, SPI, TRI, TWI, fault river, road, aspect, soil, land use, and geology are considered as the independent variables in this paper. The response variable is also considered the landslide occurrence index consisting of two values of 1 (i.e., 133 rows indicating nonlandslide) and 2 (i.e., 133

rows indicating landslide occurrence). In overall, out of 266 samples, 212 data (i.e., 80%) were randomly selected for training the proposed models. Then the accuracy of them is evaluated by means of the remaining 54 data (i.e., 20%). The mentioned landslide independent factors were produced in the geographic information system (GIS). In fact, some pre-processing actions were carried out for each layer to be created from its basic formats such as contours, polygons, and tabular data. In the next step, the values of each GIS raster were extracted for each landslide and nonlandslide point. Table 1 denotes an example of the dataset used in this study.

2.2 Conventional machine learning classification techniques

Learning algorithms have gained a huge attraction in many fields of research [57]. The machine learning models that are utilized in this work are introduced as follows:

The idea of logistic regression (LR) is drawn on determining a target with dichotomous variables such as true and false or 0 and 1 influenced by some independent factors [58]. For every classification usage, it aims to find a reasonable fit to establish a relationship between the presence or absence of the proposed target event and its key factors. Finally, it

Table 1 An example of the provided dataset for perdition of actual landslide target

Aspect	Curvature	Elevation	Distance to faults (m)	Land use	Lithology	Distance to river (m)	Distance to road (m)	Slope	Soil	TWI	SPI	Target
61.3	-0.1	1744.0	2093.4	3.0	3.0	1870.4	1555.4	10.0	5.0	7.5	58.7	Nonlandslide
9.1	0.3	1367.0	2979.9	3.0	5.0	364.0	398.1	11.6	6.0	1.6	0.2	Nonlandslide
152.2	1.2	1692.0	2729.0	3.0	5.0	531.5	444.2	12.1	5.0	-5.4	0.0	Nonlandslide
173.7	0.2	1483.0	4895.7	3.0	3.0	1456.0	250.0	2.6	5.0	-3.8	0.0	Nonlandslide
28.9	1.8	1475.0	2398.7	2.0	3.0	872.8	1447.0	25.9	5.0	-6.2	0.0	Nonlandslide
139.1	-0.1	2079.0	6848.6	3.0	3.0	1560.6	2130.1	19.2	5.0	-5.9	0.0	Nonlandslide
133.3	0.0	1355.0	1100.2	3.0	5.0	374.4	255.0	6.9	5.0	-4.8	0.0	Landslide
271.2	1.9	1961.0	2185.2	3.0	3.0	3480.4	3228.8	19.7	5.0	-5.9	0.0	Landslide
249.8	0.1	1343.0	31.6	3.0	5.0	190.3	120.0	5.8	5.0	5.3	2.1	Landslide
184.4	0.9	1780.0	1166.2	3.0	9.0	1354.2	147.6	14.6	5.0	-5.6	0.0	Landslide
18.4	0.3	1548.0	2193.6	2.0	5.0	379.5	212.6	9.0	5.0	5.4	5.7	Landslide
226.8	1.3	1263.0	1115.0	3.0	5.0	538.5	353.4	12.8	5.0	3.3	1.4	Landslide
51.4	1.7	1747.0	6395.8	3.0	3.0	878.0	1483.0	24.0	5.0	-6.1	0.0	Landslide
295.3	0.0	1560.0	1894.9	3.0	1.0	90.0	734.1	3.0	6.0	-4.0	0.0	Landslide
119.1	-2.6	1474.0	487.5	3.0	3.0	1077.1	1918.5	27.2	5.0	3.2	6.7	Landslide
270.0	0.5	1851.0	10,066.0	3.0	3.0	1356.9	4890.7	6.8	5.0	-4.8	0.0	Nonlandslide
312.4	-1.3	1539.0	4570.5	3.0	3.0	1592.5	819.3	47.7	5.0	1.5	5.5	Nonlandslide
78.9	0.6	1356.0	954.1	3.0	5.0	338.4	400.0	13.9	5.0	-5.5	0.0	Nonlandslide
56.2	1.3	1774.0	7437.8	3.0	3.0	2088.9	2231.4	30.3	5.0	-6.4	0.0	Nonlandslide
317.9	-0.3	1354.0	1775.5	3.0	5.0	113.1	250.0	8.1	5.0	4.3	1.4	Nonlandslide
229.2	-1.3	1735.0	2992.0	3.0	3.0	1810.4	1674.9	18.8	5.0	1.8	0.7	Nonlandslide
332.0	0.2	1312.0	828.6	3.0	5.0	313.0	76.2	7.6	5.0	-4.9	0.0	Landslide
135.7	-3.6	1494.0	1049.6	3.0	3.0	228.5	860.1	28.9	5.0	3.9	15.5	Landslide

