Using Hybrid Artificial Intelligence for Landslide Modeling

Solmaz Abdollahizad[1*], Kayvan Asghari[2],Majid Samadzamini[3]
[1,2,3] Department of Computer Engineering, Sardroud Center, Tabriz Branch, Islamic Azad University, Tabriz, Iran
Email: Solmaz.abdollahizad@iau.ac.ir (Corresponding Author)[1*] ,k.asghari@iau.ac.ir[2],zamini.m@iau.ac.ir[3]

Receive Date: 13 December 2024 Revise Date: 3 January 2025 Accept Date: 27 January 2025

Abstract

The current analysis exploited a hybrid ANFIS model that was optimized using PSO, GWO, and SFLA, three evolutionary algorithms. Three common models, MLP, RF, and SVM, were used to test and evaluate their performance on the same training and validation datasets, to build a LSM in EAP, Iran. For analyzing the associations between landslides and landslide conditioning variables, the PCF model was exploited as a bivariate statistical test. Furthermore, in the present analysis, the Pearson correlation test was used to measure the predictive strength of ten landslide condition variables. The fuzzy c-means clustering approach was then used to construct an initial fuzzy inference system for LSM. In addition, three wise algorithms, namely GWO, SFLA, and PSO, were used to train the ANFIS in the current analysis. One of the most significant benefits of these approaches is that they improve precision by optimizing and calculating ANFIS parameters. Indeed, it has the potential to reduce dimension dangers and the problems of local minimum, thus improving the ANFIS model accuracy. Lastly, ROC curves were used to test the LSMs generated by ANFIS-GWO, ANFIS-SFLA, and ANFIS-PSO. According to the results, the AUC values for the ANFIS-PSO, ANGIS-SFLA, and ANFIS-GWO models were 0.89, 0.88, and 0.88, respectively.

Keywords: Machine learning, Landslide, ANFIS

1. Introduction

Landslides are complex natural disasters that are frequently initiated several fatalities and casualties globally. These occurrences have the potential to risk the lives of people and infrastructures of the nations in various areas around the world with immense social-economic on sequences [1]. To this end, pinpointing the zones with landslide vulnerability is an effective technique to avert and decrease plausible damages.

Landslide susceptibility modeling is extensively acknowledged, and prediction accuracy outcomes highly rely on the exploited data quality, conditioning factors, environmental conditions, topographic features of the region, and landslide inventory. Therefore, LSM requires a multi-criteria approach and high levels of accuracy and reliability in the resulting maps in order to be relevant for decision-making and the design of disaster management plans [2].

A review of research background signifies that various methods of preparing hazard maps and landslide susceptibility have recently been established using statistics, deterministic, and heuristic Statistical models, such as multivariate analysis, weights of evidence[3], Probabilistic models e.g., FR[4] evidential belief function [5], analytical hierarchy process [6], and Certainty factor [7]have been exploited by a great deal of the aforementioned studies. The heuristic method is developed in accordance with experts-related ideas and experiences with the intention of assigning various weights to different influencing factors. Nevertheless, the method would not produce satisfactory results due to the limited study area data and low rate of reproducibility [4].

Increasing applications of data mining and machine learning algorithms have been reported in landslide susceptibility assessments, involving fuzzy logic algorithms, artificial neural networks (ANN) and evolutionary population-based algorithms [8]. As reported in previous studies machine learning models in most cases outperform conventional methods as they appear sufficient in handling non-linear data with different scales and from different type of sources [9]. Also, it has been well established that the integration of conventional statistical methods and machine learning methods in most cases perform better than individual machine learning techniques in susceptibility assessments. For instance, models such as Naive Bayes [10], Multi-Layer Perceptron [11], Decision Tree [12], Neuro-Fuzzy [13], Support Vector Machine [14], Logistic Regress in [15], and Reduced Error Pruning Trees [16] are among the so-called ML methods. Therefore, ML approaches are considered to be promising in spatial prediction of landslide.

Nowadays, deep learning neural network (DLNN) is obtaining remarkable state-of-the art accomplishment for LSM [17]. It is rapidly getting popular in LS domain due to the production of precisely spatial information from the raw input. Lots of researches have been developing on several landslide occurrence prediction using DLNN such as deep convolutional neural network [15]. Considering that lots of synthetic data and very expensive resources are needed to execute the deep learning approaches.[18], reported that DLNN's used for detecting landslides do not automatically outperform ML methods, as they strongly depend by the depth of the layers, the input window sizes and training strategies. The high bias and over-fitting problem are major issues in deep learning algorithm due to nature of data. Although the DLNN model could be considered as an alternative approach with highly predictive accuracy, there are several limitations that should be considered. The most important one that may influence the decision of using DLNN models in landslide assessments is the process of tuning the structural parameters. Finding the optimal number of hidden layers and processing elements is a very computational demanding task.

The idea of integrating state-of-the art models, such as artificial neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS), and metaheuristic optimization algorithms with the aim of developing enhanced models has been suggested by many researchers. Regarding the ANFIS model, several integrations of this model with GA[19], ant colony optimization (ACO), differential evolution (DE) [19], and biogeography-based optimization (BBO) [20] have been suggested to obtain accurate estimates of natural hazards. Therefore, Fuzzy reasoning was proposed in order to address the issue [21]. Nevertheless, since fuzzy member-ship values are determined subjectively, the approach does not provide results with high precision. As a result, modern landslide prediction models must be developed in order to cope with imprecisions and uncertainties and enhance landslide prediction capabilities [22].

However, despite widespread applications of the machine learning methods, some drawbacks exist in these methods that limit their performance [23]. The main weakness of the machine learning methods is that their application involves tuning of several parameters that make them challenging and time-consuming to apply [1]. While, for many years, the modelers either had to manually tune the parameters during a time-consuming trial and error process or for simplicity they used the default settings which are usually far from optimal, metaheuristic optimization algorithms have recently emerged as a remedy to alleviate the difficulty associated with the machine learning methods [1]. The optimization algorithms provide an intelligent framework to automatically and properly define the parameters of the base model.
Although plenty of methods and models have been exploited to create maps of landslide susceptibility by geographic information systems (GIS), there is not a compromised method to be accepted as the most appropriate one due to the possible limitations of the qualitative techniques caused by unplanned occurrences or inadequate knowledge upon which the expert decisions are centered on [24]. Conversely, inaccuracy, and imprecision of data are among the shortfalls of quantitative methods [25]. As indicated, due to a number of methods and techniques, recognizing and identifying the most efficient methods and techniques are still challenging.

2. Background

For the sake of clarity, this section will briefly introduce the background concepts of machine learning and landslides prevention.

2.1. Machine learning

Machine Learning algorithms are mainly divided into four categories: Supervised learning, unsupervised learning, Semi-supervised learning, and Reinforcement learning, as shown in Fig. 1. In the following, we briefly discuss each type of learning technique with the scope of their applicability to solve real-world problems [26].

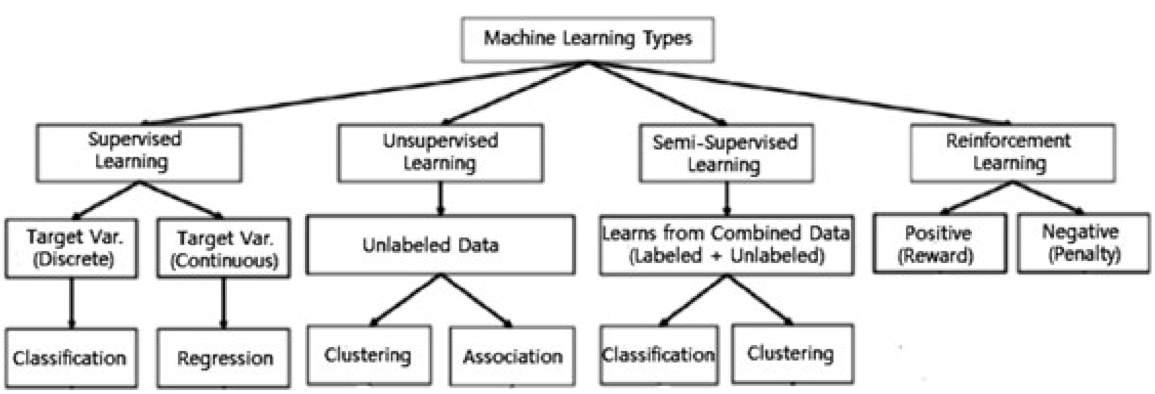

Fig.1. Type of ML

The difference between supervised learning and unsupervised learning is that supervised learning requires labeled training data. Unsupervised machine learning involves analyzing data that has not been labeled or processed, while supervised machine learning involves training using data that has been tagged at both its entrance and its exit. During the process of machine learning, the model is responsible for discovering the link between the labeled input data and the output data. Models are refined until they can accurately predict the outcomes of experiments using data that has not yet been collected. The production of labeled training data often calls for a significant investment of resources. Unsupervised machine learning acquires knowledge via the use of raw training data that has not been labeled. They are often used in the process of discovering the underlying trends in a given dataset since unsupervised models can discover relationships and patterns within unlabeled datasets [27].

