information

Article

Stacking Ensemble Technique Using Optimized Machine Learning Models with Boruta–XGBoost Feature Selection for Landslide Susceptibility Mapping: A Case of Kermanshah Province, Iran

Zeynab Yousefi [1], Ali Asghar Alesheikh [1,2,*], Ali Jafari [1], Sara Torktatari [1] and Mohammad Sharif [3,*]

[1] Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran 19967-15433, Iran; z.yousefikalestan@email.kntu.ac.ir (Z.Y.); a.jafari2@email.kntu.ac.ir (A.J.); tork_tataari@email.kntu.ac.ir (S.T.)
[2] Geospatial Big Data Computations and Internet of Things (IoT) Lab, K. N. Toosi University of Technology, Tehran 19967-15433, Iran
[3] Institute of Mobility and Urban Planning, University of Duisburg-Essen, 45127 Essen, Germany
* Correspondence: alesheikh@kntu.ac.ir (A.A.A.); mohammad.sharif@uni-due.de (M.S.)

Abstract: Landslides cause significant human and financial losses in different regions of the world. A high-accuracy landslide susceptibility map (LSM) is required to reduce the adverse effects of landslides. Machine learning (ML) is a robust tool for LSM creation. ML models require large amounts of data to predict landslides accurately. This study has developed a stacking ensemble technique based on ML and optimization to enhance the accuracy of an LSM while considering small datasets. The Boruta–XGBoost feature selection was used to determine the optimal combination of features. Then, an intelligent and accurate analysis was performed to prepare the LSM using a dynamic and hybrid approach based on the Adaptive Fuzzy Inference System (ANFIS), Extreme Learning Machine (ELM), Support Vector Regression (SVR), and new optimization algorithms (Ladybug Beetle Optimization [LBO] and Electric Eel Foraging Optimization [EEFO]). After model optimization, a stacking ensemble learning technique was used to weight the models and combine the model outputs to increase the accuracy and reliability of the LSM. The weight combinations of the models were optimized using LBO and EEFO. The Root Mean Square Error (RMSE) and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) parameters were used to assess the performance of these models. A landslide dataset from Kermanshah province, Iran, and 17 influencing factors were used to evaluate the proposed approach. Landslide inventory was 116 points, and the combined Voronoi and entropy method was applied for non-landslide point sampling. The results showed higher accuracy from the stacking ensemble technique with EEFO and LBO algorithms with AUC-ROC values of 94.81% and 94.84% and RMSE values of 0.3146 and 0.3142, respectively. The proposed approach can help managers and planners prepare accurate and reliable LSMs and, as a result, reduce the human and financial losses associated with landslide events.

Keywords: landslide susceptibility mapping; stacking ensemble technique; machine learning; Boruta–XGBoost; feature selection; meta-heuristic algorithms

Citation: Yousefi, Z.; Alesheikh, A.A.; Jafari, A.; Torktatari, S.; Sharif, M. Stacking Ensemble Technique Using Optimized Machine Learning Models with Boruta–XGBoost Feature Selection for Landslide Susceptibility Mapping: A Case of Kermanshah Province, Iran. *Information* **2024**, *15*, 689. https://doi.org/10.3390/info15110689

Academic Editors: Marko Gulić, Antonio Jiménez-Martín and Heming Jia

Received: 17 September 2024
Revised: 18 October 2024
Accepted: 31 October 2024
Published: 2 November 2024

1. Introduction

Natural hazards and disasters pose serious threats to human society. Over the last decade, these disasters have caused approximately 45,000 deaths annually [1,2]. Landslide phenomena in mountainous areas involve the downhill movement of debris, soil, and rocks under the force of gravity [3,4]. The types of landslides include debris flow, rockfall, rock slide, mudslide, and rock avalanche [5]. Natural factors (including rainfall, rapid melting of snow, and earthquakes) and human activities (including construction of roads and infrastructure, destruction of vegetation, and change in land use) contribute to landslides [6].

Therefore, these factors make this area prone to landslides. Landslides are widespread, and 4.9% of all natural disasters were related, causing 1.3% of natural disaster casualties between 1990 and 2015 [7]. Global statistics show that more than 3876 landslides were reported worldwide from 1995 to 2014, resulting in 11,689 injuries and 163,658 deaths [8].

Although natural hazards cannot be prevented, their negative impacts can be reduced by developing effective planning approaches. Identifying landslide-prone areas in different regions to prepare a landslide susceptibility map (LSM) is an effective solution for minimizing its consequences [9]. The LSM shows the probability of a future landslide in areas with similar geological, topographic, and hydrological characteristics based on past landslide events. The points with landslides have the highest susceptibility (equal to 1), and those without have the lowest susceptibility (equal to 0). Finally, for the entire region, we have the range from 0 to 1 to estimate the susceptibility. An LSM has been recognized as an important criterion in many research studies and practical actions at global and regional levels in land use management and spatial planning, directly affecting policies, especially in mountainous areas [10]. Hence, it is necessary to develop efficient and reliable approaches and tools for LSM preparation [11].

Researchers have used various approaches to prepare an LSM. In previous studies, machine learning (ML) models combined with geographic information systems (GIS) were used to prepare an LSM [12,13]. The advantage of ML models is the discovery of complex and nonlinear relationships between types of natural hazards and their factors [14]. In the literature, models such as Random Forest (RF) [15], Support Vector Regression (SVR) [16], Adaptive Neuro-Fuzzy Inference System (ANFIS) [1,6], Artificial Neural Networks (ANNs) [17], and K-Nearest Neighbors (KNNs) [18] have been continuously observed.

The performance of ML models, owing to the difference in (i) the complexity of each area, (ii) the required data, and (iii) the amount of available information on factors influencing landslides, shows very diverse results and efficiency in different areas [7,10]. Therefore, ensemble learning techniques have been introduced to overcome the limitations of individual models and obtain more accurate landslide susceptibility maps than those of individual models [19,20]. Ensemble learning techniques optimize landslide hazard susceptibility modeling; thus, the resulting model performs better [21,22]. Ensemble learning techniques can be divided into two types: homogeneous and heterogeneous. The homogeneous group uses a model as the base [23]. In the heterogeneous group, several models are trained with the same dataset [24].

Hong combined the best first decision tree (BFT) with Bagging, Cascade Generalization, Decorate, MultiboostAB, and Random SubSpace techniques to prepare an LSM. The results showed that in this homogeneous model, the five combined models performed better than the single BFT model [25]. Arab Ameri et al. implemented the credit decision tree (CDT) base model along with three ensemble techniques, CDT–Bagging, CDT–Multiboost, and CDT–SubSpace, to prepare the LSM of the Taleghan Basin of Iran. The AUC results of the homogeneous ensemble techniques were 0.9890, 0.9950, and 0.9950, respectively, which were better than those of the independent model with an AUC of 0.9190 [19].

Li et al. applied three ensemble techniques of Bagging, Dagging, and Decorate on the Radial Basis Function (RBF) model in the CaoBang province of Vietnam. According to the evaluation parameters, Bagging and Dagging had the highest AUC, with values of 0.9800 and 0.9690, respectively, and the other two models, with a value of 0.8990, showed equal performance [23]. Tin Bi et al. implemented a decision tree (DT) in combination with Bagging, Adaboost, and MultiBoost along the national road northwest of Vietnam. The combined models using Bagging and Multiboost, with AUC values of 0.9170 and 0.9100, respectively, performed better than the other two models [26]. Zhao et al. implemented four models in China: ANN, C5.0, Bagging–ANN, and Boosting–C5.0. The Bagging–ANN and Boosting–C5.0 with Root Mean Square Error (RMSE) values of 0.2300 and 0.2350, respectively, produced the best results [27]. In another study in the Karakoram National Park region of Pakistan, Ali et al. combined logistic regression (LR), KNN, and SVR with ensemble techniques, including XGBoost, Dagging, AdaBoost, Cascade Generalization,

Random Forest, and Light Gradient-Boosting Machine. XGBoost had the best output among the combined models, with an AUC of 0.9100. Yu et al. selected LR, SVR, RF, and Deep Learning in the Three Valleys region of China and applied the ensemble heterogeneous stacking approach to these models. The AUC values for the models were 0.7960, 0.8050, 0.6650, and 0.8540, respectively, which were weaker than those for stacking with an AUC of 0.8800 [16]. Lu et al. used three classical neural network models, including the Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), and Gated Recurrent Unit (GRU), as the base models for the stacking technique. The results of the AUC for individual MLP, CNN, and GRU models were 0.84, 0.86, and 0.86, respectively, while the stacking achieved a value of 0.88 [28].

According to the reviewed literature, ensemble learning techniques have resulted in higher accuracy in LSM preparation. In heterogeneous ensemble learning, owing to the use of different base models, it is possible to overcome the limitations of different ML models and improve their accuracy and performance. However, the success of ensemble learning techniques depends on the base models and the weight of the base models to combine their results at the decision level. SVR, Extreme Learning Machine (ELM), and ANFIS were used in this study. These base models have hyper-parameters whose optimal values strongly affect the performance of the model. Therefore, the new Ladybug Beetle Optimization (LBO) and Electric Eel Foraging Optimization (EEFO) meta-heuristic algorithms have been used to obtain the optimal values of the hyper-parameters and the weight of the base models. Additionally, machine learning models highly depend on suitable training data (landslide and non-landslide) and the optimal selection of features. Non-landslide data are usually estimated for locations where landslides have not occurred. For example, it is possible to sample data in non-landslide areas as non-landslide, although these data have many environmental similarities to landslide areas. This issue conflicts with the purpose of the study, which is to identify landslide-prone areas based on the similarity of environmental factors with landslide occurrence areas. Therefore, the first initiative in this study is to determine non-landslide points with a combination of the Voronoi diagram and Shannon entropy. After determining the most important influential features through Borota-XGBoost, we make sure that there is no dependence between them using VIF. With LBO and EEFO new meta-heuristic algorithms, we determine the hyper-parameters of three machine learning models, and in the last step, using both algorithms, we apply a stacking ensemble approach to three optimized basic models.