calculates the results through developing a linear equation in which a weight is multiplied by each conditioning factor [59]. Multi-layer perceptron (MLP) is the most common notion of ANNs which its idea was first designed in 1943 [60]. The MLP is capable to discover the non-linear relationship between the proposed variables. An MLP is composed of some layers including one input layer, one or more hidden layer(s), and one output layer containing several computational units called neurons. Each neuron receives the input vectors and assigns some weights and biases to establish a mathematical equation. The name SGD implies stochastic gradient descent learning method [61] which is a common iteration-based optimizer. The SGD breaks a set of samples into mini batches for calculating the gradient on each batch separately, instead of computing the gradient of the cost entirely. In other words, it optimizes a pre-defined cost function to achieve the most accurate parameters of the problem [62]. The name decision table (DT) indicates a tabular classification model in which the data are sorted to find an exact match in the table. In this sense, two responses are likely to appear: (1) if the desired value is met it will be considered as the response, (2) otherwise, the answer is no match found [63]. Generally, four major sections that construct the DT are condition stubs, condition entries, action stubs, and action entries. During the validation stage, the DT checks cases like incompleteness and contradiction [64]. Self-organizing maps (SOM) denote a special notion of ANNs, devoid of the hidden layer [65]. In this model, the input vector is mapped into a lower dimensional map. Considering two inputs, if they are closely related in the reference dataset, they remain closely related in the mentioned lower dimensional map. In such cases, they are mapped into a similar map-unit [65]. Locally weighted learning (LWL) [66] is a well-known lazy learning model. Lazy LWL responses the queries by creating a so-called local model “Naive Bayes (BS)”. The training samples that are similar to the query data are used to develop this model. Regarding the distance between the proposed training point and the prediction point, a weight is assigned to each training sample. In other words, training data which are located closer to the estimation point receive a larger weight [67]. This model is well detailed in [66]. REP tree is a fast notion of decision tree learners. In this model, with respect to the type of the problem, a regression tree or a decision tree is build using the information gain as a splitting criterion. Remarkably, it is pruned using the reduced error pruning. There is only one chance for numeric attributes to be sorted, and the missing ones are dealt with by splitting the related instances into pieces [68].

2.3 Metaheuristic evolutionary techniques

Due to the advances in soft computing, diverse optimization techniques have been successfully used for different

applications [69]. For the prediction of the landslide occurrence, many natural-inspired algorithms have been employed by researchers. These algorithms are utilized to pre-defined objective function (OF). OF function is commonly used for optimizing algorithms and calculating the precision of different algorithms. Minimizing the results of the OF can enhance the accuracy of estimation including regression or classification. First, in the evolutionary approaches, a random population should be defined as the involved relations. The advantages of a determined solution can be evaluated during a repetitive procedure. If the next solution has high accuracy, it is deserted and this should be continuing up to one of the stopping criteria occurred. In the case of the productive tasks, MLP neural network can be used that consists of a general function. In the case of the MLP, the foundation of the MLP optimization is its activation functions and also the weights and biases of the MLP. For achieving a proper performance, enough number of iterations is needed in a defined approach. In the process of the MLP optimization, the computational error reduces in each iteration and also the enhanced weights and biases of the MLP can be utilized for generating new outputs. In addition to the MLP algorithm, there are various evolutionary algorithms that are used in the lecture as follows.

Ant Colony Optimization (ACO) is first introduced in Ref. [70]. It is known as a novel branch of the hybrid evolutionary. This algorithm can mime the foraging life approaches of ant herds. The observed relations are extremely in touch to collaborate. Each ant chooses the path by predicting a possibility. They leave a chemical pheromone trail on the way for the other ants and they guided using this smell and it assistances them to choose the more promising path. This relation helps ants to discover the shortest path among the nectars, for example, the food sources as well as the nest. In Ref. [71], according to the geographical dispensation of a species, biogeography-based optimization (BBO) algorithm has been suggested. In this regard, two different factors of this model are a habitat as well as a habitat suitability indicator which, respectively, produce the possible solution of the suggested issue and its advantage. In this regard, a possible solution is introduced (for example, habitat), which includes some features (decision variables). This method is based on sharing the features among the possible solutions and can enhance the advantage of the possible solutions. This method is based on the migration operators for mutating the calculated habitat suitability index and is properly detailed in Refs. [72, 73]. This method operates via enhancing the diversity of the population for preventing trapping I the local minima [74]. For the first one, the algorithm named evolutionary strategy (ES), which points a stochastic metaheuristic approach has been proposed in Ref. [75]. This method was developed in Ref. [76]. The algorithm of ES uses two different selection and mutation operators and is based on two approaches of

evolution and adaptation. The canonical version consists of six major steps with distinct notions of this method. The population is initialized and it will be analyzed. Then a population is produced with offspring variables. For selecting the elite population, the modality of the offspring is compared to the parents. Holland [77] first designed genetic algorithm (GA) that is known as a robust search approach. This method has been widely utilized for different optimization issues [78, 79]. The GA method proposes an initial group of number strings that each one might be a possible solution. This method is similar to other metaheuristic algorithms. The modality of these parameters is then analyzed and the algorithm distributes to change these groups to a more promising string set. For producing a new population and achieving the excellent generation, a reproduction method can be considered. More details about this method are presented in Refs. [80, 81]. In the lecture, there are different special types of the GA method like probability-based incremental learning (PBIL). This algorithm is proper to conclude the genotype of a probability vector rather than relation. It is introduced as a combination of evolutionary calculation as well as reinforcement learning. The learning method of the PBIL is similar to the algorithm of GA. It gets initiated using a possibility initialization. Next, the instances are produced by the present possibility matrix and the best instance is recognized. After that, the expanded probability matrix is enhanced by the elite sample. In addition, a mutation operator operates probabilistically. Lastly, a termination criterion ends the algorithm [82]. Particle swarm optimization (PSO) is first proposed in Ref. [83], which is based on mimicking the social behavior along with herd lifestyle of the animal. This method is utilized for increasing different typical intelligent methods [84, 85]. Higher learning speed and using less memory are the considerable merits of this algorithm compared to other optimization methods of Imperialist Competition Algorithm (ICA), Artificial Bee Colony (ABC), GA, etc. [37]. In the PSO algorithm, the candidate solution and also population take the name of “particle” and “swarm”. In addition, these particles have the position and velocity, which are known as the determinative factors. Each particle is analyzed against the total population in the case of its position, for enhancing its change. More details about this method are presented in [86, 87].

