Research Article

GIS-Based Soft Computing Models for Landslide Susceptibility Mapping: A Case Study of Pithoragarh District, Uttarakhand State, India

Trung-Hieu Tran,[1] Nguyen Duc Dam [iD],[1] Fazal E. Jalal,[2] Nadhir Al-Ansari [iD],[3] Lanh Si Ho,[1,4] Tran Van Phong [iD],[5] Mudassir Iqbal,[2,6] Hiep Van Le,[1] Hanh Bich Thi Nguyen,[1] Indra Prakash [iD],[7] and Binh Thai Pham [iD][1]

[1]*University of Transport Technology, Ha Noi 100000, Vietnam*
[2]*Department of Civil Engineering, State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*
[3]*Department of Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, Lulea 971 87, Sweden*
[4]*Civil and Environmental Engineering Program, Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8527, Japan*
[5]*Institute of Geological Sciences, Vietnam Academy of Sciences and Technology, 84 Chua Lang Street, Dong Da, Hanoi 100000, Vietnam*
[6]*Department of Civil Engineering, University of Engineering and Technology, Peshawar, Pakistan*
[7]*DDG (R) Geological Survey of India, Gandhinagar 382010, India*

Correspondence should be addressed to Nadhir Al-Ansari; nadhir.alansari@ltu.se and Binh Thai Pham; binhpt@utt.edu.vn

The main objective of the study was to investigate performance of three soft computing models: Naïve Bayes (NB), Multilayer Perceptron (MLP) neural network classifier, and Alternating Decision Tree (ADT) in landslide susceptibility mapping of Pithoragarh District of Uttarakhand State, India. For this purpose, data of 91 past landslide locations and ten landslide influencing factors, namely, slope degree, curvature, aspect, land cover, slope forming materials (SFM), elevation, distance to rivers, geomorphology, overburden depth, and distance to roads were considered in the models study. Thematic maps of the Geological Survey of India (GSI), Google Earth images, and Aster Digital Elevation Model (DEM) were used for the development of landslide susceptibility maps in the Geographic Information System (GIS) environment. Landslide locations data was divided into a 70 : 30 ratio for the training (70%) and testing/validation (30%) of the three models. Standard statistical measures, namely, Positive Predicted Values (PPV), Negative Predicted Values (NPV), Sensitivity, Specificity, Mean Absolute Error (MAE), Root Mean Squire Error (RMSE), and Area under the ROC Curve (AUC) were used for the evaluation of the models. All the three soft computing models used in this study have shown good performance in the accurate development of landslide susceptibility maps, but performance of the ADT and MLP is better than NB. Therefore, these models can be used for the construction of accurate landslide susceptibility maps in other landslide-prone areas also.

1. Introduction

Landslide occurrences in Himalayan region are 30% of the total landslides in the world. Landslides cause huge loss of economy, damage to property, and loss of life all over the world [1]. Every year casualties in the Himalayan region are about 200 people, and economic loss due to landslides is more than US\$ 1 billion [2]. Causes of landslides are natural and anthropogenic. Landslides are affected by geological, topographical, meteorological, and geo-environmental

factors. Land-use pattern changes due to anthropogenic activities such as road construction, power houses excavations, and open-pit/cast mining cause landslides [3]. Natural causes of landslides include earthquakes, glacier melting, and rainfall. Currently, impact of the climate change has drastically amplified the landslides phenomenon globally [4, 5].

Landslide susceptibility map is regarded as an integral part of landslide management of a particular region [6, 7]. Such maps find their applications in minimizing landslide risk to human life, man-made infrastructures, particularly network of roads in mountainous regions, and land-use planning [8]. Landslide susceptibility means that the area is susceptible to landslides and probability of occurrence of landslides is there on the basis of past experience under similar geo-environment conditions [9]. There are several statistical techniques aided by Geographic Information System (GIS) and Remote Sensing methods that have been widely employed by many researchers to formulate models in order to generate landslide susceptibility maps. In general, there are two main approaches used in modeling and mapping of the landslide susceptibility, namely, quantitative and qualitative methods [10]. In qualitative methods, expert's opinion-based models are often used and applied such as Analytical Hierarchy Process (AHP) and Weighted Linear Combination (WLC) [11]. Whereas, quantitative methods can be classified as deterministic and probabilistic methods which involve the mathematical relationship between landslides' occurrences and their associated affecting factors. Statistical and Machine Learning methods such as Binary Logistic Regression, Information Value, Likelihood Ratio, Logistic Regression (LR), and Multivariate Regression are some the quantitative methods applied for landslide susceptibility modeling and mapping [12–15]. Furthermore, few other algorithms with optimum performance and a higher degree of accuracy used in the landslide susceptibility mapping include Discriminant Analysis, Generalized Additive Models, Evidential Belief Functions, Weighted Linear Combinations, and Weights of Evidence [16–20]. Also, Certainty Factors, Probabilistic Models, Modified Bayesian Estimation, Information Values [21–24], Dempster–Shafer models, and Index of Entropy are also widely used as the statistical quantitative techniques to map the landslide susceptible areas [25–28].