2. Study Area and Spatial Dataset

The study area is Kermanshah Province, located in the western region of Iran (Figure 1). This province has an area of 25,045 km^2 and is located in the middle of the western side of the country at 45°24′ and 48°07′ E and 33°40′ and 35°18′ N. The lowest and highest elevations are 116 m and 3359 m, respectively. Kermanshah is exposed to humid Mediterranean fronts, which, when colliding with the Zagros highlands, cause snow and rain. Generally, this province has two climates: tropical and cold. The average rainfall in different regions of the province fluctuates between 300 and 800 mm/yr. The province has diverse vegetation, including forests and tropical, cold, and temperate pastures. Kermanshah is located on the high Zagros fault, which is one of the region's most active faults; therefore, this region is prone to landslides [29]. These faults are in the northeast–southwest direction [30]. Landslides are exacerbated by triggering factors such as rainfall and earthquakes.

2.1. Landslide Inventory Map

Data on 116 historical landslides scattered throughout the province were collected from the Forestry and Watershed Organization [31]. The landslide index map documents the landslide types. A landslide inventory is required to formulate models related to landslide risk or susceptibility.

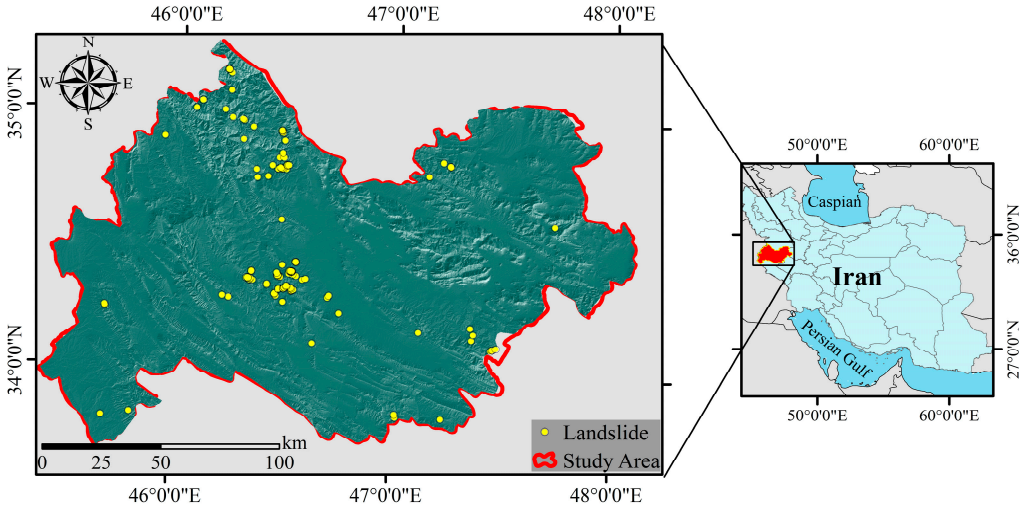


Figure 1. Study area: Kermanshah Province, Iran, and historical landslide events.

2.2. Influential Factors on Landslide

A landslide is a complex phenomenon, and discovering its controlling factors and mechanisms of development in a region is not easy. Therefore, there are no single guidelines for selecting these factors [32]. These factors can be classified from various perspectives. In terms of structure, they can be classified as geological, hydrological, topographic, and environmental. As shown in Table 1, a wide range of factors were used in this study to prepare the LSM.

Table 1. Description of conditioning factors.

Class	Factor	Source	Resolution/Scale
Topographical	Elevation	Iran National Cartographic Center	85 m
	Slope	DEM Derived	
	Aspect		
	Valley Depth		
	Profile Curvature		
	Plan Curvature		
Geological	Lithology	Geological Survey and Mineral Exploration of Iran	1:100,000
	Soil Type		
	Soil Texture		
	Distance to Faults		
Environmental	Land Use		
	Distance to Roads		
Hydrological	Stream Power Index	DEM Derived	85 m
	Topographic Wetness Indices		
	Distance to Drainage	Geological Survey and Mineral Exploration of Iran	1:100,000
	Drainage Density		
	Rainfall	Iran Meteorological Organization	

Elevation is an influential factor in landslides that is used in most models because it affects humidity and temperature. Subsequently, both temperature and precipitation influence soil moisture, making the region prone to landslides [15]. Slope is another important factor, one that indicates the amount of elevation change on the surface [33]. If all the conditions are the same, landslides occur more frequently on steep slopes. A larger slope leads to an increase in shear stress and a reduction in shear strength [34]. The aspect obtained from the elevation map is an important variable that is related to weather conditions (solar radiation, soil moisture, and temperature) [35]. It affects landslides because wind direction, the amounts of sunlight, evaporation and transpiration of snow, humidity and thickness of the soil, and vegetation growth rate vary across the different directions of the slope [7,24].

Among hydrological factors, rainfall is a strong driver of landslides [36]. The Stream Power Index (SPI) determines the erosion capacity of rivers that exacerbate landslides in a region [33]. The Topographic Wetness Index (TWI) indicates the amount of soil moisture and surface saturation [26]. Lower values are associated with steep areas, whereas higher values are associated with flat or valley areas [4]. Distance to streams or rivers is another influencing factor in modeling [19]. River flow reduces the soil's shear strength, making the land prone to landslides [37]. Geological parameters, including the lithology of the area, soil texture, and soil type, also affect landslide occurrence [35]. Rock types vary in their physical and mechanical properties, such as soil resistance, weathering intensity, porosity, and permeability [7]. Soil texture affects the soil's strength and permeability [29]. Landslides typically occur along faults [4]. Because faults are the sites of tectonic activity that can cause landslides, they are considered a factor [29].

Excavation during road construction causes slope instability. Therefore, the distance from the road is considered [8]. Land use factors directly or indirectly contribute to slope stability because different types of land cover vary in infiltration rates, surface streams, evaporation and transpiration rates, and vegetation types [34]. For example, paddy fields lead to increased water infiltration into the soil and reduced shear strength, consequently increasing landslide susceptibility [19]. The study region's dense vegetation and forests stretch from the northwest to the southeast. Most of these forests are composed of oak. Of course, due to the indiscriminate cutting of trees, we see sparse vegetation around dense areas. Sparse vegetation is clearly visible in areas of the northwest. This type of vegetation is also present in the low-height areas of the west and southwest. The pastures and agricultural lands are located in the center of the province (within the height range of 1650–1500 m), which is the main activity of the people in this sector. There are dry agricultural lands in the northeastern highlands (height 2000 m and above).

To create criteria maps, raster maps of the slope, aspect, plan curvature, profile curvature, valley depth, SPI, and TWI were obtained from the DEM layer using ArcMap version 10.8.2 software tools. Using the Euclidean distance tool, a distance raster map was obtained from the spatial layers of the river, faults, and roads. Four polygon layers, including land use, lithology, soil type, and soil texture, were classified into specific classes and converted into raster maps. Figure 2 shows the criteria maps.