The issue that the approach is used to address. Unsupervised learning is typically used in the process of determining connections across databases while supervised machine learning is often used to organize information or select a factor. Supervised machine learning requires labeled data; hence it requires a significant increase in the number of resources. There is less human supervision when it comes to unsupervised machine learning, therefore it may be more difficult to reach the proper levels of understandability[27].

2.1.1. Supervised learning

Supervised machine learning requires labeled input and output data during the training phase of the ML workflow to be effective. A data scientist will often label the training data when preparing to train and assess a model. Once a system has learned the connection between input and output data, it may be used to classify and predict data that has not yet been examined by researchers [28, 29]

The term "supervised machine learning" denotes the fact that at least part of this method requires human monitoring. The vast majority of data provided is unlabeled and unprocessed. Supervised learning refers to building a model for connecting known inputs to unknown outputs. Consequently, the output values for new data can be predicted based on those relationships learned from the previously labeled training data. Supervised learning can be divided into classification and regression problems. In classification problems, the intended output is a semantic label or class. For example, to identify potential landslides, classification problems would label each pixel in an image as ‘landslide’ or ’non-landslide’. Regression problems aim to predict a continuous variable. Common supervised learning algorithms include the logistic regression (LR), decision tree (DT), support vector machine (SVM), Naive Bayes (NB), artificial neural networks,(RF) random forest

– LR: A supervised learning algorithm that uses a logistic function to map the input variables to categorical dependent variables.

– DT: A supervised learning algorithm is commonly used in classification problems. The structure resembles a tree. The branch node represents several alternatives. Each leaf node represents a decision.
– SVM: A supervised learning algorithm is also commonly used in classification problems by constructing a separating line to distinguish between objects in a multidimensional space.
– NB: A supervised learning algorithm is based on Bayes' theorem and widely used in classification problems, which assumes that features are independent and have no correlations.
– ANN: ANN consists of a set of connected processing units that work together, can found an association of patterns among input and output.
kNN: A supervised learning algorithm uses 'feature similarity' to predict the values of new data points, in which the new data point will be assigned a value based on the distance it matches the points in the training set.
-RF: Random Forest is a commonly-used machine learning that combines the output of multiple decision trees to reach a single result. Its ease of use and flexibility have fueled its adoption, as it handles both classification and regression problems.

2.1.2. Unsupervised Machine Learning

Unsupervised learning adopts a more latent strategy as opposed to directed AI. For instance, the number of groups focuses will be chosen by an individual; however, the model will dissect gigantic measures of information productively and without human management.

Unsupervised AI is subsequently very much adjusted to give replies to questions concerning stowed-away examples and associations in the actual information. The tremendous heft of the information that is open is unlabeled, crude information. Unsupervised gaining is an intense method for making inferences from this information by putting together the information into bunches because of shared credits or looking at data sets for hidden designs. The prerequisite for labeled information, be that as it may, makes directed AI more asset-concentrated [30].
– K-Means clustering: An unsupervised learning algorithm divides all input data into k clusters, in which data in the same cluster are as similar to each other as possible.
- Clustering: clustering is designed to group unlabeled examples based on their similarity to each other. (If the examples are labeled, this kind of grouping is called classification.

2.1.3. Semi-supervised:

Semi-supervised machine learning is a combination of supervised and unsupervised machine learning methods. It can be fruit-full in those areas of machine learning and data mining where the unlabeled data is already present and getting the labeled data is a tedious process. With more common supervised machine learning methods, you train a machine learning algorithm on a "labeled" dataset in which each record includes the outcome information.
-**Transudative SVM**: Transudative support vector machines (TSVM) has been widely used as a means of treating

partially labeled data in semi supervised learning [26].

-Self-Training: In self-training, a classifier is trained with a portion of labeled data. The classifier is then fed with unlabeled data. The unlabeled points and the predicted labels are added together in the training set. This procedure is then repeated further. Since the classifier is learning itself, hence the name self-training [26].

2.1.4. Reinforcement:

Reinforcement learning is a type of machine learning algorithm that enables software agents and machines to automatically evaluate the optimal behavior in a particular context or environment to improve its efficiency, i.e., an *environment-driven approach*. This type of learning is based on reward or penalty, and its ultimate goal is to use insights obtained from environmental activists to take action to increase the reward or minimize the risk [31]. It is a powerful tool for training AI models that can help increase automation or optimize the operational efficiency of sophisticated systems such as robotics, autonomous driving tasks, manufacturing and supply chain logistics, however, not preferable to use it for solving the basic or straightforward problems [26].

2.2. Deep learning methods

Deep learning (DL) is playing an increasingly important role in our lives. It has already made a huge impact in areas, such as cancer diagnosis, precision medicine, self-driving cars, predictive forecasting, and speech recognition. The painstakingly handcrafted feature extractors used in traditional learning, classification, and pattern recognition systems are not scalable for large-sized data sets. In many cases, depending on the problem complexity, DL can also overcome the limitations of earlier shallow networks that prevented efficient training and abstractions of hierarchical representations of multi-dimensional training data.

Deep neural network (DNN) uses multiple (deep) layers of units with highly optimized algorithms and architectures. implementations. The review also covers different types of deep architectures, such as deep convolution networks, deep residual networks, recurrent neural networks, reinforcement learning, variation auto encoders, and others. Deep neural network consists of several layers of nodes. Different architectures have been developed to solve problems in different domains or use-cases. E.g., CNN is used most of the time in computer vision and image recognition, and RNN is commonly used in time series problems/forecasting. On the other hand, there is no clear winner for general problems like classification as the choice of architecture could depend on multiple factors. Nonetheless [32] evaluated 179 classifiers and concluded that parallel random forest or parRF_t, which is essentially parallel implementation of variation of decision tree, performed the best. Below are four of the most common architectures of deep neural networks. **

1. Convolution Neural Network
2. Auto encoder
3. Restricted Boltzmann Machine (RBM)
4. Long Short-Term Memory (LSTM)**

-Convolution Neural Network: CNN is based on the human visual cortex and is the neural network of choice for computer vision (image recognition) and video recognition. It is also used in other areas such as NLP, drug discovery, etc. a CNN consists of a series of convolution and sub-sampling layers followed by a fully connected layer and a normalizing (e.g., SoftMax function) layer [31].

- Auto encoder: is a neural network that uses unsupervised algorithm and learns the representation in the input data set for dimensionality reduction and to recreate the original data set. The learning algorithm is based on the implementation of the backpropagation

-Restricted Boltzmann Machine (RBM): RBM is an artificial neural network where we can apply unsupervised learning algorithm to build non-linear generative models from unlabeled data [33]. The goal is to train the network to increase a function (e.g., product or log) of the probability of vector in the visible units so it can probabilistically reconstruct the input. It learns the probability distribution over its inputs. RBM is made of two-layer network called the visible layer and the hidden layer. Each unit in the visible layer is connected to all units in the hidden layer and there are no connections between the units in the same layer.

Long Short-Term Memory (LSTM): LSTM is an implementation of the Recurrent Neural Network and was first proposed by Hoch Reiter et al. in 1997 [34]. Unlike the earlier described feed forward network architectures, LSTM can retain knowledge of earlier states and can be trained for work that requires memory or state awareness. LSTM partly addresses a major limitation of RNN, i.e., the problem of vanishing gradients by letting gradients to pass unaltered. LSTM consists of blocks of memory cell state through which signal flows while being regulated by input, forget and output gates. These gates control what is stored, read and written on the cell. LSTM is used by Google, Apple and Amazon in their voice recognition platforms**

2.3. Landslide

A geo-hazard is a devastating phenomenon that is directly and indirectly caused by activity in the earth's interior or geological environment changes, including human activity or climate change. As one type of global geo-hazard, landslides are geological phenomena related to ground movements of rock fall and debris flow and can refer to the movement of a mass of rock, debris, or earth down a slope under the influences of gravity, rainfall, and earthquake. Lithology, tectonics, climate change, and anthropogenic pressure may cause slope instability that could progress to landslides [35]. Heavy rainfall, rapid snowmelt, or earthquakes could also trigger a landslide occurrence. Landslides are ubiquitous in any terrestrial environment with slopes.

In most cases, landslide occurrence means catastrophic results.it has brought out the massive destruction of infrastructure and even thousands of fatalities every year [36]. From 2004 to 2010, 2620 fatal landslides were recorded, causing 32,322 fatalities. At least 17% of all natural-hazard fatalities around the world can be attributed to landslides. In the most affected areas financial costs and countermeasures are on the order of billions of dollars.

Recently, as a consequence of human disturbance (e.g. Deforestation, mineral mining, and intensive exploitation of land for construction) and extreme weather, the frequency and intensity of landslides have increased dramatically. With the advent of extreme natural events, the prevention of landslides has become an urgent task. landslides prevention involves an assessment of slope instability phenomena and the change in the occurrence of slopes by means of effective geological engineering principles and other existing and emerging technologies. landslides prevention can provide valuable information for government agencies, planners, decision makers, and local landowners to make emergency plans that reduce the negative effects on economics and human life. Typically, the study of landslides prevention is divided into two aspects: detection and prediction.