Nowadays, researchers are extensively using Machine Learning (ML) algorithms in studies related to natural hazards including landslide studies, such as Multivariate Adaptive Regression Spline [29], Support Vector Machine (SVM) [30], Classification and Regression Trees [31], Boosted Regression [32], Quadratic Discriminant Analysis [33], Naïve Bayes Tree (NBT) [34], Maximum Entropy [35], Random Forest (RF) [36], Artificial Neural Networks (ANN) [37], and Generalized Linear Model [38]. In addition, Naïve Bayes (NB), Multilayer Perceptron (MLP) neural network classifier, and Alternating Decision Tree (ADT) are widely employed algorithms to investigate landslide susceptibility [39–41]. Recently, Senouci et al. [42] deployed a knowledge-driven approach alongside Analytical Hierarchy Process (AHP) in a GIS-based environment to evaluate the landslide

susceptibility map in Algeria. Hong et al. [41] evaluated two-class Kernel Logistic Regression (KLR), SVM, and ADT for landslide susceptibility mapping at the Yihuang area (China). It was found that ADT has better capabilities concerning to the prediction compared to SVM and KLR. Sahana et al. [43] employed the MLP neural network to estimate rainfall induced landslide susceptibility. Chen and Wang [44] compared KLR, NBT, and ADT models for landslide susceptibility assessment at Taibai County, China.

In the present research, the main objective was to use three soft computing ML algorithms, namely, NB, MLP, and ADT-based models, which are popular ML models, to predict the landslide susceptibility of Pithoragarh District of Uttarakhand State, India, which is one of the landslide-prone areas in the Himalayas. In literature, there are several studies have been carried out in landslide susceptibility mapping in Himalayas areas [45–47]. However, the application of ML models is still limited. Moreover, during literature survey, it is noticed that, for the first time, these models have been applied in this area for landslide susceptibility mapping. Thus, the main objective of this study was to apply best ML models for the accurate landslide susceptibility mapping of the study area.

In this study, standard statistical measures including Receiver Operating Characteristic (ROC) curve techniques were used to evaluate performance of the models for the selection of the best model for the accurate landslide susceptibility mapping. The maps produced would be useful for regional planning and disaster management. Weka software was used for the models study and GIS software for the development of thematic maps and visualization.

2. Material and Methods

2.1. Study Area. In the present study, the Pithoragarh district of Uttarakhand, India, has been selected as it is one of the prominent landslide-prone areas of the Himalayas [48]. This area is mountainous, characterized by high hills and deep valleys having elevation ranges between 1500 m and 2500 m gradually increasing from south to north (Figure 1). Two main rivers Sarju and Ramganga flow in this area forming deep valleys/gorges. The region is susceptible to occurrence of landslides on the excavated slopes of road sections particularly on Bansura-Rameshwar Ghat Road, in the northeast part, in the suburbs of Muwani, Dewal Thal, northern Chera, and Gangolihat places. Geologically, the aforementioned area is complex due to presence of folded and faulted rocks, belonging to Almora crystalline (granitoid) group and Garhwal group (shale, slate, phyllite, quartzite, dolomite, limestone, magnesite, occasional calc slate, and metavolcanics). The region witnesses tectonic activities that ultimately lead to folding and faulting of the rocks [49]. Quaternary sediments are located in river valleys, on the hill slopes and also as glacial deposits.