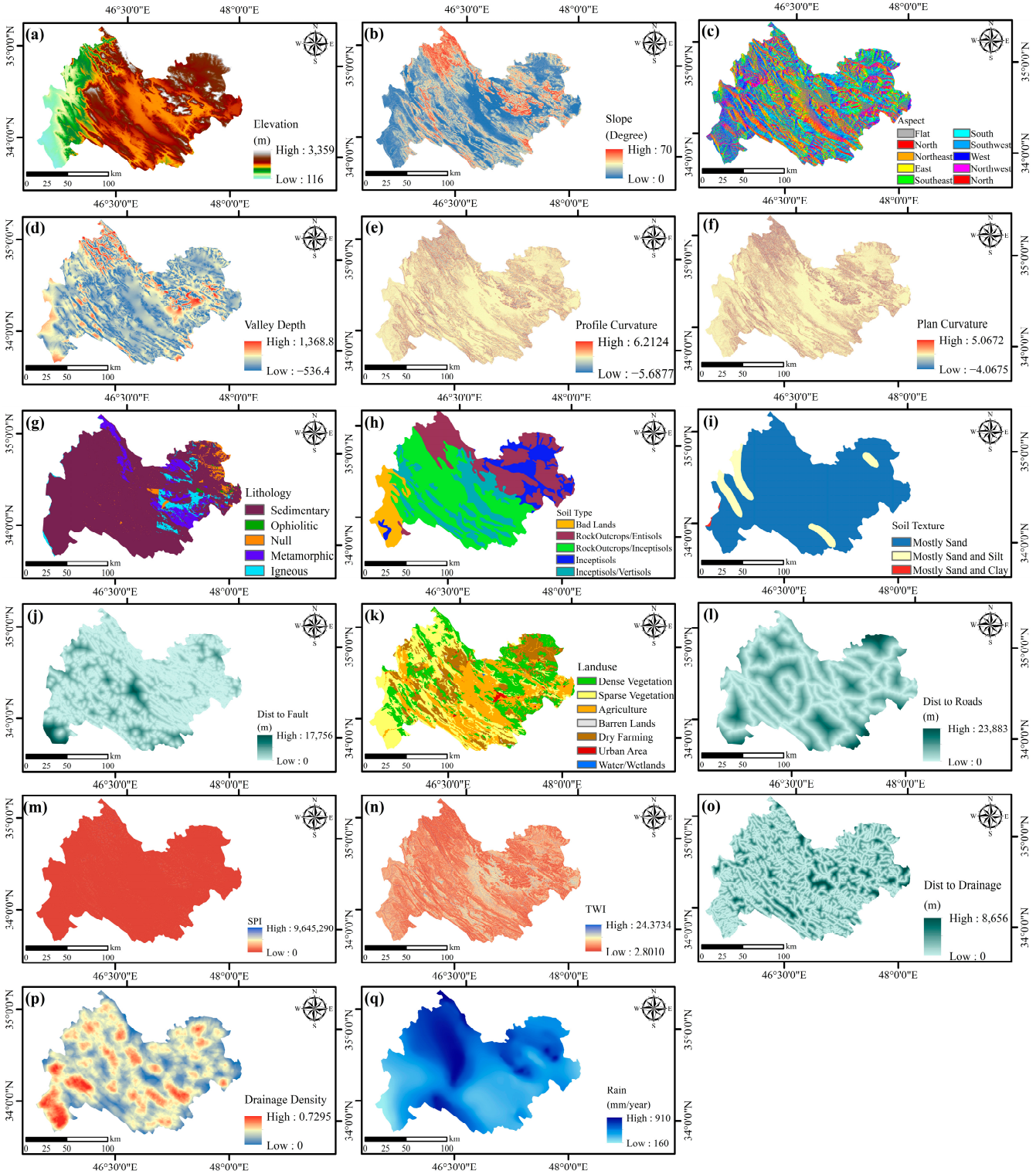


Figure 2. Map of landslide conditioning factors: (**a**) elevation; (**b**) slope; (**c**) aspect; (**d**) valley depth; (**e**) profile curvature; (**f**) plan curvature; (**g**) lithology; (**h**) soil type; (**i**) soil texture; (**j**) distance to faults; (**k**) land use; (**l**) distance to roads; (**m**) SPI; (**n**) TWI; (**o**) distance to drainage; (**p**) drainage density; (**q**) rainfall.

3. Materials and Methods

3.1. Variance Inflation Factor (VIF)

Eliminating factors with a high correlation helps reduce the data dimensions and model complexity. Multicollinearity is a statistical analysis in which the number of independent factors in a multiple regression model is specified. Tolerance (TOL) and VIF represent the correlation effects in the regression between conditional factors, as shown in Equations (1) and (2).

$$\text{Tolerance} = 1 - R_j^2, \tag{1}$$

$$\text{VIF} = \left[\frac{1}{\text{Tolerance}}\right], \tag{2}$$

where R_j is the multiple correlation coefficient between factor j and other conditional factors. Factors with a VIF > 10 are correlated with other factors and thus removed from the modeling process.

3.2. Boruta–XGBoost Feature Selection Method

This hybrid feature selection method is based on the Boruta method, first proposed by Kursa and Rudnicki in 2010 [38]. To improve the performance of Boruta, the XGBoost algorithm can be considered as its core [39]. The algorithm operates as follows: building shadow features (iteratively), training XGBoost models, and iteratively obtaining feature importance [40]. Using the z-score, the most important variables for each input predictor are determined for the repeated features. Original features with z-scores that are smaller than the maximum z-score of the shadow features are removed. The z-score is obtained from Equation (3).

$$\text{z-score} = \frac{\text{MDA}}{\text{SD}}, \tag{3}$$

where MDA represents the average accuracy loss of the input and shadow variables, and SD indicates the standard deviation of the accuracy loss.

3.3. Base Models

3.3.1. Adaptive Fuzzy Inference System (ANFIS)

Fuzzy theory is inspired by uncertainty in human life. Although a fuzzy inference system can model complex processes using if–then rules, it cannot be trained like ML models. To address this issue, an ANFIS, which is a combination of an artificial neural network and a fuzzy system, was proposed by Zhang [41]. This method uses a neural network to determine the appropriate parameters for the membership functions of the fuzzy model. The values of these membership functions, as the ANFIS hyper-parameters, affect the accuracy and efficiency of the model [42]. Therefore, in this study, the values of the membership functions are optimized using meta-heuristic algorithms.

3.3.2. Support Vector Regression (SVR)

SVR is a supervised ML model that is used for regression and classification. SVR creates a hyperplane with the largest marginal distance from the samples of two classes [43]. The larger this margin, the less influence from the outlier data. This model uses a linear combination of the kernel function to model high dimensions and detect the complexity of data in high dimensions [44]. The advantages of SVR include high accuracy, the ability to manage high-dimensional data, and acceptable performance on small datasets [45]. SVR has three parameters: the kernel, cost function, and gamma. The RBF kernel function, defined in Equation (4), is used in this study.

$$K(x, xi) = \exp\left(-\gamma \|x - x_i\|^2\right), \tag{4}$$

where γ is the kernel function. The three parameters, which include kernel scale parameters (KS), the maximum deviation from targets (ε), and the positive compensation parameter

(which regulates the balance between complexity and estimation error [C]), are the hyper-parameters of the SVR that should be optimized.

3.3.3. Extreme Learning Machine (ELM)

An ELM is a single hidden layer feedforward neural network (SLFN), an advanced and more efficient version of the perceptron neural network. The neurons in the hidden layer are randomly generated and are independent of the training data. The weights and biases of the input layer are randomly selected. The output weight can be obtained using the generalized Moore–Penrose inverse analytical relation [46]. Owing to the type of training process of this network, an ELM can solve regression problems with less runtime and higher speed than conventional ANNs [47].

For N distinct samples (X_i,T_i), $X_i = [x_{i1},x_{i2},\cdots x_{in}]^T \in R_n$ and $T_i = [t_{i1},t_{i2},\cdots t_{im}]^T \in R_m$. Standard SLFNs with $\tilde{N}$ hidden neurons and an activation function g(x) are mathematically modeled using Equation (5).

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot X_i + b_i) = o_j \quad j = 1, 2, \ldots N \tag{5}$$

where $w_i = [w_{i1},w_{i2},\cdots w_{in}]^T$ is the weight vector connecting inputs to the hidden layer, $\beta_i = [\beta_{i1},\beta_{i2},\cdots \beta_{im}]^T$ is the weight vector of the i-th hidden neuron with output neurons, $o_j = [o_{j1},o_{j2},\cdots o_{jm}]^T$ is the j-th output vector, and b_i is the threshold of the i-th neuron. In this study, a hyperbolic function is used as the activation function. The weights and biases are our hyper-parameters, which are adjusted using meta-heuristic algorithms.

3.4. Meta-Heuristic Algorithms

3.4.1. Ladybug Beetle Optimizer (LBO)

This algorithm, proposed by Safiri and Nikofard in 2023, was inspired by the natural behavior of ladybugs [48]. Ladybugs look for a warm place to settle in winter. The search process is simulated using this algorithm. The proposed algorithm comprises three main parts: (1) determining the amount of heat at the position of each ladybug, (2) updating the position of the ladybugs, and (3) ignoring the destroyed ladybug(s). Ladybugs move in winter to search for a suitable place. They follow each other by emitting signals. They tend to move towards the ladybugs in front of them (the ladybugs that managed to find a warmer place compared to the others). Several ladybugs deviate from the proper path as they wander in the environment and freeze to death. As a result, their number always decreases when searching for a warm place, which the algorithm considers to increase its speed. To balance exploitation and exploration, a mutation step is considered for some individuals in the population.

3.4.2. Electric Eel Foraging Optimizer (EEFO)

An electric eel is an aquatic animal with an amazing biological structure. This animal can generate electricity in the range of 300–800 volts. The hunting method for electric eels inspired the idea of the Electric Eel Foraging Optimizer algorithm presented by Zhao et al. in 2023 [49]. The amount of electricity emitted by eels varies, so they generate low electric charges for communication and searching and high electric charges during hunting and defending against predators. Their hunting method is collective; therefore, this is a particle swarm algorithm. Mathematical modeling of eel behavior is implemented in four phases: interaction, rest, hunting, and migration [50]. When eels encounter a group of fish, they interact by swimming and moving. The eels then swim in an electrified circular path to trap numerous small fish in the center of the circle. The interaction phase is considered the global exploration phase in the algorithm. When eels do not hunt, they remain in a resting area. The resting phase increases the efficiency of the algorithm. When eels find prey, they collectively swim in a large circle to surround the prey. With the electricity generated by the eels, the electrified circle becomes the hunting area. When eels find prey, they tend to move from the resting area and migrate to the hunting area, which is the migration phase

in the algorithm. If the eels sense that the prey is approaching, they will move toward the candidate position. Otherwise, the eels remain in their current position. For more details, refer to Zhao's paper [49].

3.5. Non-Landslide Sampling

To train the ML models against landslide points, points with zero susceptibility, known as non-landslide points, are required. A combined method, similar to that used by Razavi [51], was employed to create the non-landslide points. A Voronoi diagram map was prepared using the landslide point file in ArcMap version 10.8.2, which was divided into six classes based on area. Subsequently, according to Equation (6), several random points were generated in each class.

$$N = (2 \times n + 4) \times (2 \times m + 1) \quad (6)$$

where n is the number of classes, m is the influence of each class from the previous step (starting from value five), and N is the number of points in each class. In this manner, non-landslide data, twice the number of landslide points, were extracted. Creating random points based on the area of the Voronoi diagram results in more points being sampled in areas where landslides have not occurred. In the next step, the entropy of the random non-landslide points was obtained using Equations (7)–(9). Finally, the desired number of points with the lowest entropy was selected.