Related datasets for landslides prevention are generally obtained from three sources: (i) remotely sensed data acquired by Earth-observing satellites, (ii) data collected by in situ sensors, and (iii) data collected during fieldwork [37].

2.3.1. Typical data source for landslides prediction

In the present study, the ESRI ArcGIS 10.3 software was utilized with the intention of producing and displaying the data layers. All the layers of data were organized in raster format with a pixel size of 30 (m)×30 (m). The influencing factors were obtained from ASTER Global DEM , the geological map, and a topographic map with the same resolution [38]. Then the geological map was used in order to obtain the lithology map, which was then converted into a raster format. Landsat 8 (OLI images) were used in order to derive the NDVI as well as land use maps. For further analyses, all the so-called factors were standardized by a similar scale of 30 ×30 (m^2). Besides, the conditioning factors that were of continuous data were reclassified into distinct subsections with the intention of transforming continuous data to sections at specific intervals. In order to achieve the identical output scaling, the other discrete conditioning variables were reclassified into groups (Fig. 2). For training/modelling, 536 (70%) landslide locations were utilized in the present analysis, and 230 (30%) landslides were utilized for validating. A value of "1" was allocated to the landslide training instances[39]. Moreover, from the landslide-free zones, a similar amount of non-landslide points (766) was randomly produced, and a value of '0' was allocated to these instances, which were randomly divided into two sections with a ratio of 70/30 as well [40].

3. ML modeling

3.1. Random Forest (RF)

RF is regarded as an ensemble learning method that classifies unidentified samples predicated on the combined outcomes of a series of weak classifications Trees developed via bootstrapping techniques. Particularly, the learning process includes choosing the predictor variable for each iteration and resampling the data through replacement]. By means of this approach, an RF model demonstrates a higher efficient ability to prevent overfitting problems and by and large present a superior generalization output.

In a randomized forest, the best split amid a subcategory of predictors, which are haphazardly selected by the node is used to split each node. Inherently, inside huge datasets, it has been a prominent technique for identifying beneficial hitherto invisible patterns. There are n variables that could be selected as random subsections from the training data with the intention of determining the best possible node to split. The best node division could be finalized utilizing Gini criteria [41].

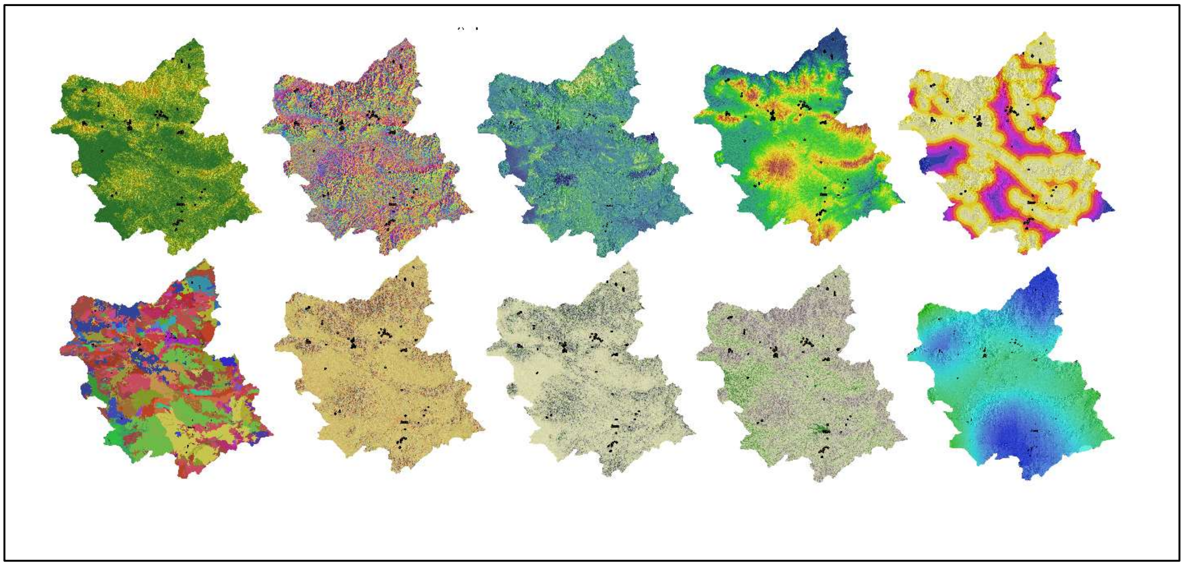

Fig.2. Ten condition factors map

3.2.Support Vector Machine (SVM)

SVM is a series of techniques for ML on the basis of the notion of an optimum hyperplane of separation. In feature space, SVM considers the widest margin between the two groups. A standard SVM model could be of a two-class or multi-class model (an amalgamation of a two-class SVMs chain). The most widely used form of ML is the two-class SVM. The separating hyperplane is among the possible planes, which divides two groups during the model performance. The line L being the in-between classification line, and L1 and L2 being lines running parallel to L across the sample points nearest to the classification line. The classification margin is considered to be the distance between them. The purpose of the optimum hyperplane classification is to correctly distinguish between the two types of samples while maximizing the support vector margin. In general, there are two categories of SVMs on the basis of object classification, viz. the two-class and multiclass. The multiclass SVM is a synthesis of a set of two-class SVMs [42]. Presently, pairwise classification and the one-to-the-other-class approach are prevalent multi-class SVM methods. The most commonly used technique is two-class SVM. Normally, along with grouping, SVM could be exploited for regression analysis.

3.3. Artificial Neural Network (ANN)

ANN are units for the processing of computational information influenced by the behavior and structure of actual biological neurons whose architecture attempts to mimic the human brain cells' acquisition of knowledge and organizational skills [43].

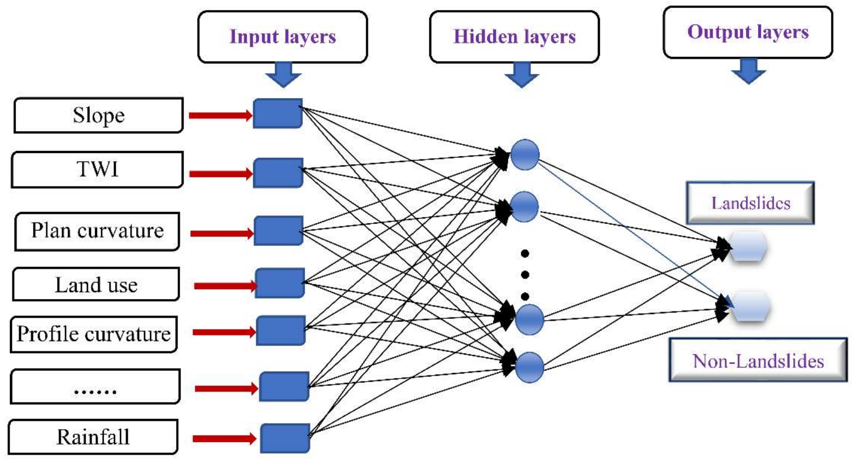

Fig.3. The layers of MLP

A three-layer feed-forward MLP has been constructed (Fig. 3). Following trial and error and cogitating the lowest error, the optimal network architecture was chosen. Initial weights in a limited range were initiated by random. Once the stopping error criteria are met, the process is terminated. The first objective in this analysis was to fulfill the stopping criteria of the root mean square error (RMSE). If RMSE is not attained, the epochs number could then be used as a termination criterion, which was set to be 1000 in the current analysis.

3.4. Hybrid modelling

3.4.1. Neuro-Fuzzy inference system

ANFIS is a hybrid paradigm for inferring associations between inputs and outputs on the basis of artificial neural networks (ANN) and fuzzy logic. The neural-fuzzy framework with TSK inference engine that is proposed by. The so-called fuzzy inference engine was chosen due to its capability in modeling complex nonlinear problems with extra accuracy and with lower number of rules compared to other fuzzy inference engines (e.g., the Mamdani engine). The five layers of the neural fuzzy structure (Fig. 4) include two groups of nodes: fixed and adaptive nodes. The first and forth layers were fabricated with adaptive nodes, while set nodes are used to design the remaining layers. Connection weights were regulated during training phases in adaptive nodes with the intention of fitting to training data, while fixed nodes simply sum or normalize all incoming signals.