3 Results and discussion

As stated supra, this study outlines the optimization of a typical machine learning model for landslide occurrence prediction. Two major steps form the body of this paper. First, the performance of seven conventional machine learning classification techniques including LR, MLP, SGD, DT, SOM, LWL, and REP tree is evaluated. Then the elite model is

selected to be optimized with six metaheuristic algorithms, namely ACO, BBO, ES, GA, PBIL, and PSO. Note that five accuracy criteria of kappa statistics (κ), mean absolute error (MAE), root mean square error (RMSE), relative absolute error (RAE in %), and root relative squared error (RRSE in %) were used. Equations (1)–(5) describe these indices:

$$\kappa = \frac{p_o - p_e}{1 - p_e}, \quad (1)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^S |Y_{i_{\text{observed}}} - Y_{i_{\text{predicted}}}|, \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^S [(Y_{i_{\text{observed}}} - Y_{i_{\text{predicted}}})]^2}, \quad (3)$$

$$\text{RAE} = \frac{\sum_{i=1}^S |Y_{i_{\text{observed}}} - Y_{i_{\text{observed}}}|}{\sum_{i=1}^S |Y_{i_{\text{observed}}} - \bar{Y}_{\text{observed}}|}, \quad (4)$$

$$\text{RRSE} = \sqrt{\frac{\sum_{i=1}^S (Y_{i_{\text{predicted}}} - Y_{i_{\text{observed}}})^2}{\sum_{i=1}^S (Y_{i_{\text{observed}}} - \bar{Y}_{\text{observed}})^2}}, \quad (5)$$

in which p_o is the percentage agreement between the classifier and ground truth, and p_e represents the chance agreement. Also, $Y_{i_{\text{observed}}}$ and $Y_{i_{\text{predicted}}}$ stand for the actual and predicted values of landslide occurrence, respectively. The term S defines the number of instances, and $\bar{Y}_{\text{observed}}$ is the average of the target landslide numbers (i.e., 1 and 2).

3.1 Conventional machine learning technique implementation

In this part, the performance of LR, MLP, SGD, DT, SOM, LWL, and REP tree classification models is evaluated for estimating the landslide occurrence using altitude, slope, total curvature, profile curvature, plan curvature, SPI, TRI, TWI, fault river, road, aspect, soil, land use, and geology are considered as landslide independent factors. The Waikato environment for knowledge analysis (WEKA) software was used to implement the mentioned models. Table 2 shows an example of the produced results by each model. Also, the implemented models are compared in Table 3 in terms of KS, MAE, RMSE, RAE (%), and RRSE (%) accuracy indices. Note that, a score-based ranking system is also developed to determine the most capable models. Based on this system, the cells that indicate more accuracy for each model are shown with more intense red color. Finally, the overall

Table 2 Assessment of conventional machine learning classification

Instances	Actual		Predicted											
	LR	MLP	SGD	DT	SOM	Lazy (LWL)	REP tree	LR	MLP	SGD	DT	SOM	Lazy (LWL)	REP tree
1	1	2	1	2	2	1	1	1	1	1	1	1	1	2
2	1	1	1	1	1	1	2	2	2	2	2	2	1	1
3	1	1	1	2	2	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	2	2	2	2	1	2	2	2
6	1	1	1	1	1	1	1	2	2	2	2	2	2	1
7	1	2	1	2	2	1	2	2	2	2	1	1	1	1
8	1	2	1	2	2	2	1	1	1	1	1	2	1	1
9	1	2	1	1	1	2	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	2	2	2	2	2	2	2	1
11	1	1	1	1	1	1	1	2	2	2	2	2	1	1
12	1	2	1	2	2	1	1	2	2	2	2	2	2	1
13	1	1	1	1	1	1	1	1	2	2	2	2	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	2
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	2	2	2	2	2	1	2
17	1	1	1	1	1	1	2	1	1	1	1	1	1	1
18	1	2	1	2	2	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	2	2	1	1	1	1	1	1	1	1	2
22	1	2	1	2	2	2	1	2	2	2	2	2	1	1
23	1	1	1	1	2	1	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	2	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1

No landslide = 1; landslide = 2

Table 3 The obtained results of the conventional machine learning models for the landslide occurrence prediction

Model no.	Classification model	Statistical indexes					Ranking the predicted models					Total ranking
		KS	MAE	RMSE	RAE (%)	RRSE (%)	KS	MAE	RMSE	RAE (%)	RRSE (%)	
1	LR	0.368	0.426	0.460	85.224	92.009	4	2	4	2	4	16
2	MLP	0.796	0.219	0.312	53.976	62.392	7	7	7	7	7	35
3	SGD	0.293	0.353	0.595	70.677	118.892	2	4	1	4	1	12
4	DT	0.256	0.447	0.470	89.429	94.084	1	1	3	1	3	9
5	SOM	0.293	0.353	0.595	70.677	118.892	2	4	1	4	1	12
6	Lazy (LWL)	0.504	0.403	0.429	80.582	85.866	5	3	5	3	5	21
7	REP tree	0.564	0.312	0.395	62.475	79.041	6	6	6	6	6	30

score of each model determines its ranking. According to this table, the MLP outperforms all other six models in terms of all defined indices. In this regard, the KS, MAE, RMSE, RAE, and RRSE of the MLP are obtained as 0.796, 0.219, 0.312, 53.976, and 62.392, respectively. In addition, considering the total ranking score (16, 35, 12, 9, 12, 21, and 30, respectively, obtained for LR, MLP, SGD, DT, SOM, LWL, and REP tree) it can be seen that the MLP presents the most accurate estimation, followed by REP tree and LWL as the second and third efficient models.