2.2. Methodology. The methodology followed in undertaking the formulation of models in this study is illustrated in the Figure 2, which is self-explanatory. Landslide
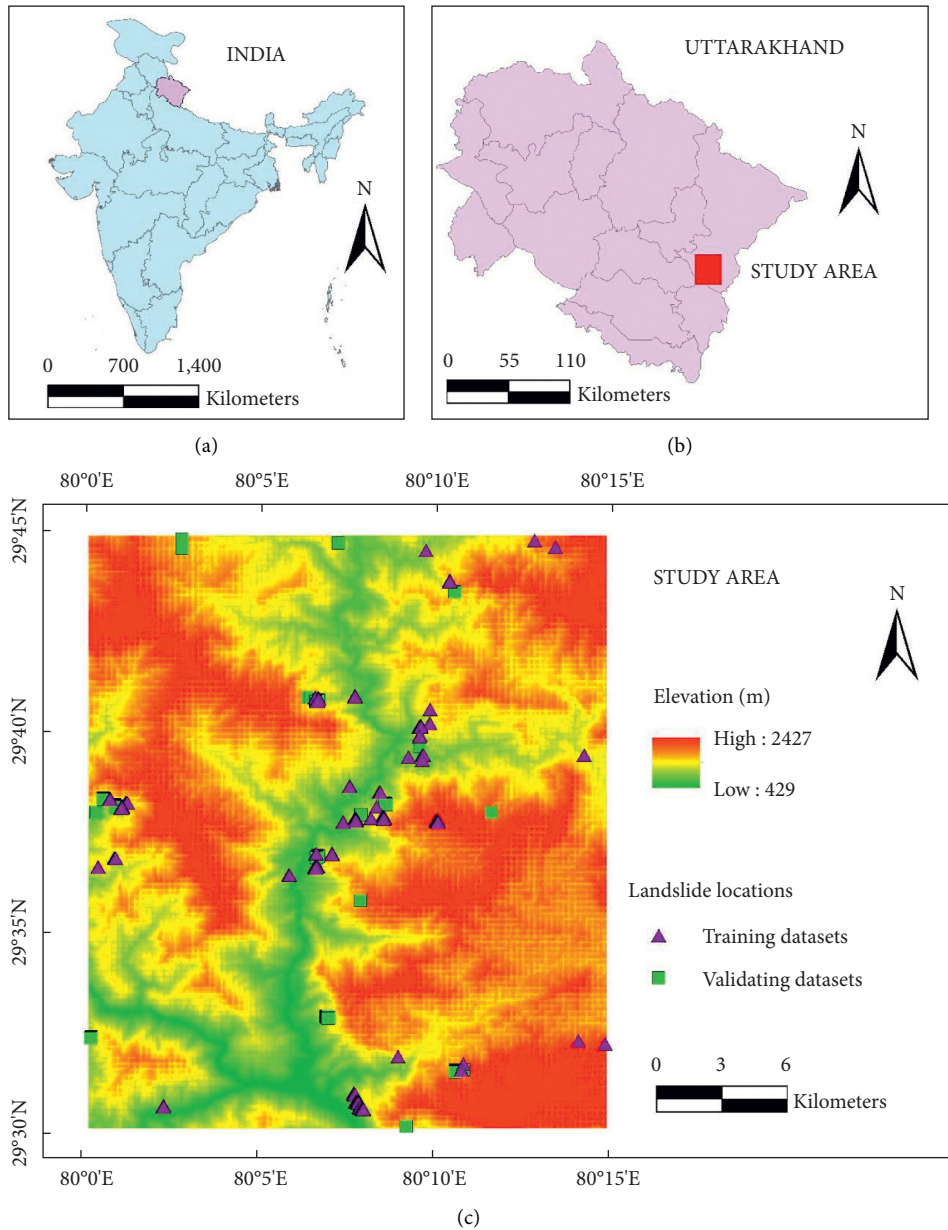
Figure 1: Location of the study area and historical landslides.

conditioning factors and landslide inventory were utilized in order to model the data using three ML techniques. Models were validated in terms of degree of accuracy and performance using various indices levels. In this study, data of GSI (http://www.portal.gsi.gov.in), Google Earth images, and Aster DEM were used for the development of thematic maps.