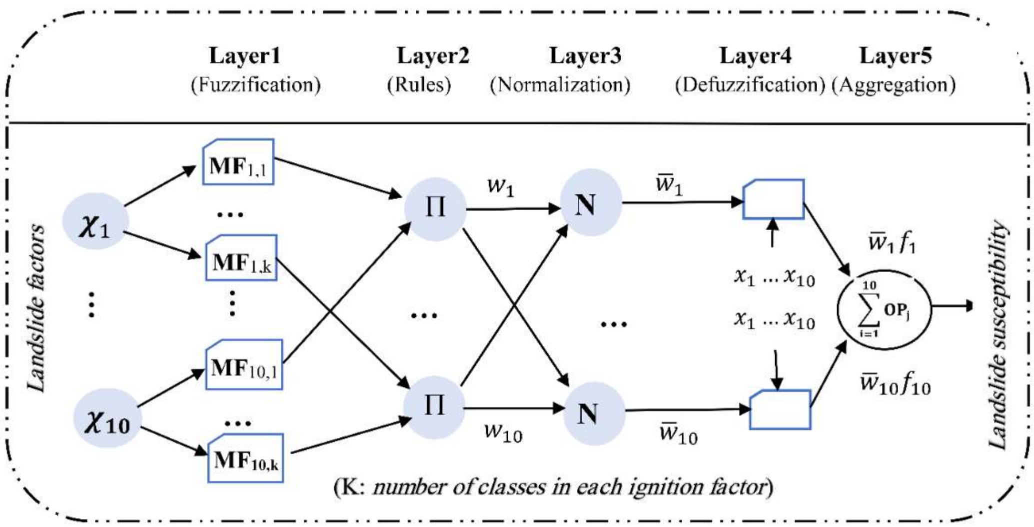

Fig.4. ANFIS structure

Figure 5 depicts the final configuration of the neural fuzzy model used in the present analysis. As it can be observed, the model holds ten input variables, a single output, and 12 rules possessed with 240 antecedent and 132 consequent parameters. Three modeling methods were utilized to evaluate the best values for the so-called parameters during the training phase.

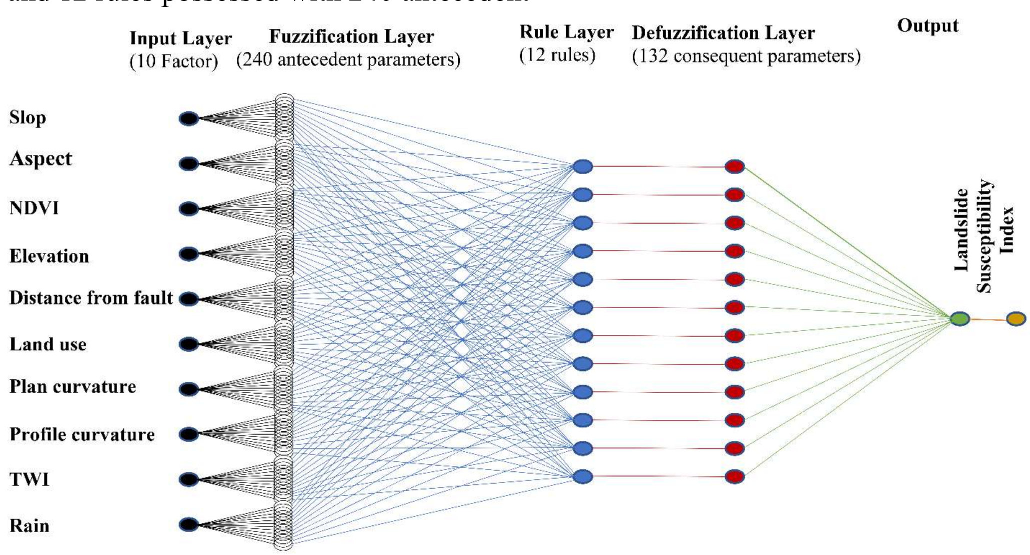

Fig.5. Proposed fuzzy model

3.4.2. Optimization algorithms

- **SFLA**

The SFLA is one of the renowned meta-heuristic algorithms. By merging the Memetic algorithm and the PSO, this algorithm has achieved to prominence in local and global search. Furthermore, owing to its simplicity, fast convergence speed, and global search capabilities, the algorithm is nowadays recognized by researchers as an effective optimization algorithm for discrete data.

The development of the initial population is the first step in the implementation of the present algorithm, as it is with every other evolutionary algorithm. This algorithm's population consists of a group of frogs, each of those frogs is used as solution to a problem and is looking for food [8]. The position of the frog i is demonstrated as Xi= (xi1, xi2, …, xiD). Initially, p number of frogs are distributed into n memeplexes. This set of frogs that have a similar structure but vary in their adaptability. As a result, the frogs' descending ordering is operated based on their suitability, in which the first frog goes to the first memeplex from the first ordered list, the next frog goes to the next memeplex, and the n+1 frog goes to the first memeplex. This pattern will continue until there are no more frogs left to assign to memeplexes. The best and worst frog positions for either of the memeplex are defined by Xb and Xw (Fig. 6).

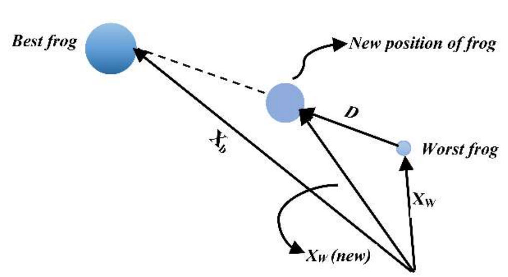

Fig.6. The best frog leaping in SFLA

This upgrade process is replicated before the update number has been satisfactory enough. In all memeplexes, both frogs are mixed and reordered into M memeplexes following the completion of the local area deep-searching of all memeplexes. Furthermore, once the local search fails to identify superior solutions with each iteration, random virtual frogs are created and replaced in the population with the aim of allowing the randomly production of knowledge. The shuffling and local search processes proceed until the specified convergence requirements are reached. The ultimate purpose of the process is to discover the right global optimal solutions.

- **GWO**

This algorithm is a population-based optimization method. Alpha, beta, delta, and omega are the four leadership groups in a pack of grey wolves. The fittest approach to an optimization problem is the alpha. The search process (i.e., locating a prey) starts with the generation of a random population from the applicant solutions, which is then modeled in order to simulate the hunting behavior of the wolves' pack. The wolves predict the possible location of the prey during the iterations of the algorithm. Flowingly, they calculate their distances from the prey and change their location depending on the prey. The wolves one by one represent a potential solution that changes during the hunt. Additionally, GWO frequently employs efficient operations to prevent being trapped in a local optimum in order to reach the global optimum (i.e., prey), as follows [42]:

$$\vec{X}(t+1) = \vec{X}_{P(t)} - \vec{A}. \mid \vec{C}.\vec{X}_{P(t)} - \vec{X}_{(t)} \mid \quad (1)$$

$$(\vec{A} = 2\vec{a}.\vec{r}_i - \vec{a}, \vec{C} = 2.\vec{r}_2)$$

Where $\vec{X}$ is the position vector of an applicant solution which is called grey wolf, t is the existing iteration, $\vec{X}_p$denotes the global optimum's position vector which is called prey, $\vec{r}_{1,2}$ denotes random vectors between 0 and 1 which enable wolves to achieve every position in the search space, and throughout the iterations, the constituents of $\vec{a}$are reduced linearly from 2 to 0 to highlight the exploration (which is searching for prey) and exploitation (which is targeting the prey):

$$\vec{a} = 2 - \frac{2t}{Max\_Iter} \tag{2}$$

Where *t* is the existing epoch and *Max_Iter* is the overall number of iterations. Throughout the exploration phase, he potential updating of positions could be given for other wolves consistent with the α, β, and δ positions:

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 .\mid \vec{C}_1 . \vec{X}_\alpha - \vec{X} \mid \tag{3}$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 .\mid \vec{C}_2 . \vec{X}_\beta - \vec{X} \mid \tag{4}$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 .\mid \vec{C}_3 . \vec{X}_\delta - \vec{X} \mid \tag{5}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1(t) + \vec{X}_2(t) + \vec{X}_3(t)}{3} \tag{6}$$

Once the termination condition is met, the GWO algorithm's search process ends [26].

- **PSO**

PSO is an optimization technique of stochastic type based on a population. It has been extensively employed to aid and solve the rapid convergence problem and global solutions in the majority of modern scientific and engineering optimization problems [44]. It uses an objective function and three mechanisms in order to discover an optimal solution by mimicking swarm behavior in tasks performances, such as that of bird flocks and fish. Each particle moves through the search space in quest of the best landing spot based on its own and neighboring particles' experiences (Fig. 7).