3.2 Metaheuristic evolutionary technique implementation

In this part, it is aimed to optimize the performance of the elite model which showed the highest accuracy of the prediction among typical machine learning approaches. As explained, the MLP outperformed other models and is coupled with ACO, BBO, ES, GA, PBIL, and PSO evolutionary algorithms to achieve a more reliable approximation of landslide occurrence risk. This is noteworthy that the mentioned methods try to find the optimal values of the weights and biases of the MLP.

The programming language of MATLAB 2014 was used for this part. Each optimization process was executed within 1000 iterations, and mean absolute error (MSE) as defined as the objective function to measure the accuracy of the ACO-MLP, BBO-MLP, ES-MLP, GA-MLP, PBIL-MLP, and PSO-MLP ensembles in each iteration. Figure 1 illustrates the convergence path of each model. According to this figure, the GA-MLP, PBIL-MLP, and BBO-MLP have reached the lower MSE in comparison with other ensembles. Notably, the MSEs obtained for the ACO-MLP and ES-MLP (around 0.0080) was a little higher than PSO-MLP (around 0.0027). Also, considering the number of iterations that each model needed to reach the minimum error, it was concluded that the ACO-MLP had the best convergence speed; however, the obtained MSE was higher than other models.

In the following, two accuracy criteria of the area under the receiving operating characteristic curve (AUROC),

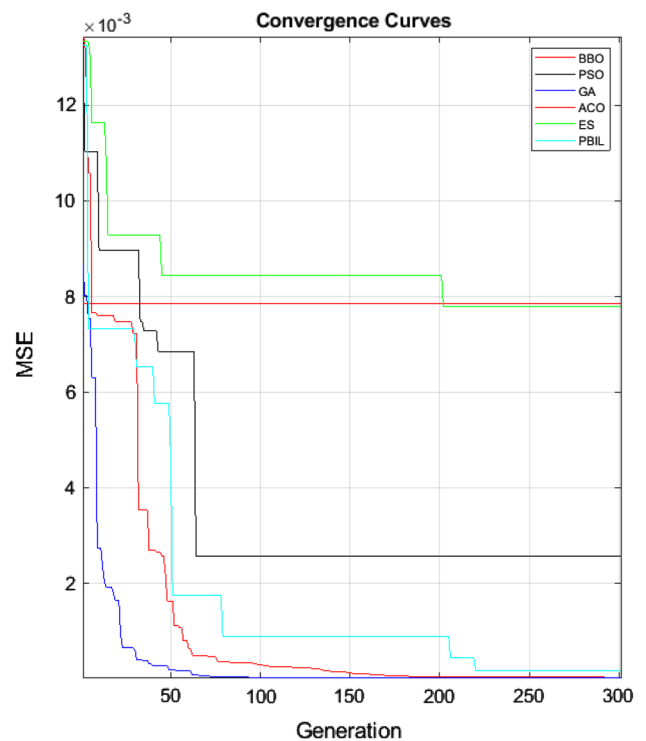


Fig. 1 The convergence curves of the implemented evolutionary models in terms of MSE

which is a well-known method for evaluating the accuracy of a diagnostic issue, and classification ratio are defined. The ROC curves related to the performance of the typical MLP, as well as metaheuristic ensembles, are shown in Fig. 2. Needless to say, the higher the AUROC, the higher the accuracy of the results. As the first result, it can be deduced that the performance of the MLP has enhanced considerably, by applying the mentioned algorithms. Moreover, the GA-MLP has gained the highest AUROC (0.983), followed by BBO-MLP (AUROC=0.971) and PBIL-MLP (0.948). After those, the PSO was a more successful technique for optimizing the MLP (0.917), in comparison with ACO (0.896) and ES (0.878).