2.3. Data Used

2.3.1. Landslide Inventory. In order to undertake the landslide susceptibility mapping for the study area, it is necessary to address the landslide inventory which indicates the geospatial distribution of several factors including locations, types, and nature of landslides [50]. For the

model development, the landslide polygon data represented on map by points was split randomly into ratio of 70 : 30 as the training dataset (70%) and validation dataset (30%), respectively [51, 52]. The training data (70%) was used for landslide susceptibility zoning whereas the remaining 30% testing dataset for the validation was used to check accuracy and robustness of models. This split ratio was selected based on the experience of other similar research studies [51, 52]. In this research, the inventory of 91 landslides marked as points in the map was prepared based on the record of (GSI) (https://www.gsi.gov.in/webcenter/portal) as well as interpretation of the Google Earth images for the modeling (Figures 1 and 3). Most of the identified landslides in the area are translational rock-cum-debris slides.
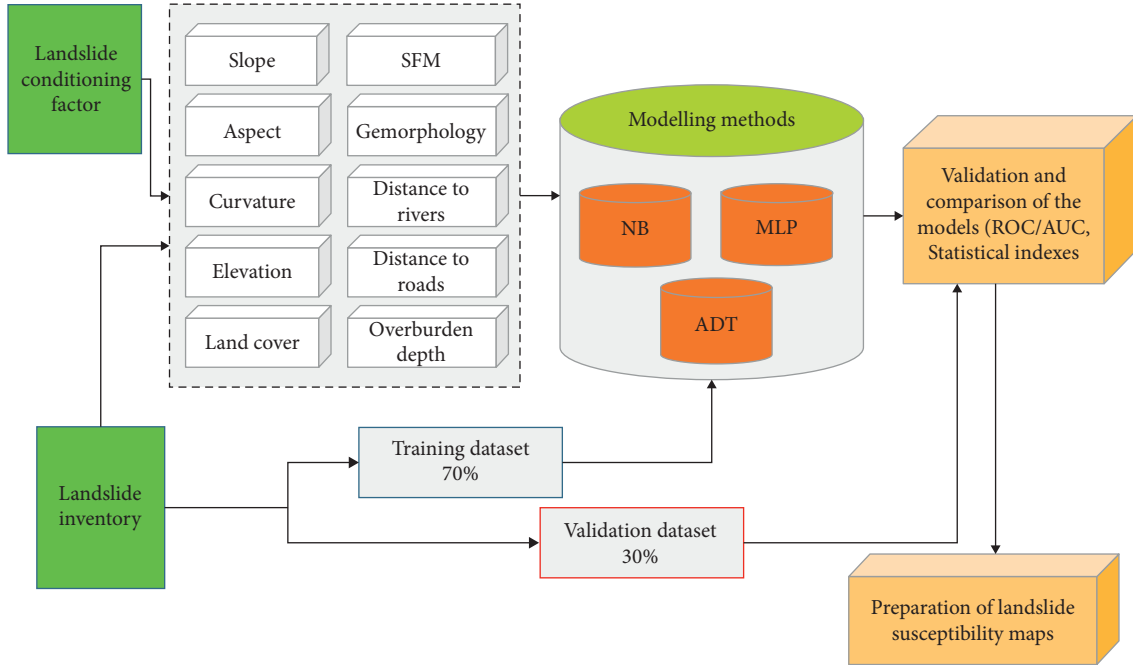
FIGURE 2: Methodological framework for landslide susceptibility mapping in this study.

FIGURE 3: Landslide photos of the study area (source: http://www.portal.gsi.gov.in).

2.3.2. Landslide Conditioning Factors. In this study, ten impacting parameters, namely, slope degree, slope aspect, slope curvature, elevation, geomorphology, land cover, distance to roads, distance to rivers, overburden depth, and Slope Forming Materials (SFM) were taken into consideration on the basis of the indigenous geo-environmental scenario. For the models' study, thematic layers of the aforementioned parameters were generated using 30 m cell size with the aid of ArcGIS software. Slope degree, aspect, curvature, drainage, and elevation maps were generated from the Aster DEM collected from the United States Geological Survey (USGS) (https://earthexplorer.usgs.gov). Geological, geomorphological, and SFM maps were extracted from the reports of the GSI (https://www.gsi.gov.in/webcenter/portal). Land cover map was generated using GSI data and Google Earth images. The binary classification method (0's and 1's) was employed for modeling landslides that is "1" for the presence of landslide instances and "0" for the absence [53, 54]. Table 1 shows the original sources of the

TABLE 1: Information of data used in this study.