Calculation of entropy measurement according to the following Equation (7):

$$E_j = -k \sum_{i=1}^{m} P_{ij} \times \mathrm{Ln}\, P_{ij}, \quad (7)$$

where P_{ij} is the normalized input matrix, and k is the equilibrium constant (between zero and one).

The determination of weight based on the goal according to Equations (8) and (9) is as follows:

$$d_j = 1 - E_j, \quad (8)$$

$$W_j = \frac{d_j}{\sum d_j}, \quad (9)$$

where E_j is the entropy density, d_j is the deviation degree, and W_j is the final entropy weight.

3.6. Proposed Methodology

A dynamic, novel, and efficient hybrid approach is used in this study to predict landslides accurately. The hybrid approach performs intelligent and accurate analysis based on non-landslide sampling and selects the best combination of influencing features, optimal and dynamic models, and ensemble learning techniques to predict landslides. This hybrid approach uses the ANFIS, SVR, and ELM predicting models as well as the new meta-heuristic algorithms LBO and EEFO. Figure 3 shows the outline of the study.

3.6.1. Feature Encoding and Normalization

After generating the landslide and non-landslide data, the datasets were randomly divided into training (80%) and testing (20%) datasets. Then, the categorical variables were converted into transformation numbers using the frequency ratio (FR) so that the ML models could handle these values. Subsequently, the z-score method was used to normalize the data.

3.6.2. Feature Selection

The Boruta–XGBoost method [52] was used to select the best features affecting the landslides. Because the Boruta–XGBoost method does not consider multicollinearity between the data, VIF analysis was initially used to identify independent factors. Finally,

the best combination of features influencing the landslides was determined using the Boruta–XGBoost algorithm.

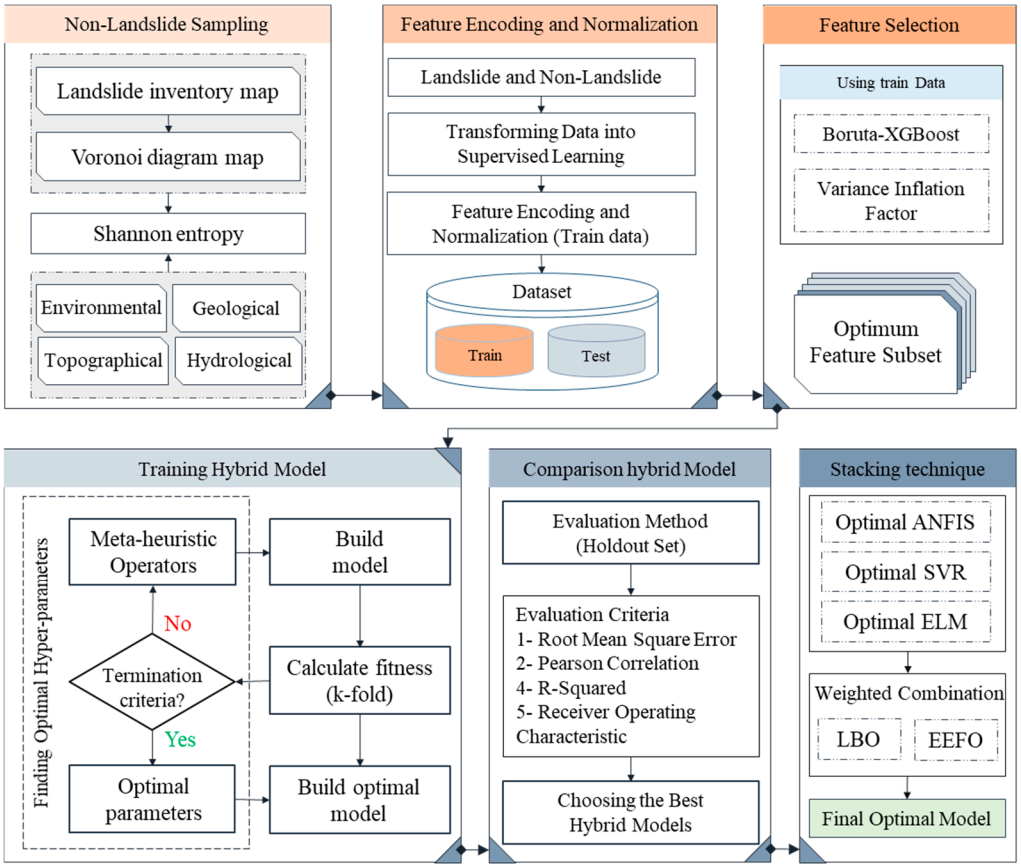


Figure 3. Flowchart of the LSM using the stacking ensemble machine learning models.

3.6.3. Model Framework

After the feature selection step, the ANFIS, SVR, and ELM models were used to predict the landslides. Using the new meta-heuristic algorithms LBO and EEFO, the optimal values of the hyper-parameters were determined to increase accuracy and efficiency. LBO and EEFO are meta-heuristic algorithms that do not require initial parameters. Algorithm 1 shows the procedure used to obtain the optimal hyper-parameters.

Algorithm 1. Steps to perform the proposed method.

1. Initial population % real encoding for hyper-parameter optimization
2. **for** 1 to 1000 **do**
    3. execution of meta-heuristic (LBO, EEFO) operators
    4. **for** 1 to 10 (10-fold cv) **do**
        5. model training using (k − 1) fold
        6. model validation using 1-fold (compute RMSE)
        7. save RMSE
    8. **end**
    9. mean of RMSE on each fold as fitness function
    10. update best solution
11. **end** (end of meta-heuristic algorithms)
12. return the best solution (with optimum hyper-parameters)

3.6.4. Ensemble Learning Techniques

As a final step, the optimal model results were combined to increase the accuracy and reliability of the LS prediction. The best of these models was used to predict landslides by evaluating the optimal ANFIS, SVR, and ELM models. The stacking technique was applied to combine the optimized base models, where the landslide is a weighted combination of the outputs of the models. In this study, the optimal weight was determined using EEFO and LBO. Thus, the final LS prediction was obtained using the following steps (see Algorithm 2). First, optimal models were used to predict the training data output. In stacking learning, the predicted outputs are the inputs, and their observed values are the outputs. Then, using meta-heuristic algorithms (EEFO and LBO), the optimal combination weights of the models and the final model, which is the optimal combination of the ANFIS, SVR, and ELM, were obtained.

Algorithm 2. Steps to perform the stacking ensemble technique.

1.**Input**
(i).$D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ % training dataset
(ii).BaseModels : B_1, B_2, B_3 % optimal ANFIS, optimal SVR, optimal ELM,
(iii).LearningAlgorithm : L_1, L_2 % LBO, EEFO
2. $D' = \Phi$ % generate empty dataset
3. **for** i = 1:n **do**
4. **for** b = 1:B **do**
5. $z_{ib} = \hat{y}_b(x_i)$
6. **end** (end of optimal models' outputs)
7.$D' = D' U((z_{1b}, \ldots, z_{ib}), y_i)$
8. **end**
9.$h' = L(D')$ % learning algorithm for obtaining the optimal combination weight
10. **Output**
i.$\hat{Y}(x) = \hat{y}'(\hat{y}_1(x), \ldots, \hat{y}_b(x))$ % $\hat{y}'$ is the weight combination of three base models

3.6.5. Validation

The evaluation of ML models is critical. The output accuracy of the model should be calculated to measure its performance and robustness. The models were evaluated using a set of quantitative criteria and common indicators for susceptibility maps, including RMSE, Pearson correlation coefficient, r-squared, and AUC-ROC [16]. Landslide occurrence and non-occurrence are inherently binary classification problems. However, susceptibility mapping is used to determine the possibility of occurrence. The model output, such as the regression problem, is a soft output instead of a hard output in the classification. Therefore, the regression criteria can be used to evaluate the model in addition to the classification criteria. This type of problem-solving attitude is common in the preparation of susceptibility maps and is also common in other areas, such as flood susceptibility mapping [53,54] and land subsidence susceptibility mapping [42,55].

4. Results and Implementation

4.1. Create Non-Landslide Points

A Voronoi diagram of landslide points for non-landslide data sampling was first created and divided into six classes based on the area. Subsequently, 230 candidate points were randomly generated (Figure 4a). In the first class with the largest area of polygons, 85 points were created; in the second class, 63 points; in the third class, 42 points; in the fourth class, 25 points; in the fifth class, 12 points; and in the sixth class, 3 points were created. Next, an entropy level of 230 points was calculated. Finally, 117 points with the lowest entropy weights were selected as non-landslide points (Figure 4b).

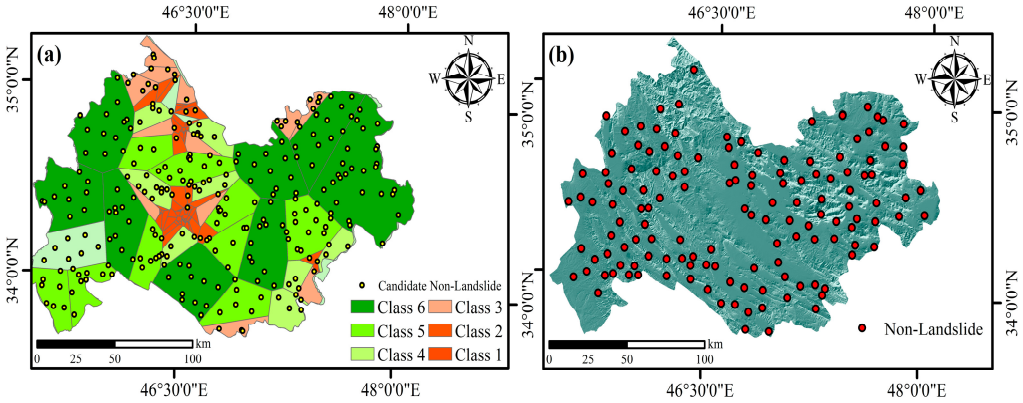


Figure 4. Steps to create non-landslide points. (**a**) Candidate points from the Voronoi map. (**b**) Final non-landslide points from the combination of the Voronoi and entropy maps.