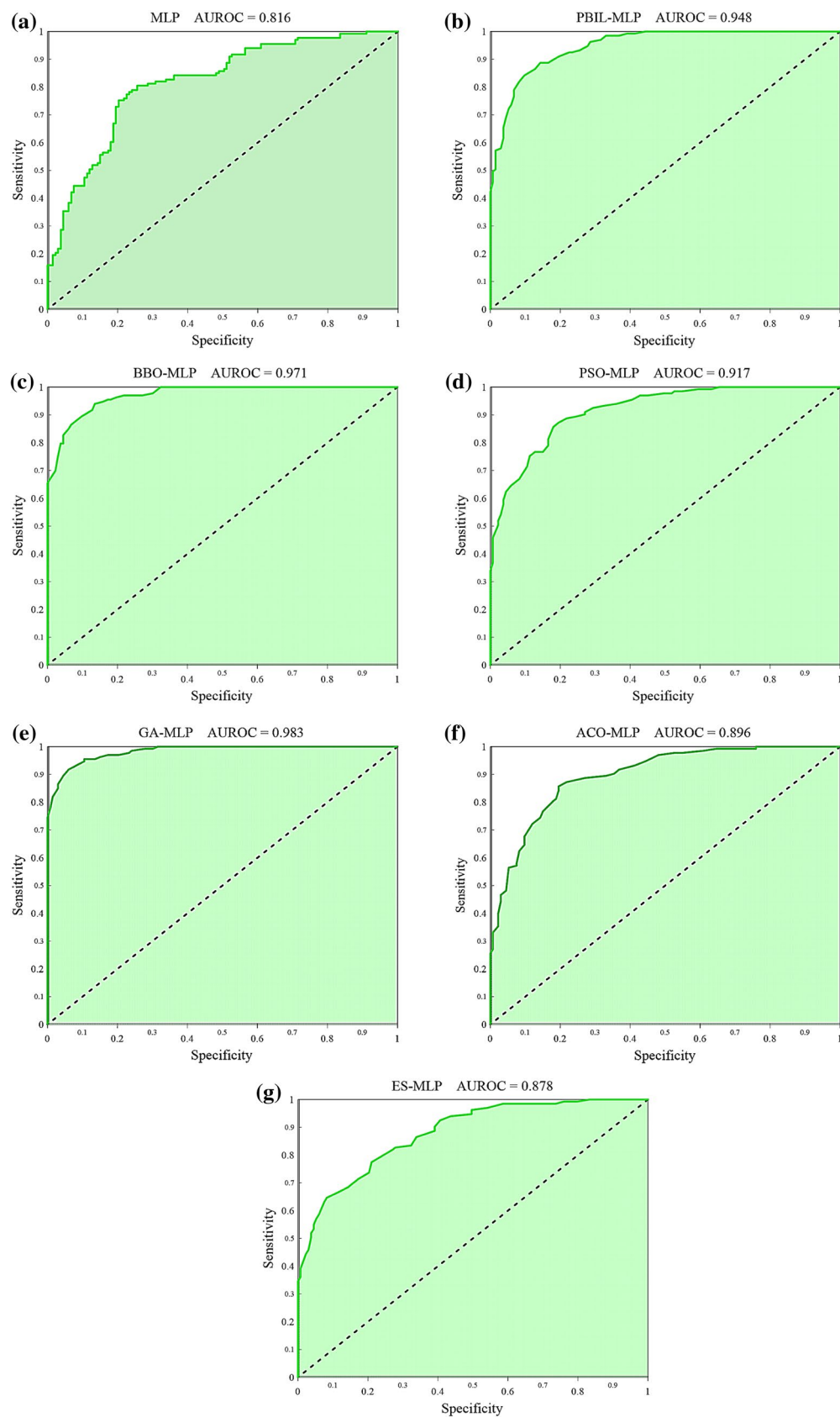


Fig. 2 The ROC curve plotted for the results of the applied models

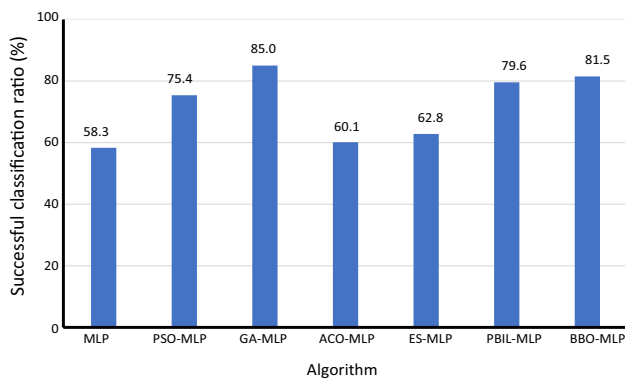


Fig. 3 The classification rate of the performance of the typical MLP and evolutionary ensembles

Furthermore, the percentage of the correctly classified samples demonstrates another method for accuracy evaluation of the applied models. Figure 3 shows the classification ratio of the proposed models. First, increasing the classification ratio of the MLP shows the efficiency of the used optimization algorithms. Also, supporting the AUROC results, the highest value of the classification ratio is obtained for the GA-MLP (85%) followed by BBO-MLP (81.5%) and PBIL-MLP (79.6%). The PSO-MLP has classified 75.4% of the samples correctly. This value was calculated as 62.8% and 60.1%, respectively, for the ES-MLP and ACO-MLP. All in all, the results of the study show that the GA, BBO, PBIL, and PSO have shown the higher capability of optimization of the MLP neural network.

4 Conclusions

Due to the importance of having an appropriate approximation of landslide occurrence risk, this study presented a comprehensive optimization of MLP neural network for landslide occurrence prediction using six capable evolutionary methods, namely ACO, BBO, ES, GA, PBIL, and PSO. To do so, a proper dataset was provided. First, seven conventional machine learning techniques including LR, MLP, SGD, DT, SOM, LWL, and REP tree were evaluated. The results showed that the MLP performed more efficiently with respective KS, MAE, RMSE, RAE, and RRSE of 0.796, 0.219, 0.312, 53.976, and 62.392. In the next step, the elite model (i.e., MLP) was coupled with the mentioned optimization algorithms to achieve a more reliable prediction. The results showed that the AUROC and classification of the MLP (i.e., 0.816 and 58.3%) increased as 0.983, 0.971, 0.948, 0.917, 0.878, and 0.896, and 85, 81.5, 79.6, 75.4, 62.8, and 60.1%, respectively, by applying GA, BBO, PBIL, PSO, ES, and ACO techniques. From the comparison viewpoint, it was revealed that GA outperformed other

optimization methods. Finally, it is worth noting that the presented paper can be improved by performing an optimization for proper selection of landslide conditioning factors, which would be a good idea for future works.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