At each iteration *iter*, two pieces of data characterize each of *i* particle on the search domain: its velocity V_i^{iter} and its position P_i^{iter}, which are updated as follows:

$$\begin{cases} V_i^{iter+1} = \omega^{iter} V_i^{iter} + c_1 r_1 \left(P_{i,Best}^{iter} - P_i^{iter}\right) + c_2 r_2 \left(P_{Global}^{iter} - P_i^{iter}\right) \\ P_i^{iter+1} = P_i^{iter} + V_i^{iter+1} \\ iter = 1{:}MaxIt, \;\; i = 1{:}N \end{cases} \tag{7}$$

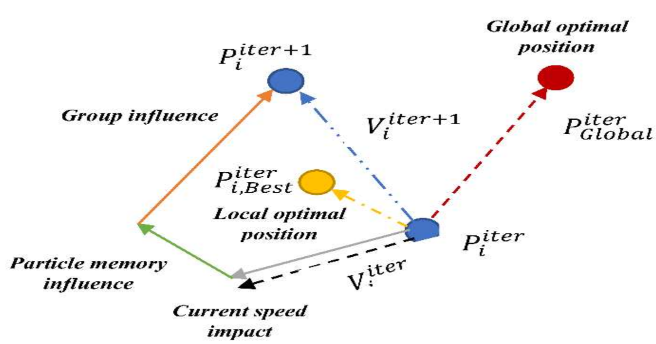

Fig.7. PSO update process

Where N is the particles' number existing in the swarm, and MaxIt is the maximum number of iterations. The inertia weight is referred to as w_iter. The term V_i_iter refers to the particle's preceding velocity. It functions as a momentum, stopping the particle from making radical changes of direction and biasing it in the direction it is actually in. The cognitive part of the second term, c_1 r_1 (P_(i,Best)^iter-P_i^iter), denotes the particle's personal experience. The particle is moved to its superior location as a result of the so-called term by which models its inclination to go back to locations that were of highest satisfaction in the past. The third term, c_2 r_2 (P_Global^iter-P_i^iter), denotes particles cooperation which is also known as social comportment. This term has the effect of attracting each particle to the best position sought by its neighbors.

3.5. Development of the models

Due to the prominence of ANFIS parameters' fine-tuning in reaching a global optimum, the PSO, SFLA, and GWO metaheuristic algorithms were employed in order to search for the most appropriate variables and develop three ANFIS types of the model which are of meta optimized type, namely ANFIS-GWO, ANFIS- PSO, and ANFIS-SFLA. With the aim of doing so, a primary fuzzy inference system was firstly created by means of training dataset and the c-means clustering method [45]. The GWO, PSO, and SFLA algorithms were then utilized in order to experiment with different combinations of the neuro-fuzzy system's basic parameters. Lastly, the basic parameter values were stored and the best values were calculated on the basis of the lowest MSE and RMSE values (Fig. 8), which were computed using n (the number of samples) and T_i (target values) existing in the validation or training dataset, as well as the landslide models' O_i (output values). The ultimate landslide models were established and employed with the aim of computing the susceptibility index for either of the pixels in the study region by means of the optimum unification of the ANFIS model's basic parameters.

ANFIS antecedent and consequent parameters were shown in Table 1 and Table 2 shows the final setting parameters of ANFIS-GWO, ANFIS- PSO, and ANFIS-SFLA.

Table 1. The ANFIS parameters a) antecedent b) consequent

(a)

$c_{1,1}$	$\sigma_{1,1}$	$c_{2,1}$	$\sigma_{2,1}$	…	$c_{10,1}$	$\sigma_{10,1}$
$c_{1,2}$	$\sigma_{1,2}$	$c_{2,2}$	$\sigma_{2,2}$	…	$c_{10,2}$	$\sigma_{10,2}$
…	…	…	…	…	…	…
$c_{1,12}$	$\sigma_{1,12}$	$c_{2,12}$	$\sigma_{2,12}$	…	$c_{10,12}$	$\sigma_{10,12}$

(b)

$p_{0,1}$	$p_{1,1}$	…	$p_{10,1}$
$p_{0,2}$	$p_{1,2}$	…	$p_{10,2}$
…	…	…	…
$p_{0,12}$	$p_{1,12}$	…	$p_{10,12}$

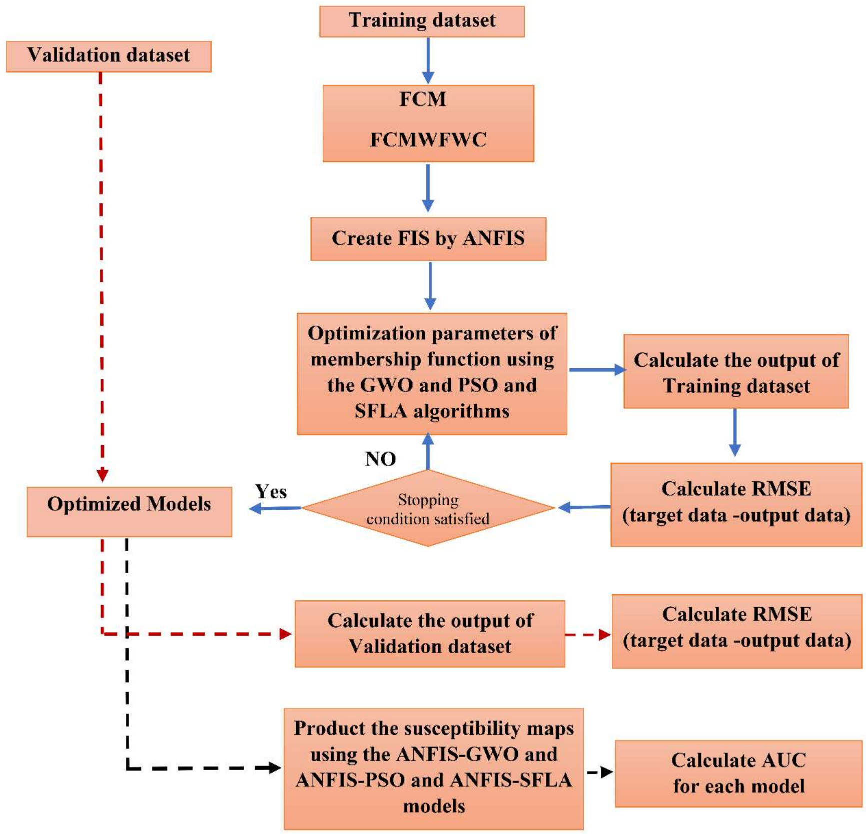

Fig.8. Optimization steps of hybrid models

Table 2. The parameters of optimization algorithms

Models	Iteration	Parameters		
ANFIS- PSO	500	w=0.9 → 0	c1=1, c2=3	Population=1000
ANFIS-SFLA	500	Frogs in Mem=50	Memeplex=20	Frogs=1000
ANFIS-GWO	500	0< (r1, r2) <1	a=2→0	Population=1000

4. Landslides susceptibility assessment

Landslide susceptibility expresses the likelihood of a landslide event occurring in a given area based on local terrain conditions or climate conditions. It usually partitions the geographical surface into zones of varying grades of stability based on the landslide inventory. The resulting output is a solely spatial distribution of the predicted categorized hazard probabilities across grid cells.

Machine learning methods applied for landslide susceptibility assessment represent a structured gathering of the available information extracted from landslide inventories, process/model with that information, and form a judgment about it in a transient workflow. This workflow unfolds through stages of preprocessing, implementation or modeling, and post processing, wherein modeling plays an essential role [37].

4.1. Workflow of a machine learning in landslides susceptibility assessment

Landslide assessments can reduce the hazard to a certain extent. However, the task of cataloging landslides to evaluate historical trends can be achieved and hazard zonation maps must be produced in this process. Such maps would make it easy for decision makers to take advantage of it to identify sensitive areas and manage regional land use. The main task of landslide assessment is preparation of landslide susceptibility maps incorporating spatial and temporal predictions of landslides on a regional scale. It is a great challenge to the global change research community. Results depend on the data used and modeling approaches employed, and the latter is the focus of this research [46].

So far, landslide susceptibility mapping methods and techniques have used simple expert knowledge first and then gradually evolved into sophisticated mathematical procedures, and using statistics with the aid of GIS. Physical-based methods are suitable for small-scale areas where detailed geomorphological and geological information are needed. These methods assess slope failure accurately because of its site-specific locations at a localized scale based on the safety factor index of slopes [46].

Supervised learning is by far the most widespread form of machine learning applied in landslide susceptibility assessment. The following are details about the workflow of supervised learning applied in landslide susceptibility assessment.

Initially, high-quality spatial data are collected from remotely sensed images or real-time monitoring for a landslide to produce landslide inventories. A landslide inventory includes historical landslide data and other related information, such as geological data, meteorological conditions, and topographical data, which can roughly clarify the relationships between predisposing factors and landslide occurrences. Based on these data, the predictive models for landslide susceptibility zonation can construct the relationships between the input and output variables. Prior to any prediction modeling, these two types of variables should be identified. Commonly, the output consists of landslides and non-landslides. The input relates to conditioning factors of landslides.

Redundant or irrelevant factors may create noise, decreasing the overall predictive capability of the models. It is essential to choose suitable factors in landslide susceptibility assessment.

To date, no universal guidelines have been agreed upon for the determination of case-specific conditioning factors [47]. Landslide conditioning factors show variation with respect to the study area and its geographical locations.