4.2. Results of Feature Selection

Determining the factors that influence landslides is important for the modeling results. The Boruta–XGBoost method was used to select the influencing features and eliminate redundancies. After implementing this algorithm in R version 2021.09.2 software, the results showed that out of the 17 considered features, 4 features, including profile curvature, valley depth, TWI, and aspect, with scores lower than the maximum shadow, should be eliminated from the model. Figure 5 shows the results of the proposed algorithm. In this algorithm, rainfall was selected as the most important feature. The distance from the drainage and the soil type were the other factors with scores over 10. The SPI factor scores, including the distance from the road, distance from the fault, slope, land use, and drainage density, were calculated within the range of 5 to 10. The factors with scores in the maximum shade range and lower than five were lithology, soil texture, plan curvature, and elevation. Since the Boruta–XGBoost method does not consider the relationship between factors such as correlation and collinearity, the VIF was calculated. The values of this coefficient for the 13 features selected in the previous step were all less than 10, as shown in Table 2. The elevation had the highest score with a VIF of 1.6425, and the slope had the lowest score with a VIF of 1.0667. All 13 desired factors had inflation coefficients of 1 to 2. These factors exhibited no significant correlations with one another and were subsequently incorporated into the modeling process.

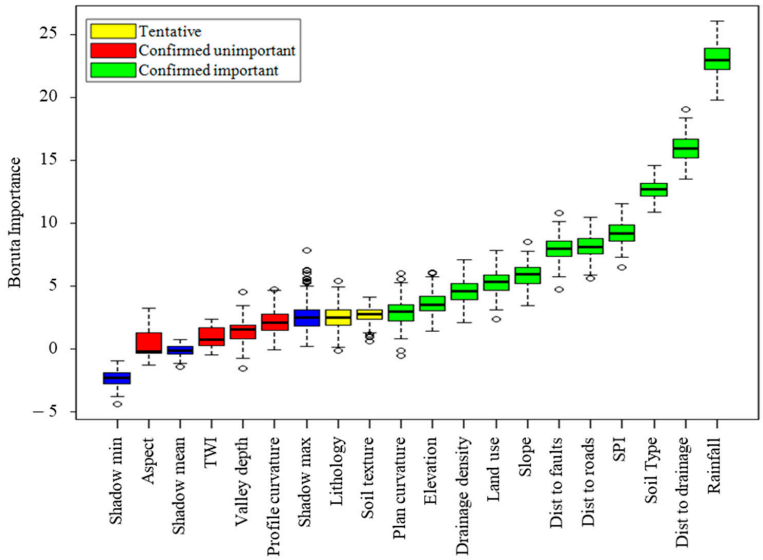


Figure 5. Factor ranks by importance extracted using the Boruta–XGBoost results.

Table 2. Variance inflation factor (VIF) of the finally selected landslide factors.

Factor	VIF	Factor	VIF
Plan curvature	1.3324	Soil type	1.1754
Land use	1.3372	Drainage density	1.3431
Distance to roads	1.5710	Slope	1.0667
SPI	1.2547	Elevation	1.6425
Distance to faults	1.1807	Lithology	1.1932
Soil texture	1.3293	Rainfall	1.4603
Distance to drainage	1.0845		

4.3. Optimizing Base Models

In this study, two new optimizers, LBO and EEFO, which have shown high performance and efficiency in various experiments [48,49], were used to optimize the hyper-parameters of three models: ANFIS, ELM, and SVR (with RMSE as the cost function). In addition, these two meta-heuristic algorithms do not require initial values to adjust, which is advantageous. The fitness function, that is, the mean RMSE with 10-fold, was used to output the fine-tuned models as the final model. The optimization process was performed after 1000 iterations or when the RMSE was 0, with a population of 200.

The hyper-parameters of the ANFIS model are membership function values. Therefore, a basic fuzzy system was generated using the training data. The values of the membership functions were then optimized during the optimization process. The optimal values for the SVR hyper-parameters (C, KS, and Epsilon) and ELM (weight between network layers w_i and β_i) were obtained. According to Figure 6, the optimal values for the SVR parameters converged at the 100th iteration. However, the optimization process for the ANFIS and ELM models continued until the 1000th iteration.

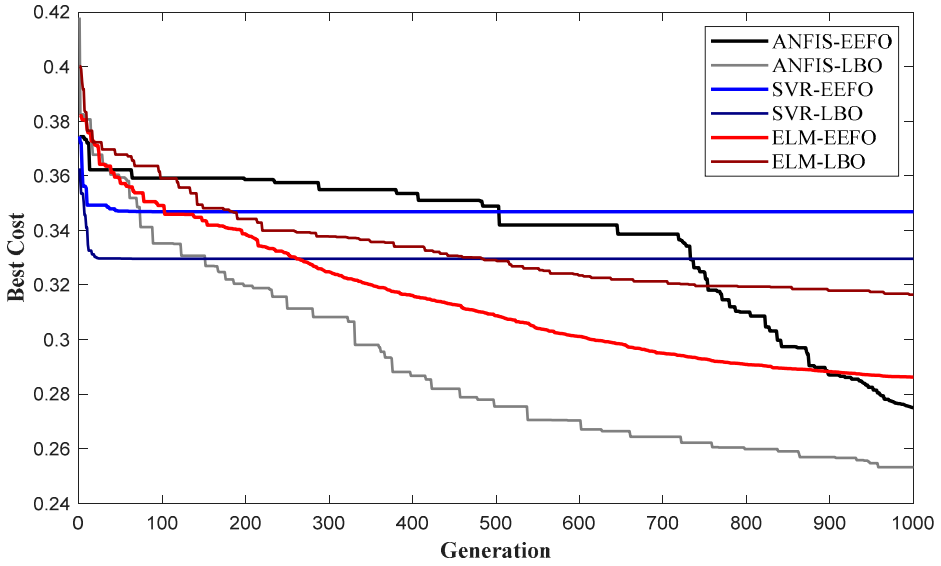


Figure 6. Convergence diagram of the ML models using meta-heuristic algorithms.

Table 3 presents the evaluation criteria values of the optimized base models. As can be seen, optimizing the hyper-parameter values improves the accuracy. Based on the results of each optimized model, they had an acceptable accuracy for the LSM preparation. According to this table, ANFIS–EEFO, with AUC = 0.9437 and RMSE = 0.3520, has a better accuracy than ANFIS–LBO. However, the SVR and ELM models, in combination with LBO, showed better results than the EEFO algorithm. Therefore, the combined ANFIS–EEFO, SVR–LBO, and ELM–LBO models were used as the basic models for stacking.

Table 3. Model performance in the training and validation datasets.

Hybrid Models	AUC		RMSE		R		R-Squared	
	Train	Test	Train	Test	Train	Test	Train	Test
ANFIS–EEFO	0.9721	0.9437 *	0.3174	0.3520 *	0.8351	0.7674 *	0.5970	0.5042 *
ANFIS–LBO	0.9799	0.9394	0.2986	0.3694	0.8624	0.7415	0.6432	0.4537
ANFIS	0.9117	0.9156	0.3586	0.3608	0.6970	0.6940	0.4857	0.4790
SVR–EEFO	0.9754	0.9221	0.2556	0.3421	0.8791	0.7429	0.7387	0.5316
SVR–LBO	0.9795	0.9416 *	0.2522	0.3173 *	0.8882	0.7846 *	0.7456	0.5970 *
SVR	1.0000	0.8983	0.0736	0.4089	0.9999	0.6086	0.9783	0.3310
ELM–EEFO	0.9733	0.8593	0.3275	0.4036	0.8233	0.6371	0.5711	0.3482
ELM–LBO	0.9684	0.8853 *	0.3484	0.3889 *	0.7973	0.6552 *	0.5145	0.3947 *
ELM	0.8935	0.8896	0.3980	0.4019	0.6600	0.6416	0.3664	0.3536
Stacking–EEFO	0.9820	0.9479	0.2443	0.3149	0.8952	0.7991	0.7613	0.6038
Stacking–LBO	0.9820	0.9481 *	0.2442	0.3146 *	0.8950	0.7994 *	0.7614	0.6039 *

* Denotes best performance.

Therefore, the ensemble stacking technique was applied to these three models to obtain a combined model. The ensemble technique combines the outputs of the base models by considering the obtained weights for each to enhance the accuracy of the output. Two algorithms, EEFO and LBO, were used to obtain the best combination of weights for the models in such a way as to minimize the RMSE (as the cost function). According to the weights of the base models presented in Table 4, a higher weight was assigned to the SVR–LBO hybrid model because it has better results than the other base models. The stacking ensemble technique improves the results based on the evaluation values.

Table 4. Weights from meta-heuristic algorithms for base models.

Ensemble Technique	ANFIS–EEFO	SVR–LBO	ELM–LBO
Stacking–EEFO	0.3275	0.5011	0.4689
Stacking–LBO	0.4932	0.8246	0.6363

Furthermore, the simple form of the ANFIS, SVR, and ELM models demonstrated higher accuracy and better performance compared to a similar study conducted by Ghayur Sadigh et al. [29]. This improvement could be due to the targeted selection of non-landslide points and the feature selection process.