1. Donald IB, Chen Z (1997) Slope stability analysis by the upper bound approach: fundamentals and methods. *Can Geotech J* 34:853–862
2. Griffiths DV, Lane PA (1999) Slope stability analysis by finite elements. *Geotechnique* 49:387–403
3. Su GS, Zhang Y, Chen GQ, Yan LB (2013) Fast estimation of slope stability based on Gaussian process machine learning. *Disaster Adv* 6:81–91
4. Rodrigues ÉO, Pinheiro VHA, Liatsis P, Conci A (2017) Machine learning in the prediction of cardiac epicardial and mediastinal fat volumes. *Comput Biol Med* 89:520–529
5. Lu P, Rosenbaum MS (2003) Artificial neural networks and grey systems for the prediction of slope stability. *Nat Hazards* 30:383–398
6. Sultan N, Savoye B, Jouet G, Leynaud D, Cochonat P, Henry P, Stegmann S, Kopf A (2010) Investigation of a possible submarine landslide at the Var delta front (Nice continental slope, southeast France). *Can Geotech J* 47:486–496
7. Zhang G, Cao J, Wang LP (2014) Failure behavior and mechanism of slopes reinforced using soil nail wall under various loading conditions. *Soils Found* 54:1175–1187
8. Latifi N, Rashid ASA, Siddiqua S, Majid MZA (2016) Strength measurement and textural characteristics of tropical residual soil stabilised with liquid polymer. *Measurement* 91:46–54
9. Moayedi H, Huat B, Thamer A, Torabihaghighi A, Asadi A (2010) Analysis of longitudinal cracks in crest of Doroodzan Dam. *Electron J Geotech Eng* 15:337–347
10. Hoang N-D, Pham A-D (2016) Hybrid artificial intelligence approach based on metaheuristic and machine learning for slope stability assessment: a multinational data analysis. *Expert Syst Appl* 46:60–68
11. Qi C, Tang X (2018) Slope stability prediction using integrated metaheuristic and machine learning approaches: a comparative study. *Comput Ind Eng* 118:112–122
12. Damiano E, Olivares L (2010) The role of infiltration processes in steep slope stability of pyroclastic granular soils: laboratory and numerical investigation. *Nat Hazards* 52:329–350
13. Moayedi H, Huat BBK, Kazemian S, Asadi A (2010) Optimization of tension absorption of geosynthetics through reinforced slope. *Electron J Geotech Eng* 15:93–104
14. Raftari M, Kassim KA, Rashid ASA, Moayedi H (2013) Settlement of shallow foundations near reinforced slopes. *Electron J Geotech Eng* 18:797–808
15. Marto A, Latifi N, Janbaz M, Kholghifard M, Khari M, Alimohammadi P, Banadaki AD (2012) Foundation size effect on modulus of subgrade reaction on sandy soils. *Electron J Geotech Eng* 17:2015
16. Gao W, Dimitrov D, Abdo H (2018) Tight independent set neighborhood union condition for fractional critical deleted graphs and ID deleted graphs. *Discret Contin Dyn Syst S* 12:711–721