Every study area has its own particular set of factors that cause landslides. According to numerous studies, common landslide causal factors can be divided into two categories: (i) internal factors, which are related to geology and topography, such as the elevation, profile curvature, slope, plan curvature, distance to faults, aspect, distance from rivers, landform and lithology; and (ii) external factors, which usually cause landslides, such as rainfall, distance from roads and the seismic intensity.

To further select the appropriate input factors, one effective method involves ranking the importance of the input variables. Popular algorithms include ReliefF [48], Genetic Algorithms (GA), Information Gain Ratio (IGR), and symmetrical uncertainty analysis. Through calculating a score for each factor, these algorithms can evaluate and rank the contributions of landslide causal factors, and the factors with lower contributions are sequentially removed. Furthermore, machine learning methods can rank these factors by their weights. Unsupervised learning methods such as cluster sampling can evaluate factors by weighting the relative importance of each conditioning factor.

The predictive model is trained. The performance of the models is usually measured through some kind of cost function. It is also important to optimize model performance. This entails the adjustment of hyper parameters that control the training process, structure, and properties of the model [49]. For example, a validation dataset is separated from the test and training sets using sampling strategies.

The generic approach that was selecting the training sets is usually made by sampling 70% of all instances randomly throughout the available data. The remaining part is reserved for testing the model.

4.2. Conventional machine learning methods for landslides susceptibility assessment

Conventional machine learning algorithms have been applied to landslide susceptibility assessment and achieve outstanding performance and are mainly classified into single base learning algorithms and ensemble learning algorithms.

4.2.1. Single base learning algorithms

The most frequently traditional single algorithms applied for landslide susceptibility assessment include (1) LR, (2) SVM, (3) DT, and (4) ANN [50]. LR has a long tradition of application in landslide susceptibility assessment [51]. A study proved that the predictive model complexity and the size of the training dataset influence the accuracy and predictive power of LR models concerning landslide susceptibility. SVM can identify the optimal boundary between the training data from two classes. Compared with other algorithms, the SVM algorithm achieves slightly better accuracies in shallower landslide assessment applications.

As an original tree-like structure, DT The standard ANN model comprises three layers, namely an input layer (i.e., landslide conditioning factors), hidden layers, and an output layer (i.e., landslide susceptibility). A case proved that ANN applied in landslide susceptibility assessment achieved fairly precise models. In summary, several drawbacks are usually identified when utilizing the aforementioned single base learning algorithms, such as overfitting and unstable performance [37].

4.2.2. Ensemble learning algorithms

Generally, ensemble learning algorithms can enhance the performance of the single base learning algorithms and improve the robustness and generalizability. A commonly used ensemble algorithm in landslide susceptibility assessment is RF. Usually, an RF model has a more predictive capability to identify landslide susceptibility zones than other models [52].

For example, Hong et al. indicated that three ensemble models (i.e., AdaBoost, bagging, and rotation forest) could significantly improve the performance of J48 DT as the base learner, and rotation forest can be considered a promising method for landslide susceptibility mapping in similar cases with better accuracy than other methods.

Other ensemble methods have been developed for landslide susceptibility assessment, including GBDT [53], Random Subspace [48], Multiboot, and Regularized Greedy Forests. These ensemble methods can reduce both the bias and variance and avoid overfitting problems compared to the base classifiers to improve predictive capability.

4.3. Landslides susceptibility assessment

Recently, with the rapid development of deep learning, state-of-the-art learning approaches have been successfully applied in landslide susceptibility assessment in the field. Indeed, deep learning has also been commonly applied to feature extraction. Deep learning can find optimal features and handle indirect relationships between features and goals and can thus simplify the feature engineering and data preprocessing steps [37].

4.4. Prediction of landslide displacement

Landslide prediction is important for mitigating geo hazards but is very challenging. In landslide evolution, displacement depends on the local geological conditions and variations in the controlling factors [54].

4.4.1. Conventional machine learning methods for predicting landslide displacement

Recently, conventional machine learning methods, including the ANN, SVM [55], Gaussian process, and ELM, have been applied to produce models for landslide displacement prediction. Here, the input is the landslide displacement and the triggering factor. The output is the predicted landslide displacement.

To optimize time-series data used as input, Li et al. [56] introduced a chaos theory-based Wavelet Analysis-Volterra filter model (chaotic WA-Volterra model) into SVM for cumulative landslide displacement prediction. The WAVolterra model aims to decompose the cumulative displacement data into different low- and high-frequency components. Chaos theory was used to reconstruct the phase space of each frequency component. Reconstructed phase spaces were selected as the input–output data to train the SVM models. The predictive results (i.e., the predictive cumulative displacements) were obtained by summing the predictive displacements of each frequency component. This study indicates the potential for chaos characteristic identification of landslide displacements to be applied in machine learning. A certain optimization has been achieved in feature processing.

Most studies verified the superiority of their proposed methods by comparing unoptimized algorithms or a small number of state-of-the-art algorithms and using the prediction results of one operation of the models for comparison; they did not, however, repeatedly test their proposed method. Inadequate method comparison reduces conclusion credibility. Since machine learning methods have a certain degree of uncertainty, model training and predictions may differ each time. Using the prediction of one operation of the model for comparison, the excellent prediction performance claimed in the studies may be accidental and not repeatable. In addition, most studies on landslide displacement prediction based on machine learning only used one landslide case to verify the applicability and superiority of their proposed algorithm. This strategy may lead to unreliable conclusions because a prediction model that performs well on a site-specific landslide may not perform well on other landslides. Therefore, it is necessary to investigate the versatility, mean prediction accuracy, and prediction stability of machine learning methods through multiple landslide cases and repeated calculations. More importantly, there is still a lack of a comprehensive comparison of available machine learning methods in reservoir landslide displacement prediction [57].

4.4.2. Deep learning methods for predicting landslide Displacement

The aforementioned approaches regard landslide displacement prediction as a static regression problem. On the other hand, landslides are considered a dynamic system in which the displacement continues to change. The influencing factors and displacement conditions in one moment affect the displacement and stability conditions in the next moment. To investigate the dynamic process, LSTM is an appropriate method since it is suitable for learning the temporal dynamics of sequential data.

The general workflow for the application of LSTM in landslide displacement prediction is as follows. The measured accumulated displacement of the landslide is first divided into a trend term (i.e., a static component) and a periodic term (i.e., a dynamic component). Selected controlling factors and periodic terms will be considered input. Generally, LSTM adds loops to the architecture, receives these inputs, and outputs a predicted result. Finally, the LSTM model was validated and estimated by comparing the predicted total displacement with the monitoring results of the total displacements The LSTM model can establish connections between landslide conditions at different times and learn rules from previous deformation time steps. The results indicated that the LSTM model achieved a more satisfactory performance than static SVM methods [37].

5. Results

The aim of this study was to evaluate the spatial prediction of landslides using ensemble ML classifiers. The current analysis exploited a hybrid ANFIS model that was optimized using PSO, GWO, and SFLA, three evolutionary algorithms. Three common models, MLP, RF, and SVM, were used to test and evaluate their performance on the same training and validation datasets, to build a LSM in

EAP, Iran. There was a total of ten conditioning factors which were investigated. The landslide inventory database had 766 locations which were split into two groups: one for training, with 536 landslides, and another for evaluation of the model with 230 landslides. For analyzing the associations between landslides and landslide conditioning variables, the PCF model was exploited as a bivariate statistical test. Furthermore, in the present analysis, the Pearson correlation test was used to measure the predictive strength of ten landslide condition variables.

Table 3. The performance of ML Models

Item	RF	MLP	SVM
True positive	418	512	426
True negative	441	410	457
False positive	118	24	110
False negative	95	126	79
PPV (%)	78.49	94.95	80.14
NPV (%)	82.35	75.74	85.29
Sensitivity (%)	81.48	81.05	84.36
Specificity (%)	78.90	94.47	80.60
Accuracy (%)	80.13	86.01	82.37

Table 4. Evaluation of three ML models on training data

Models	AUC	95 % CI	Kappa index	SE
MLP	0.934	0.878-0.969	0.721	0.021
SVM	0.907	0.845-0.950	0.647	0.025
RF	0.868	0.800-0.920	0.603	0.030

Table 5. Evaluation of three ML models on the validation data

models	AUC	95 % CI	Kappa index	SE
MLP	0. 911	0.798-0.964	0.633	0.038
SVM	0.887	0.778-0.954	0.633	0.043
RF	0.871	0.759-0.944	0.630	0.046

Table 6. The validation of ML Models

Item	RF	MLP	SVM
True positive	176	199	184
True negative	199	181	192
False positive	54	31	46
False negative	31	49	38
PPV(%)	76.67	87.06	80.00
NPV (%)	86.67	78.69	83.33
Sensitivity (%)	85.02	80.24	82.88
Specificity (%)	78.66	85.38	80.67
Accuracy (%)	81.52	82.61	81.74

Table 7. Performance result of the all model

	MLP	SVM	RF	ANFIS-PSO	ANFIS-SFLA	ANFIS-GWO
AUC	0.911	0.887	0.871	0.89	0.88	0.88
Sen.	%80.24	%82.88	%85.02	%82.34	%81.22	%81.24
Spec.	%85.36	%80.67	%78.66	%81.62	%82.45	%82.64
Acc.	%82.61	%81.74	%81.52	%83.28	%80.95	%81.63

The fuzzy c-means clustering approach was then used to construct an initial fuzzy inference system for LSM. In addition, three wise algorithms, namely GWO, SFLA, and PSO, were used to train the ANFIS in the current analysis. One of the most significant benefits of these approaches is that they improve precision by optimizing and calculating ANFIS parameters. Indeed, it has the potential to reduce dimension dangers and the problems of local minimum, thus improving the ANFIS model accuracy. Lastly, ROC curves were used to test the LSMs generated by ANFIS-GWO, ANFIS-SFLA, and ANFIS-PSO. According to the results, the AUC values for the ANFIS-PSO, ANGIS-SFLA, and ANFIS-GWO models were 0.89, 0.88, and 0.88, respectively.