4.4. LSM Preparation

Susceptibility maps were prepared after training and evaluating the models. For this purpose, the factors influencing the landslides were extracted for the entire study area. Next, the models were run for the entire area, which produced an output showing each point's susceptibility to landslides. The final maps show the susceptibility of landslide occurrence, which varies between 0 and 1. The resulting maps were classified into five susceptibility classes using the natural break method: very low, low, moderate, high, and very high (Figure 7). The natural break method has been consistently observed in susceptibility mapping and effectively arranges data clusters most efficiently [19,53,55].

According to the maps obtained, the northwest, central, and parts of the northeast regions, where the density of landslide points is high, are classified as having very high susceptibility in all five models. The southern, western, and parts of the northeast, which lack landslide points, are estimated as having two classes of very low and low susceptibility. The percentages of the area of the classes in the five models are shown in Figure 8. The largest areas of the very high and high susceptibility classes belong to the ANFIS model, covering 21.19% and 26.82% of the entire area, respectively. For this reason, it also had the lowest percentage of the very low susceptibility class, with a value of 9.8% among all models.

The two ensemble models produced almost identical estimates for all the susceptibility classes. These two models predicted that the low and very low susceptibility classes comprise 45%, which is the highest compared to the base models. The SVR model obtained the highest value in the moderate susceptibility class, with a value of 27%, whereas the other four models had similar estimates in the range of 20% to 22% for this class. The SVR model predicted 54.87% of the entire region in the low and moderate susceptibility classes.

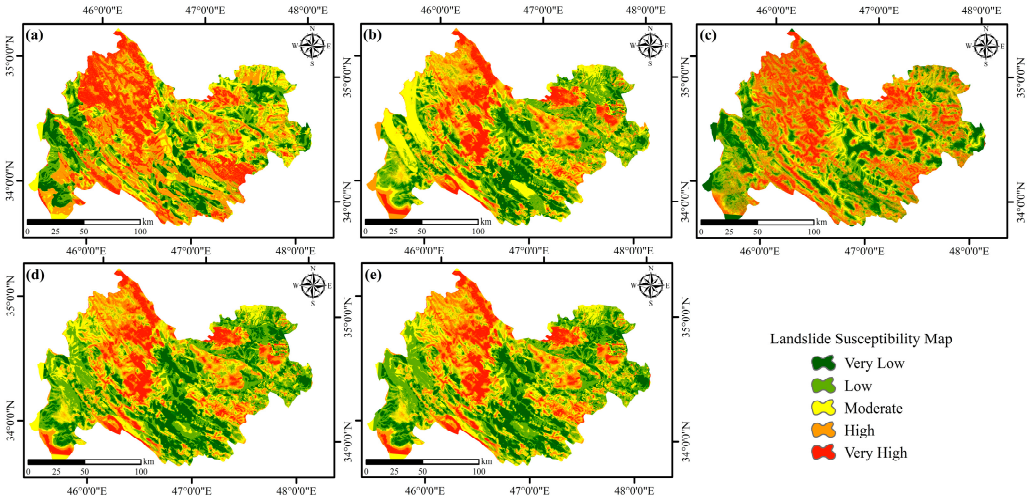
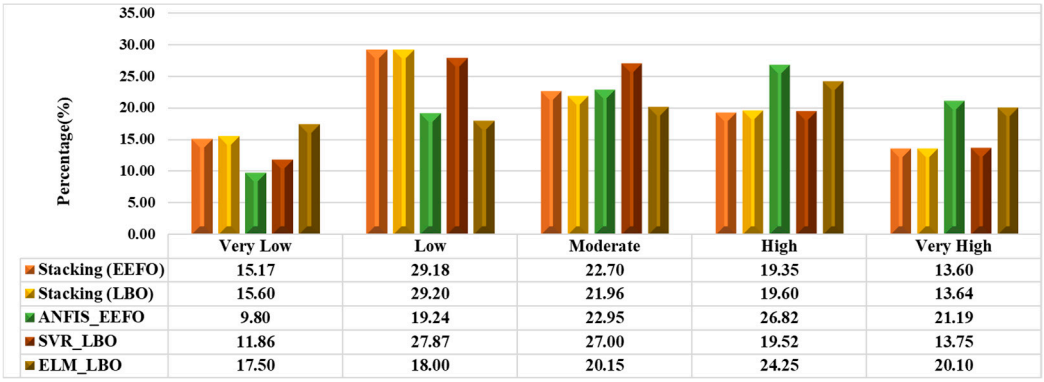


Figure 7. Landslide susceptibility mapping derived from different models: (**a**) ANFIS–EEFO; (**b**) SVR–LBO; (**c**) ELM–LBO; (**d**) Stacking–EEFO; (**e**) Stacking–LBO.



	Very Low	Low	Moderate	High	Very High
Stacking (EEFO)	15.17	29.18	22.70	19.35	13.60
Stacking (LBO)	15.60	29.20	21.96	19.60	13.64
ANFIS_EEFO	9.80	19.24	22.95	26.82	21.19
SVR_LBO	11.86	27.87	27.00	19.52	13.75
ELM_LBO	17.50	18.00	20.15	24.25	20.10

Figure 8. Percentage of area for susceptibility classes in the ML models.

Table 5 lists the number of landslide points in different classes. According to this table, all the models detected most landslide points in the high and very high susceptibility classes. Additionally, the stacking ensemble technique detected more points in the high and very high susceptibility classes, indicating better accuracy of the stacking ensemble technique.

Table 5. Number of landslide points in different classes.

	Very Low	Low	Moderate	High	Very High
ANFIS–EEFO	1	1	13	38	63
SVR–LBO	1	2	4	27	82
ELM–LBO	1	2	8	21	84
Stacking–EEFO	1	2	3	22	88
Stacking–LBO	1	1	4	19	91

5. Discussion

Landslides are natural hazards that have harmful and inevitable impacts. Therefore, finding a solution to mitigate the damage caused is crucial. A susceptibility map, which estimates the probability of landslide occurrence, is considered an effective method. Owing to topographical and environmental conditions, the Kermanshah province in Iran is prone to landslides. Thus, preparing an LSM helps managers, government officials, and other people manage this natural and physical phenomenon. ML models, which are efficient for preparing LSMs, face various challenges. One of the first challenges is the selection of non-landslide points. The random selection of points is a common method for selecting non-landslide points [33], but it is unreliable due to its random nature. In some studies, a landslide point density map was used to ensure that non-landslide points were not selected near the landslide points. This study used a combined method based on the Voronoi diagram and Shannon entropy to select the non-landslide points.

The Boruta–XGBoost algorithm was used to select the factors influencing landslides. Four features, including aspect, valley depth, profile curvature, and TWI, were removed from the modeling as shadows or duplicate features. According to this algorithm, rainfall, distance to the drainage, soil type, SPI, distance to roads, and distance to faults were identified as the most important factors influencing landslides. Despite efforts to calculate the importance of landslide conditioning factors, discrepancies have persisted. Therefore, in Hwan's study [56], profile curvature was introduced as the most important factor affecting landslides and was excluded from our modeling. Li identified elevation, distance to the drainage, lithology, and TWI as the most important factors [57], while TWI was removed from our conditioning factors. In Nhu's study [15], slope and TWI were the most important factors, whereas aspect and rainfall factors were the least important factors. Ghayur Sadigh [29], who worked in the same study area, used the relief feature selection method and introduced rainfall, distances from roads, and drainage, followed by elevation, slope, TWI, valley depth, and distance from faults as the most important features of landslide conditioning. The first three factors are consistent with our results, but valley depth and TWI were excluded from our feature selection process. Different study areas or algorithms lead to significant differences in the identification of the conditioning factors. The importance of the factors is specific to a region and cannot be generalized to other regions. In addition, different feature selection algorithms may yield different results for the same region.

After selecting the influencing factors, three models were chosen: ANFIS, SVR, and ELM. Setting the hyper-parameters of ML models is crucial for the final accuracy of the model. Determining the hyper-parameters in the model is computationally expensive and time-consuming. The trial-and-error method also requires user experience to set the initial configurations. In this study, we used two new meta-heuristic algorithms, EEFO and LBO, to improve the accuracy of the base models. The results showed that all the optimized models achieved high learning and prediction performance. Our optimization results were consistent with those of studies in the field of model optimization for preparing susceptibility maps [1,16,42,58]. Hybrid models reduce noise and variance and prevent problems such as overfitting and underfitting.

In the final implementation stage, we applied the ensemble stacking technique, which is a heterogeneous method, to the optimized base models. The ensemble technique increases accuracy and robustness while reducing uncertainty and overfitting problems. In the stacking technique, the algorithm determines the weight of each model based on its accuracy and combines the models to reach a suitable output. To obtain the weights, two meta-heuristic algorithms were used instead of a single machine learning model. The results of the ensemble approach were consistent with those of other studies on ensemble methods [56,57].

Finally, LSMs for five models were prepared, which were divided into five classes: very low, low, moderate, high, and very high susceptibility, using the natural breaks method. The optimal ANFIS, SVR, and ELM models estimated the total areas of the moderate, high, and

very high classes as 70.96, 60.27, and 64.5, respectively, a significant portion of the province's area. However, the ensemble techniques, Stacking–EEFO and Stacking–LBO, predicted more than 50% of the region with moderate to high susceptibility. The environmental and topographical conditions of Kermanshah Province, such as rainfall levels and fault location, make this area prone to landslides, which is consistent with the LSM obtained in this study area [29]. Managers and government officials should pay close attention to managing land use and infrastructure to prevent the aggravating factors affecting landslides. The obtained LSMs can help with better planning and saving lives.