17. Gao W, Guirao JLG, Basavanagoud B, Wu J (2018) Partial multi-dividing ontology learning algorithm. *Inf Sci* 467:35–58
18. Gao W, Wang W, Dimitrov D, Wang Y (2018) Nano properties analysis via fourth multiplicative ABC indicator calculating. *Arab J Chem* 11:793–801
19. Zhang ZF, Liu ZB, Zheng LF, Zhang Y (2014) Development of an adaptive relevance vector machine approach for slope stability inference. *Neural Comput Appl* 25:2025–2035
20. Cheng M-Y, Hoang N-D (2014) Slope collapse prediction using Bayesian framework with k-nearest neighbor density estimation: case study in Taiwan. *J Comput Civ Eng* 30:04014116
21. Pinheiro M, Sanches S, Miranda T, Neves A, Tinoco J, Ferreira A, Correia AG (2015) A new empirical system for rock slope stability analysis in exploitation stage. *Int J Rock Mech Min Sci* 76:182–191
22. Lyu Z, Chai J, Xu Z, Qin Y (2018) Environmental impact assessment of mining activities on groundwater: case study of copper Mine in Jiangxi Province, China. *J Hydrol Eng* 24:05018027
23. Gao W, Guirao JLG, Abdel-Aty M, Xi W (2019) An independent set degree condition for fractional critical deleted graphs. *Discret Contin Dyn Syst S* 12:877–886
24. Gao W, Wu H, Siddiqui MK, Baig AQ (2018) Study of biological networks using graph theory. *Saudi J Biol Sci* 25:1212–1219
25. Aqeel A, Zaman H, El Aal AA (2018) Slope stability analysis of a rock cut in a residential area, Madinah, Saudi Arabia: a case study. *Geotech Geol Eng* 2018:1–14
26. Xiao T, Li D-Q, Cao Z-J, Au S-K, Phoon K-K (2016) Three-dimensional slope reliability and risk assessment using auxiliary random finite element method. *Comput Geotech* 79:146–158
27. Varnes DJ, Radbruch-Hall D (1976) Landslides cause and effect. *Bull Int Assoc Eng Geol* 13:205–216
28. Pourghasemi HR, Mohammady M, Pradhan B (2012) Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. *Catena* 97:71–84
29. Hong H, Miao Y, Liu J, Zhu AX (2019) Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping. *Catena* 176:45–64
30. Chen W, Zhang S, Li R, Shahabi H (2018) Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. *Sci Total Environ* 644:1006–1018
31. Kornejady A, Pourghasemi HR (2019) Producing a spatially focused landslide susceptibility map using an ensemble of Shannon's entropy and fractal dimension (case study: Ziarat Watershed, Iran), spatial modeling in GIS and R for earth and environmental sciences. Elsevier, Oxford, pp 689–732
32. He Q, Shahabi H, Shirzadi A, Li S, Chen W, Wang N, Chai H, Bian H, Ma J, Chen Y, Wang X, Chapi K, Ahmad BB (2019) Landslide spatial modelling using novel bivariate statistical based Naïve Bayes, RBF Classifier, and RBF Network machine learning algorithms. *Sci Total Environ* 663:1–15
33. Chen W, Pourghasemi HR, Panahi M, Kornejady A, Wang J, Xie X, Cao S (2017) Spatial prediction of landslide susceptibility using an adaptive neuro-fuzzy inference system combined with frequency ratio, generalized additive model, and support vector machine techniques. *Geomorphology* 297:69–85
34. Bui DT, Tuan TA, Hoang N-D, Thanh NQ, Nguyen DB, Van Liem N, Pradhan B (2017) Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. *Landslides* 14:447–458
35. Jaafari A, Panahi M, Pham BT, Shahabi H, Bui DT, Rezaie F, Lee S (2019) Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility. *Catena* 175:430–445
36. Tien Bui D, Shahabi H, Shirzadi A, Chapi K, Hoang N-D, Pham B, Bui Q-T, Tran C-T, Panahi M, Bin Ahamd B (2018) A novel integrated approach of relevance vector machine optimized by imperialist competitive algorithm for spatial modeling of shallow landslides. *Remote Sens* 10:1538
37. Moayedhi H, Mehrabi M, Mosallanezhad M, Rashid ASA, Pradhan B (2018) Modification of landslide susceptibility mapping using optimized PSO-ANN technique. *Eng Comput* 2018:1–18
38. Chen W, Panahi M, Tsangaratos P, Shahabi H, Ilia I, Panahi S, Li S, Jaafari A, Ahmad BB (2019) Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility. *Catena* 172:212–231
39. Termeh SVR, Kornejady A, Pourghasemi HR, Keesstra S (2018) Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Sci Total Environ* 615:438–451
40. Tien Bui D, Khosravi K, Li S, Shahabi H, Panahi M, Singh V, Chapi K, Shirzadi A, Panahi S, Chen W (2018) New hybrids of ANFIS with several optimization algorithms for flood susceptibility modeling. *Water* 10:1210
41. Hong H, Panahi M, Shirzadi A, Ma T, Liu J, Zhu A-X, Chen W, Kougiaris I, Kazakis N (2018) Flood susceptibility assessment in Hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution. *Sci Total Environ* 621:1124–1141
42. Pham BT, Prakash I, Singh SK, Shirzadi A, Shahabi H, Bui DT (2019) Landslide susceptibility modeling using reduced error pruning trees and different ensemble techniques: hybrid machine learning approaches. *Catena* 175:203–218
43. Chen W, Panahi M, Pourghasemi HR (2017) Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling. *Catena* 157:310–324
44. Tien Bui D, Pham BT, Nguyen QP, Hoang N-D (2016) Spatial prediction of rainfall-induced shallow landslides using hybrid integration approach of Least-Squares Support Vector Machines and differential evolution optimization: a case study in Central Vietnam. *Int J Dig Earth* 9:1077–1097
45. Demir G, Aytakin M, Akgun A (2015) Landslide susceptibility mapping by frequency ratio and logistic regression methods: an example from Niksar-Resadiye (Tokat, Turkey). *Arab J Geosci* 8:1801–1812
46. Chen W, Chai H, Sun X, Wang Q, Ding X, Hong H (2016) A GIS-based comparative study of frequency ratio, statistical index and weights-of-evidence models in landslide susceptibility mapping. *Arab J Geosci* 9:204
47. 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* 73:3745–3761
48. Chen W, Yan X, Zhao Z, Hong H, Bui DT, Pradhan B (2019) Spatial prediction of landslide susceptibility using data mining-based kernel logistic regression, naive Bayes and RBFNetwork models for the Long County area (China). *Bull Eng Geol Environ* 78:247–266
49. Yang J, Song C, Yang Y, Xu C, Guo F, Xie L (2019) New method for landslide susceptibility mapping supported by spatial logistic regression and GeoDetector: a case study of Duwen Highway Basin, Sichuan Province, China. *Geomorphology* 324:62–71
50. Wang Q, Li W, Chen W, Bai H (2015) GIS-based assessment of landslide susceptibility using certainty factor and index of entropy models for the Qianyang County of Baoji city, China. *J Earth Syst Sci* 124:1399–1415