In solving high-dimensional and non-linear problems including landslide prediction, the fostered novel method effectively merged evolutionary algorithms, neuro-fuzzy inference mechanisms, and expert knowledge. The GWO, SFLA, and PSO usage in optimizing the structural parameters of ANFIS ensured that there would be lower rate of issues throughout the modeling processes owing to the local minimum and dimension dangers.

Comparing the performance of the ensemble method reveals that, the stacking method had a more robust classification efficiency. The ensemble ML, despite its excellent efficiency, is primarily affected by the structural parameters tuning processes.

Based on the optioned results, our future work will take place to apply the integrated approach of deep learning methods and object-based image analysis methods for semi/automate landsli'de detecting and delineation from earth observation satellite image. Application of remote sensing and ensemble ML techniques would help prudent planning for land development and assist engineers and authorities to adapt their decisions to take landslide hazard into account and to limit its consequences. The greater the ability of engineers and land-use managers to understand which landscapes have a high susceptibility to landslide occurrences, the greater is their ability to improve mitigation and risk assessment strategies [58].

Predictive models developed by machine learning for landslides prevention can be under constrained. For instance, models that perform well in datasets and are consequently viewed as high quality probably deviate strongly for situations and data outside their valid local areas because of the complex physical earth system. The challenges and opportunities in the applications of machine learning for landslides prevention will be discussed below [37].

6. Discussion

Meaningless or corrupted data in datasets is known as noise. This noise can result in errors in the predictions by the machine learning algorithms and can impact their performance in terms of accuracy, size of the model and the time taken to build the model.

Zhu and Wu have conducted a quantitative study of impact of noise on machine learning classification algorithms. The authors have shown that with increase in feature noise the accuracy of the classification algorithms decreases linearly. In terms of attribute noise, the study shows that the lowest level of classification accuracy is given by classifiers trained using noisy dataset compared to clean dataset on both clean and noisy target and the performance deteriorates linearly with addition in noise.

In another study, the effect of noise on 4 classification algorithms with various degrees and types of noise is studied. Of the 4 algorithms used, the authors conclude that Naïve Bayes and C4.5 are resistant towards noise with the former being the strongest, and IBk and SMO the least resistant to noise with the latter the worst of all [59].

6.1. Dataset heterogeneity

Climate and ecosystem processes reveal a high level of heterogeneity due to differences in geography, topography, and climatic conditions in diverse areas of the earth. For example, some regions are mountainous, some regions are dry and experience severe, long-term droughts, and some regions are quite wet and covered with dense forests. The patterns, mechanisms, and driving forces of landslides vary among these regions.

Currently, this heterogeneity in the data emphasizes the idea that landslides prevention models primarily apply to local or regional zones. Various factors correspond to a homogeneous group of locations. Developing a universal model that can be applied to global regions remains a challenge. Moreover, another source of heterogeneity is presented in different multi-sensors, which exhibit different imaging geometries, spatial and temporal resolutions, physical meanings, contents, and statistics [37].

6.2. Challenges from class imbalance

Machine learning methods require a vast amount of data to train a model. The data necessary for landslide susceptibility mapping is a collection of landslide causative factors as predictors and landslide inventory as a response variable; however, landslides do not occur everywhere, and the occurrence of landslides is limited in an area. This geophysical phenomenon leads to severely skewed class distribution, wherein the number of landslide samples (minority class) is significantly less than non-landslide locations (majority class). The imbalance in landslide data hampers the predictive ability of learning algorithms, and hence, the final models show poor performance in the class with fewer samples [60].

The training areas into two classes (i.e., landslides and non-landslides).There are fewer areas in the training regions in which landslides appear than non-landslides. Such imbalances can cause a model to be biased towards classifying the susceptible areas as safe since there is a larger number of nonland slide samples. After investigating the mapping of landslides with an imbalanced training sample (i.e., the sample contained more examples of non-landslide areas than landslide areas), a study indicated that the RF method underestimated landslide occurrences. For overcoming this problem, some typical solutions can be divided into three categories: data-level techniques, algorithm-level methods, and hybrid approaches [61].

7. Conclusion

The aim of this study was to evaluate the spatial prediction of landslides using ensemble ML classifiers. The current analysis exploited a hybrid ANFIS model that was optimized using PSO, GWO, and SFLA, three evolutionary algorithms. Three common models, MLP, RF, and SVM, were used to test and evaluate their performance on the same training and validation datasets, to build a LSM in EAP, Iran. There was a total of ten conditioning factors which were investigated. The landslide inventory database had 766 locations which were split into two groups: one for training, with 536 landslides, and another for evaluation of the model with 230 landslides. For analyzing the associations between landslides and landslide conditioning variables, the PCF model was exploited as a bivariate statistical test. Furthermore, in the present analysis, the Pearson correlation test was used to measure the predictive strength of ten landslide condition variables.

The fuzzy c-means clustering approach was then used to construct an initial fuzzy inference system for LSM. In addition, three wise algorithms, namely GWO, SFLA, and PSO, were used to train the ANFIS in the current analysis. One of the most significant benefits of these approaches is that they improve precision by optimizing and calculating ANFIS parameters. Indeed, it has the potential to reduce dimension dangers and the problems of local minimum, thus improving the ANFIS model accuracy. Lastly, ROC curves were used to test the LSMs generated by ANFIS-GWO, ANFIS-SFLA, and ANFIS-PSO. According to the results, the AUC values for the ANFIS-PSO, ANGIS-SFLA, and ANFIS-GWO models were 0.89, 0.88, and 0.88, respectively.

In solving high-dimensional and non-linear problems including landslide prediction, the fostered novel method effectively merged evolutionary algorithms, neuro-fuzzy inference mechanisms, and expert knowledge. The GWO, SFLA, and PSO usage in optimizing the structural parameters of ANFIS ensured that there would be lower rate of issues throughout the modeling processes owing to the local minimum and dimension dangers.

Comparing the performance of the ensemble method reveals that, the stacking method had a more robust classification efficiency. The ensemble ML, despite its excellent efficiency, is primarily affected by the structural parameters tuning processes.

Based on the optioned results, our future work will take place to apply the integrated approach of deep learning methods and object-based image analysis methods for semi/automate landslide detecting and delineation from earth observation satellite image. Application of remote sensing and ensemble ML techniques would help prudent planning for land development and assist engineers and authorities to adapt their decisions to take landslide hazard into account and to limit its consequences. The greater the ability of engineers and land-use managers to understand which landscapes have a high susceptibility to landslide occurrences, the greater is their ability to improve mitigation and risk assessment strategies.

References

[1] Chen, W., H.R. Pourghasemi, and Z. Zhao, A GIS-based comparative study of Dempster-Shafer, logistic regression and artificial neural