The main limitation of our study was the data used. We employed a small dataset of landslide points (116) and conditioning factors with low spatial accuracy. Although we tried to improve the process of our ML models using cross-validation, increasing the dataset using techniques such as remote sensing data is recommended. Environmental factors were selected based on exploratory and quantitative procedures and a literature review; however, the mechanism between conditioning factors and land was not considered. Further refinement of the input data, such as continuous updates and implementation of new factors, is desirable to obtain a susceptibility model that considers all the factors influencing landslides. In preparing landslide data to non-landslide data, the ratio was 1:1; therefore, we did not consider the imbalance between them. If the actual ground conditions are assessed and there is an imbalance between the two groups of points, their influence in our models is unknown.

6. Conclusions

Kermanshah is a mountainous province in Iran with high average rainfall. It is located on the highly active fault of Zagros. These factors make this area prone to landslides. In this study, we developed a stacking ensemble technique based on machine learning and optimization algorithms to increase the accuracy of LSMs and manage small datasets. The feature selection approach helped eliminate duplicate features from modeling so that the dimension of the feature space did not increase unnecessarily and overfitting did not occur. The optimization of the hyper-parameters for base models was performed using two new algorithms, leading to promising results. The ensemble stacking technique, aimed at increasing the accuracy and robustness, used the EEFO and LBO algorithms, consequently improving the results. Our results are remarkable, considering the small dataset. The landslide susceptibility maps obtained effectively indicate the susceptibility of the study area to landslides. This methodology can be applied to other areas with various environmental parameters that affect the modeling performance. Furthermore, the results can assist planners and decision-makers in land use planning.

Author Contributions: Conceptualization, A.A.A. and A.J.; methodology, Z.Y. and A.J.; software, Z.Y., A.J. and S.T.; validation, Z.Y. and A.J.; formal analysis, Z.Y. and A.J.; investigation, Z.Y., A.J. and S.T.; resources, A.J.; data curation, A.J.; writing—original draft preparation, Z.Y., A.J. and S.T.; writing—review and editing, A.A.A., A.J., S.T. and M.S.; visualization, Z.Y., A.J. and S.T.; supervision, A.A.A. and A.J.; project administration, A.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Panahi, M.; Gayen, A.; Pourghasemi, H.R.; Rezaie, F.; Lee, S. Spatial Prediction of Landslide Susceptibility Using Hybrid Support Vector Regression (SVR) and the Adaptive Neuro-Fuzzy Inference System (ANFIS) with Various Metaheuristic Algorithms. *Sci. Total Environ.* **2020**, *741*, 139937. [CrossRef] [PubMed]
2. Pourghasemi, H.R.; Rahmati, O. Prediction of the Landslide Susceptibility: Which Algorithm, Which Precision? *Catena* **2018**, *162*, 177–192. [CrossRef]

3. Hussain, M.A.; Chen, Z.; Zheng, Y.; Zhou, Y.; Daud, H. Deep Learning and Machine Learning Models for Landslide Susceptibility Mapping with Remote Sensing Data. *Remote Sens.* **2023**, *15*, 4703. [CrossRef]
4. Pourghasemi, H.R.; Gayen, A.; Park, S.; Lee, C.-W.; Lee, S. Assessment of Landslide-Prone Areas and Their Zonation Using Logistic Regression, Logitboost, and Naïvebayes Machine-Learning Algorithms. *Sustainability* **2018**, *10*, 3697. [CrossRef]
5. Ahmad, H.; Alam, M.; Yinghua, Z.; Najeh, T.; Gamil, Y.; Hameed, S. Landslide Risk Assessment Integrating Susceptibility, Hazard, and Vulnerability Analysis in Northern Pakistan. *Discov. Appl. Sci.* **2024**, *6*, 7. [CrossRef]
6. Moayedi, H.; Mehrabi, M.; Kalantar, B.; Abdullahi Mu'azu, M.; A. Rashid, A.S.; Foong, L.K.; Nguyen, H. Novel Hybrids of Adaptive Neuro-Fuzzy Inference System (ANFIS) with Several Metaheuristic Algorithms for Spatial Susceptibility Assessment of Seismic-Induced Landslide. *Geomat. Nat. Hazards Risk* **2019**, *10*, 1879–1911. [CrossRef]
7. Kumar, C.; Walton, G.; Santi, P.; Luza, C. An Ensemble Approach of Feature Selection and Machine Learning Models for Regional Landslide Susceptibility Mapping in the Arid Mountainous Terrain of Southern Peru. *Remote Sens.* **2023**, *15*, 1376. [CrossRef]
8. Achu, A.; Aju, C.; Di Napoli, M.; Prakash, P.; Gopinath, G.; Shaji, E.; Chandra, V. Machine-Learning Based Landslide Susceptibility Modelling with Emphasis on Uncertainty Analysis. *Geosci. Front.* **2023**, *14*, 101657. [CrossRef]
9. Chang, Z.; Huang, F.; Huang, J.; Jiang, S.-H.; Liu, Y.; Meena, S.R.; Catani, F. An Updating of Landslide Susceptibility Prediction from the Perspective of Space and Time. *Geosci. Front.* **2023**, *14*, 101619. [CrossRef]
10. Kavzoglu, T.; Colkesen, I.; Sahin, E.K. Machine Learning Techniques in Landslide Susceptibility Mapping: A Survey and a Case Study. In *Landslides: Theory, Practice and Modelling*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 283–301.
11. Fallah-Zazuli, M.; Vafaeinejad, A.; Alesheykh, A.A.; Modiri, M.; Aghamohammadi, H. Mapping Landslide Susceptibility in the Zagros Mountains, Iran: A Comparative Study of Different Data Mining Models. *Earth Sci. Inform.* **2019**, *12*, 615–628. [CrossRef]
12. Reichenbach, P.; Rossi, M.; Malamud, B.D.; Mihir, M.; Guzzetti, F. A Review of Statistically-Based Landslide Susceptibility Models. *Earth-Sci. Rev.* **2018**, *180*, 60–91. [CrossRef]
13. Badola, S.; Parkash, S. Landslide Susceptibility Mapping Using Machine Learning in Himalayan Region: A Review. In *Geo-Information for Disaster Monitoring and Management*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 123–143.
14. He, Y.; Huo, T.; Gao, B.; Zhu, Q.; Jin, L.; Chen, J.; Zhang, Q.; Tang, J. Thaw Slump Susceptibility Mapping Based on Sample Optimization and Ensemble Learning Techniques in Qinghai-Tibet Railway Corridor. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 5443–5459. [CrossRef]
15. Nhu, V.-H.; Shirzadi, A.; Shahabi, H.; Chen, W.; Clague, J.J.; Geertsema, M.; Jaafari, A.; Avand, M.; Miraki, S.; Asl, D.T.; et al. Shallow Landslide Susceptibility Mapping by Random Forest Base Classifier and Its Ensembles in a Semi-Arid Region of Iran. *Forests* **2020**, *11*, 421. [CrossRef]
16. Yu, L.; Wang, Y.; Pradhan, B. Enhancing Landslide Susceptibility Mapping Incorporating Landslide Typology via Stacking Ensemble Machine Learning in Three Gorges Reservoir, China. *Geosci. Front.* **2024**, *15*, 101802. [CrossRef]
17. Mohanty, B.; Sarkar, R.; Saha, S. Preparing Coastal Erosion Vulnerability Index Applying Deep Learning Techniques in Odisha State of India. *Int. J. Disaster Risk Reduct.* **2023**, *96*, 103986. [CrossRef]
18. Lee, S.-M.; Lee, S.-J. Landslide Susceptibility Assessment of South Korea Using Stacking Ensemble Machine Learning. *Geoenviron. Disasters* **2024**, *11*, 7. [CrossRef]
19. Arabameri, A.; Chandra Pal, S.; Rezaie, F.; Chakrabortty, R.; Saha, A.; Blaschke, T.; Di Napoli, M.; Ghorbanzadeh, O.; Thi Ngo, P.T. Decision Tree Based Ensemble Machine Learning Approaches for Landslide Susceptibility Mapping. *Geocarto Int.* **2022**, *37*, 4594–4627. [CrossRef]
20. Jafari, A.; Delavar, M.R.; Stein, A. Stacking-Based Uncertainty Modelling of Statistical and Machine Learning Methods for Residential Property Valuation. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *4*, 49–55. [CrossRef]
21. Liu, B.; Guo, H.; Li, J.; Ke, X.; He, X. Application and Interpretability of Ensemble Learning for Landslide Susceptibility Mapping along the Three Gorges Reservoir Area, China. *Nat. Hazards* **2024**, *120*, 4601–4632. [CrossRef]
22. Li, Y.; Yang, J.; Han, Z.; Li, J.; Wang, W.; Chen, N.; Hu, G.; Huang, J. An Ensemble Deep-Learning Framework for Landslide Susceptibility Assessment Using Multiple Blocks: A Case Study of Wenchuan Area, China. *Geomat. Nat. Hazards Risk* **2023**, *14*, 2221771. [CrossRef]
23. Le Minh, N.; Truyen, P.T.; Van Phong, T.; Jaafari, A.; Amiri, M.; Van Duong, N.; Van Bien, N.; Duc, D.M.; Prakash, I.; Pham, B.T. Ensemble Models Based on Radial Basis Function Network for Landslide Susceptibility Mapping. *Environ. Sci. Pollut. Res.* **2023**, *30*, 99380–99398. [CrossRef]
24. Matougui, Z.; Djerbal, L.; Bahar, R. A Comparative Study of Heterogeneous and Homogeneous Ensemble Approaches for Landslide Susceptibility Assessment in the Djebahia Region, Algeria. *Environ. Sci. Pollut. Res.* **2024**, *31*, 40554–40580. [CrossRef] [PubMed]
25. Hong, H. Assessing Landslide Susceptibility Based on Hybrid Best-First Decision Tree with Ensemble Learning Model. *Ecol. Indic.* **2023**, *147*, 109968. [CrossRef]
26. Tien Bui, D.; Ho, T.-C.; Pradhan, B.; Pham, B.-T.; Nhu, V.-H.; Revhaug, I. GIS-Based Modeling of Rainfall-Induced Landslides Using Data Mining-Based Functional Trees Classifier with AdaBoost, Bagging, and MultiBoost Ensemble Frameworks. *Environ. Earth Sci.* **2016**, *75*, 1–22. [CrossRef]
27. Zhao, F.; Miao, F.; Wu, Y.; Ke, C.; Gong, S.; Ding, Y. Refined Landslide Susceptibility Mapping in Township Area Using Ensemble Machine Learning Method under Dataset Replenishment Strategy. *Gondwana Res.* **2024**, *131*, 20–37. [CrossRef]