51. Yao X, Tham LG, Dai FC (2008) Landslide susceptibility mapping based on support vector machine: a case study on natural slopes of Hong Kong, China. *Geomorphology* 101:572–582
52. Zare M, Pourghasemi HR, Vafakhah M, Pradhan B (2013) Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arab J Geosci* 6:2873–2888
53. Polykretis C, Chalkias C, Ferentinou M (2017) Adaptive neuro-fuzzy inference system (ANFIS) modeling for landslide susceptibility assessment in a Mediterranean hilly area. *Bull Eng Geol Environ* 2017:1–15
54. Tian Y, Xu C, Hong H, Zhou Q, Wang D (2019) Mapping earthquake-triggered landslide susceptibility by use of artificial neural network (ANN) models: an example of the 2013 Minxian (China) Mw 5.9 event. *Geomat Nat Hazards Risk* 10:1–25
55. Lee S, Hong S-M, Jung H-S (2017) A support vector machine for landslide susceptibility mapping in Gangwon Province, Korea. *Sustainability* 9:48
56. Pradhan B, Lee S (2010) Regional landslide susceptibility analysis using back-propagation neural network model at Cameron Highland, Malaysia. *Landslides* 7:13–30
57. Wu J, Yu X, Gao W (2017) Disequilibrium multi-dividing ontology learning algorithm. *Commun Stat Theory Methods* 46:8925–8942
58. Menard S (1995) *Applied logistic regression analysis*. Sage University Series, Thousand Oaks
59. Ayalew L, Yamagishi H (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* 65:15–31
60. Hebb DO (1949) *The organization of behavior*. Wiley, New York
61. Bottou L (2010) Large-scale machine learning with stochastic gradient descent, *Proceedings of COMPSTAT'2010*. Springer, Berlin, pp 177–186
62. Choudhury A, Eksioğlu B (2019) Using predictive analytics for cancer identification. In: *IISE annual conference*, Institute of Industrial and Systems Engineers, Orlando, USA
63. Kohavi R (1995) *The power of decision tables*. Springer, Berlin
64. Nguyen TA, Perkins WA, Laffey TJ, Pecora D (1987) Knowledge-base verification. *AI Mag* 8:69–75
65. Larose DT, Larose CD (2014) *Discovering knowledge in data: an introduction to data mining*. Wiley, New York
66. Atkeson CG, Moore AW, Schaal S (1997) Locally weighted learning for control. *Lazy learning*. Springer, Berlin, pp 75–113
67. Friedman JH (1995) Intelligent local learning for prediction in high dimensions
68. Dhakate PP, Patil S, Rajeswari K, Abin D (2014) Preprocessing and classification in WEKA using different classifiers. *Int J Eng Res Appl* 4:91–93
69. Gao W, Zhu L, Wang K (2015) Ontology sparse vector learning algorithm for ontology similarity measuring and ontology mapping via ADAL technology. *Int J Bifur Chaos* 25:1540034
70. Dorigo M (1992) *Optimization, learning and natural algorithms*. PhD Thesis, Politecnico di Milano, Italy
71. Simon D (2008) Biogeography-based optimization. *IEEE Trans Evol Comput* 12:702–713
72. Ma H, Simon D (2011) Blended biogeography-based optimization for constrained optimization. *Eng Appl Artif Intell* 24:517–525
73. Ergezer M, Simon D, Du D (2009) Oppositional biogeography-based optimization. In: *2009 IEEE international conference on systems, man and cybernetics*, San Antonio, TX, USA
74. Ahmadlou M, Karimi M, Alizadeh S, Shirzadi A, Parvinnejhad D, Shahabi H, Panahi M (2018) Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). *Geocarto Int* 34:1–21
75. Bianchi L, Dorigo M, Gambardella LM, Gutjahr WJ (2009) A survey on metaheuristics for stochastic combinatorial optimization. *Nat Comput* 8:239–287
76. Schwefel H-PP (1993) *Evolution and optimum seeking: the sixth generation*. Wiley, Oxford
77. Holland JH (1975) *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. University of Michigan Press, Ann Arbor
78. Moayedi H, Raftari M, Sharifi A, Jusoh WAW, Rashid ASA (2019) Optimization of ANFIS with GA and PSO estimating α ratio in driven piles. *Eng Comput* 2019:1–12
79. Bui X-N, Moayedi H, Rashid ASA (2019) Developing a predictive method based on optimized M5Rules–GA predicting heating load of an energy-efficient building system. *Eng Comput* 2019:1–10
80. Davis L (1991) *Handbook of genetic algorithms*, 1st edn. Van Nostrand Reinhold, New York
81. Whitley D (1994) A genetic algorithm tutorial. *Stat Comput* 4:65–85
82. Ling LY (2016) *Participatory search algorithms and applications*. Doctorate thesis, School of Electrical and Computer Engineering, Universidade Estadual De Campinas, Brazil
83. Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: *Proceedings of the sixth international symposium on micro machine and human science, 1995 (MHS'95)*
84. Nguyen H, Moayedi H, Foong LK, Al Najjar HAH, Jusoh WAW, Rashid ASA, Jamali J (2019) Optimizing ANN models with PSO for predicting short building seismic response. *Eng Comput* 2019:1–15
85. Nguyen H, Moayedi H, Jusoh WAW, Sharifi A (2019) Proposing a novel predictive technique using M5Rules-PSO model estimating cooling load in energy-efficient building system. *Eng Comput* 2019:1–10
86. Kennedy J (2010) Particle swarm optimization. *Encyclop Mach Learn* 2010:760–766
87. Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization. *Swarm Intell* 1:33–57

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.