network models for landslide susceptibility mapping. Geocarto international, 2017. 32(4): p. 367-385.
[2] Feizizadeh, B. and T. Blaschke, An uncertainty and sensitivity analysis approach for GIS-based multicriteria landslide susceptibility mapping. International Journal of Geographical Information Science, 2014. 28(3): p. 610-638.
[3] Chen, W., Z. Sun, and J. Han, Landslide susceptibility modeling using integrated ensemble weights of evidence with logistic regression and random forest models. Applied sciences, 2019. 9(1): p. 171.
[4] Kose, D.D. and T. Turk, GIS-based fully automatic landslide susceptibility analysis by weight-of-evidence and frequency ratio methods. Physical Geography, 2019. 40(5): p. 481-501.
[5] Li, Y. and W. Chen, Landslide susceptibility evaluation using hybrid integration of evidential belief function and machine learning techniques. Water, 2020. 12(1): p. 113.
[6] Yan, F., et al., A novel hybrid approach for landslide susceptibility mapping integrating analytical hierarchy process and normalized frequency ratio methods with the cloud model. Geomorphology, 2019. 327: p. 170-187.
[7] Wang, Y., et al., Comparative study of landslide susceptibility mapping with different recurrent neural networks. Computers & Geosciences, 2020. 138: p. 104445.
[8] Pham, B.T., et al., A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. Geocarto International, 2020. 35(12): p. 1267-1292.
[9] Pourghasemi, H.R. and O. Rahmati, Prediction of the landslide susceptibility: Which algorithm, which precision? Catena, 2018. 162: p. 177-192.
[10] Lee, S., et al., Landslide susceptibility mapping using Naïve Bayes and Bayesian network models in Umyeonsan, Korea. Geocarto international, 2020. 35(15): p. 1665-1679.
[11] Ebrahimy, H., et al., A comparative study of land subsidence susceptibility mapping of Tasuj plane, Iran, using boosted regression tree, random forest and classification and regression tree methods. Environmental Earth Sciences, 2020. 79: p. 1-12.
[12] Park, S.-J., et al., Landslide susceptibility mapping and comparison using decision tree models: A Case Study of Jumunjin Area, Korea. Remote Sensing, 2018. 10(10): p. 1545.
[13] Gheshlaghi, H.A. and B. Feizizadeh, An integrated approach of analytical network process and fuzzy based spatial decision making systems applied to landslide risk mapping. Journal of African Earth Sciences, 2017. 133: p. 15-24.
[14] Zhu, A.-X., et al., A similarity-based approach to sampling absence data for landslide susceptibility mapping using data-driven methods. Catena, 2019. 183: p. 104188.
[15] Fang, Z., et al., Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. Computers & Geosciences, 2020. 139: p. 104470.
[16] Phong, T.V., et al., Landslide susceptibility modeling using different artificial intelligence methods: A case study at Muong Lay district, Vietnam. Geocarto International, 2021. 36(15): p. 1685-1708.
[17] Bui, D.T., et al., Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. Catena, 2020. 188: p. 104426.
[18] Ghorbanzadeh, O., et al., Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. Remote Sensing, 2019. 11(2): p. 196.
[19] Hong, H., et al., Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). Catena, 2018. 163: p. 399-413.
[20] Ahmadlou, M., et al., Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). Geocarto International, 2019. 34(11): p. 1252-1272.
[21] Pourghasemi, H.R., et al., Assessment of landslide-prone areas and their zonation using logistic regression, logitboost, and naïvebayes machine-learning algorithms. Sustainability, 2018. 10(10): p. 3697.

[22] Kan, G., et al., A new hybrid data-driven model for event-based rainfall–runoff simulation. Neural Computing and Applications, 2017. 28: p. 2519-2534.
[23] Nguyen, A.T. and L. Hens, Human ecology of climate change hazards in Vietnam. Cham, Switzerland: Springer International Publishing, 2019.
[24] Aditian, A., T. Kubota, and Y. Shinohara, Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. Geomorphology, 2018. 318: p. 101-111.
[25] Tian, Y., et al., Geomorphometry and statistical analyses of landslides triggered by the 2015 Mw 7.8 Gorkha earthquake and the Mw 7.3 aftershock, Nepal. Frontiers in Earth Science, 2020. 8: p. 407.
[26] Sarker, I.H., Machine learning: Algorithms, real-world applications and research directions. SN computer science, 2021. 2(3): p. 160.
[27] Nagar, A. and J. Sharma, A Comparative Evaluation of Supervised Learning and Unsupervised Learning. resmilitaris, 2022. 12(5): p. 1498-1507.
[28] Chou, S.-Y., J.-S.R. Jang, and Y.-H. Yang, FrameCNN: A weakly-supervised learning framework for frame-wise acoustic event detection and classification. Recall, 2017. 14: p. 55-64.
[29] Du, Y., et al. Semi-Supervised Learning for Surface EMG-based Gesture Recognition. in IJCAI. 2017.
[30] Farag, M.A., et al., Anti-acetylcholinesterase activity of essential oils and their major constituents from four Ocimum species. Zeitschrift für Naturforschung C, 2016. 71(11-12): p. 393-402.
[31] Srinivas, M., G. Sucharitha, and A. Matta, Machine learning algorithms and applications. 2021: John Wiley & Sons.
[32] Fernández-Delgado, M., et al., Do we need hundreds of classifiers to solve real world classification problems? The journal of machine learning research, 2014. 15(1): p. 3133-3181.
[33] Teh, Y.W. and G.E. Hinton, Rate-coded restricted Boltzmann machines for face recognition. Advances in neural information processing systems, 2000. 13.
[34] Graves, A. and A. Graves, Long short-term memory. Supervised sequence labelling with recurrent neural networks, 2012: p. 37-45.
[35] Stoffel, M., D. Tiranti, and C. Huggel, Climate change impacts on mass movements—case studies from the European Alps. Science of the Total Environment, 2014. 493: p. 1255-1266.
[36] Froude, M.J. and D.N. Petley, Global fatal landslide occurrence from 2004 to 2016. Natural Hazards and Earth System Sciences, 2018. 18(8): p. 2161-2181.
[37] Ma, Z., G. Mei, and F. Piccialli, Machine learning for landslides prevention: a survey. Neural Computing and Applications, 2021. 33(17): p. 10881-10907.
[38] Chen, W., et al., A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena, 2017. 151: p. 147-160.
[39] Tien Bui, D., et al., 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. International Journal of Digital Earth, 2016. 9(11): p. 1077-1097.
[40] Abdollahizad, S., et al., Using hybrid artificial intelligence approach based on a neuro-fuzzy system and evolutionary algorithms for modeling landslide susceptibility in East Azerbaijan Province, Iran. Earth Science Informatics, 2021. 14: p. 1861-1882.
[41] Elmoulat, M., et al., Edge computing and artificial intelligence for landslides monitoring. Procedia Computer Science, 2020. 177: p. 480-487.
[42] Tien Bui, D., et al., A novel integrated approach of relevance vector machine optimized by imperialist competitive algorithm for spatial modeling of shallow landslides. Remote Sensing, 2018. 10(10): p. 1538.
[43] Alizadeh, M., et al., Social vulnerability assessment using artificial neural network (ANN) model for earthquake hazard in Tabriz city, Iran. Sustainability, 2018. 10(10): p. 3376.
[44] Van Dao, D., et al., A spatially explicit deep learning neural network model for the prediction of landslide susceptibility. Catena, 2020. 188: p. 104451.

[45] Li, W., Z. Fang, and Y. Wang, Stacking ensemble of deep learning methods for landslide susceptibility mapping in the Three Gorges Reservoir area, China. Stochastic Environmental Research and Risk Assessment, 2021: p. 1-22.

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

[47] Kavzoglu, T., E. Kutlug Sahin, and I. Colkesen, An assessment of multivariate and bivariate approaches in landslide susceptibility mapping: a case study of Duzkoy district. Natural Hazards, 2015. 76: p. 471-496.

[48] Luo, X., et al., Coupling logistic model tree and random subspace to predict the landslide susceptibility areas with considering the uncertainty of environmental features. Scientific reports, 2019. 9(1): p. 15369.

[49] Schmidt, J., et al., Recent advances and applications of machine learning in solid-state materials science. npj computational materials, 2019. 5(1): p. 83.

[50] Kalantar, B., et al., Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). Geomatics, Natural Hazards and Risk, 2018. 9(1): p. 49-69.

[51] Kuhn, S., M.J. Cracknell, and A.M. Reading, Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia. Geophysics, 2018. 83(4): p. B183-B193.

[52] Goetz, J., et al., Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers & geosciences, 2015. 81: p. 1-11.

[53] Song, Y., et al., Landslide susceptibility mapping based on weighted gradient boosting decision tree in Wanzhou section of the Three Gorges Reservoir Area (China). ISPRS International Journal of Geo-Information, 2018. 8(1): p. 4.

[54] Miao, F., et al., Prediction of landslide displacement with step-like behavior based on multialgorithm optimization and a support vector regression model. Landslides, 2018. 15: p. 475-488.

[55] Zhu, C.H. and G.D. Hu, Time series prediction of landslide displacement using SVM model: application to Baishuihe landslide in Three Gorges reservoir area, China. Applied Mechanics and Materials, 2013. 239: p. 1413-1420.

[56] Li, Y., et al., Forecasting of landslide displacements using a chaos theory based wavelet analysis-Volterra filter model. Scientific reports, 2019. 9(1): p. 19853.

[57] Wang, Y., et al., A comparative study of different machine learning methods for reservoir landslide displacement prediction. Engineering Geology, 2022. 298: p. 106544.

[58] Abdollahizad, S., et al., Stacking ensemble approach in data mining methods for landslide prediction. The Journal of Supercomputing, 2023. 79(8): p. 8583-8610.

[59] Saseendran, A.T., et al., Impact of noise in dataset on machine learning algorithms. Mach. Learn. Res, 2019. 1: p. 1-8.

[60] Gupta, S.K. and D.P. Shukla, Handling data imbalance in machine learning based landslide susceptibility mapping: a case study of Mandakini River Basin, North-Western Himalayas. Landslides, 2023. 20(5): p. 933-949.

[61] Krawczyk, B., Learning from imbalanced data: open challenges and future directions. Progress in artificial intelligence, 2016. 5(4): p. 221-232.