28. Lu, J.; He, Y.; Zhang, L.; Zhang, Q.; Gao, B.; Chen, H.; Fang, Y. Ensemble Learning Landslide Susceptibility Assessment with Optimized Non-Landslide Samples Selection. *Geomat. Nat. Hazards Risk* **2024**, *15*, 2378176. [CrossRef]
29. Ghayur Sadigh, A.; Alesheikh, A.A.; Bateni, S.M.; Jun, C.; Lee, S.; Nielson, J.R.; Panahi, M.; Rezaie, F. Comparison of Optimized Data-Driven Models for Landslide Susceptibility Mapping. *Environ. Dev. Sustain.* **2024**, *26*, 14665–14692. [CrossRef]
30. Arian, M.; Aram, Z. Relative Tectonic Activity Classification in the Kermanshah Area, Western Iran. *Solid Earth* **2014**, *5*, 1277–1291. [CrossRef]
31. Forestry and Watershed Organization. Available online: https://frw.ir (accessed on 1 September 2024).
32. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Selecting Optimal Conditioning Factors in Shallow Translational Landslide Susceptibility Mapping Using Genetic Algorithm. *Eng. Geol.* **2015**, *192*, 101–112. [CrossRef]
33. Hong, H.; Liu, J.; Zhu, A.-X. Modeling Landslide Susceptibility Using LogitBoost Alternating Decision Trees and Forest by Penalizing Attributes with the Bagging Ensemble. *Sci. Total Environ.* **2020**, *718*, 137231. [CrossRef]
34. Pham, Q.B.; Ekmekcioğlu, Ö.; Ali, S.A.; Koc, K.; Parvin, F. Examining the Role of Class Imbalance Handling Strategies in Predicting Earthquake-Induced Landslide-Prone Regions. *Appl. Soft Comput.* **2023**, *143*, 110429. [CrossRef]
35. Sharma, N.; Saharia, M.; Ramana, G. High Resolution Landslide Susceptibility Mapping Using Ensemble Machine Learning and Geospatial Big Data. *Catena* **2024**, *235*, 107653. [CrossRef]
36. Zeng, T.; Wu, L.; Peduto, D.; Glade, T.; Hayakawa, Y.S.; Yin, K. Ensemble Learning Framework for Landslide Susceptibility Mapping: Different Basic Classifier and Ensemble Strategy. *Geosci. Front.* **2023**, *14*, 101645. [CrossRef]
37. Vahidnia, M.H.; Alesheikh, A.A.; Alimohammadi, A.; Hosseinali, F. A GIS-Based Neuro-Fuzzy Procedure for Integrating Knowledge and Data in Landslide Susceptibility Mapping. *Comput. Geosci.* **2010**, *36*, 1101–1114. [CrossRef]
38. Kursa, M.B.; Rudnicki, W.R. Feature Selection with the Boruta Package. *J. Stat. Softw.* **2010**, *36*, 1–13. [CrossRef]
39. Heidari, A.A.; Akhoondzadeh, M.; Chen, H. A Wavelet PM2. 5 Prediction System Using Optimized Kernel Extreme Learning with Boruta-XGBoost Feature Selection. *Mathematics* **2022**, *10*, 3566. [CrossRef]
40. Yuan, X.; Chen, F.; Xia, Z.; Zhuang, L.; Jiao, K.; Peng, Z.; Wang, B.; Bucknall, R.; Yearwood, K.; Hou, Z. A Novel Feature Susceptibility Approach for a PEMFC Control System Based on an Improved XGBoost-Boruta Algorithm. *Energy AI* **2023**, *12*, 100229. [CrossRef]
41. Jang, J.-S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [CrossRef]
42. Alesheikh, A.A.; Chatrsimab, Z.; Rezaie, F.; Lee, S.; Jafari, A.; Panahi, M. Land Subsidence Susceptibility Mapping Based on InSAR and a Hybrid Machine Learning Approach. *Egypt. J. Remote Sens. Space Sci.* **2024**, *27*, 255–267. [CrossRef]
43. Cortes, C.; Vapnik, V. Support-Vector Networks. In *Machine Learning*; AT&T Bell Labs.: Hohndel, NJ, USA, 1995.
44. Shikhteymour, S.R.; Borji, M.; Bagheri-Gavkosh, M.; Azimi, E.; Collins, T.W. A Novel Approach for Assessing Flood Risk with Machine Learning and Multi-Criteria Decision-Making Methods. *Appl. Geogr.* **2023**, *158*, 103035. [CrossRef]
45. Drucker, H.; Burges, C.J.; Kaufman, L.; Smola, A.; Vapnik, V. Support Vector Regression Machines. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1996; Volume 9.
46. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme Learning Machine: Theory and Applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]
47. Adnan, R.M.; Mostafa, R.R.; Kisi, O.; Yaseen, Z.M.; Shahid, S.; Zounemat-Kermani, M. Improving Streamflow Prediction Using a New Hybrid ELM Model Combined with Hybrid Particle Swarm Optimization and Grey Wolf Optimization. *Knowledge-Based Syst.* **2021**, *230*, 107379. [CrossRef]
48. Safiri, S.; Nikoofard, A. Ladybug Beetle Optimization Algorithm: Application for Real-World Problems. *J. Supercomput.* **2023**, *79*, 3511–3560. [CrossRef]
49. Zhao, W.; Wang, L.; Zhang, Z.; Fan, H.; Zhang, J.; Mirjalili, S.; Khodadadi, N.; Cao, Q. Electric Eel Foraging Optimization: A New Bio-Inspired Optimizer for Engineering Applications. *Expert Syst. Appl.* **2024**, *238*, 122200. [CrossRef]
50. Mehta, P.; Yildiz, B.S.; Sait, S.M.; Yıldız, A.R. Optimization of Electric Vehicle Design Problems Using Improved Electric Eel Foraging Optimization Algorithm. *Mater. Test.* **2024**, *66*, 1230–1240. [CrossRef]
51. Razavi Termeh, V. Modeling Asthma Prone Areas Using Machine Learning with Emphasis on Data Management (in Persian). Ph.D. Thesis, K.N. Toosi University of Technology: Tehran, Iran, 2021.
52. Junninen, H.; Niska, H.; Tuppurainen, K.; Ruuskanen, J.; Kolehmainen, M. Methods for Imputation of Missing Values in Air Quality Data Sets. *Atmos. Environ.* **2004**, *38*, 2895–2907. [CrossRef]
53. Zandi, I.; Pahlavani, P.; Bigdeli, B. Preparation of Flood Susceptibility Map Using Multi-Criteria Spatial Analysis and Data Fusion (A Case Study: Maneh and Samalqan County). *J. Geomat. Sci. Technol.* **2023**, *12*, 53–76. [CrossRef]
54. Ahmadlou, M.; Karimi, M.; Alizadeh, S.; Shirzadi, A.; Parvinnejhad, D.; Shahabi, H.; Panahi, M. Flood Susceptibility Assessment Using Integration of Adaptive Network-Based Fuzzy Inference System (ANFIS) and Biogeography-Based Optimization (BBO) and BAT Algorithms (BA). *Geocarto Int.* **2019**, *34*, 1252–1272. [CrossRef]
55. Jafari, A.; Alesheikh, A.A.; Rezaie, F.; Panahi, M.; Shahsavar, S.; Lee, M.-J.; Lee, S. Enhancing a Convolutional Neural Network Model for Land Subsidence Susceptibility Mapping Using Hybrid Meta-Heuristic Algorithms. *Int. J. Coal Geol.* **2023**, *277*, 104350. [CrossRef]
56. Huan, Y.; Song, L.; Khan, U.; Zhang, B. Stacking Ensemble of Machine Learning Methods for Landslide Susceptibility Mapping in Zhangjiajie City, Hunan Province, China. *Environ. Earth Sci.* **2023**, *82*, 35. [CrossRef]

57. Li, W.; Fang, Z.; Wang, Y. Stacking Ensemble of Deep Learning Methods for Landslide Susceptibility Mapping in the Three Gorges Reservoir Area, China. *Stoch. Environ. Res. Risk Assess.* **2021**, *8*, 2207–2228. [CrossRef]
58. Benbouras, M.A. Hybrid Meta-Heuristic Machine Learning Methods Applied to Landslide Susceptibility Mapping in the Sahel-Algiers. *Int. J. Sediment Res.* **2022**, *37*, 601–618. [CrossRef]