Data used	Resolution	Source
Land cover, geomorphology, overburden depth, and SFM	1 : 50,000	GSI report (https://www.gsi.gov.in/webcenter/portal) and Google Earth Image (2020)
Elevation, slope angle, slope aspect, curvature, road, and rivers	30 m	DEM (https://earthexplorer.usgs.gov)

data used. Figure 4 shows the maps of the landslide conditioning factors.

(1) Slope Angle. It is an extremely vital parameter while evaluating the occurrence of landslides. Different types of landslides depend on the slope angle. Most of the landslides are reported between 25 and 45 degree slopes [55]. The slope angle also regulates the surface and subsurface water flow that has direct
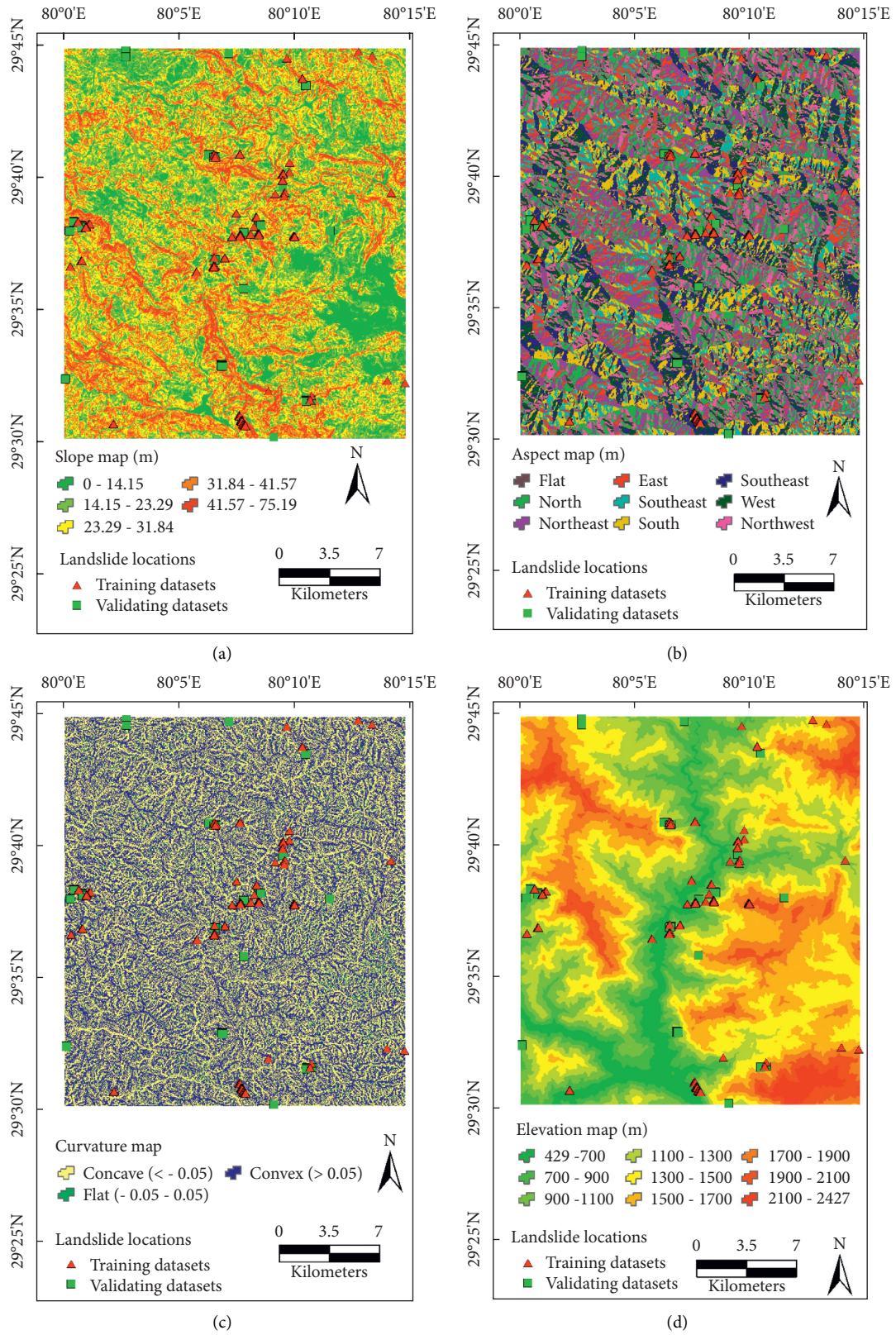
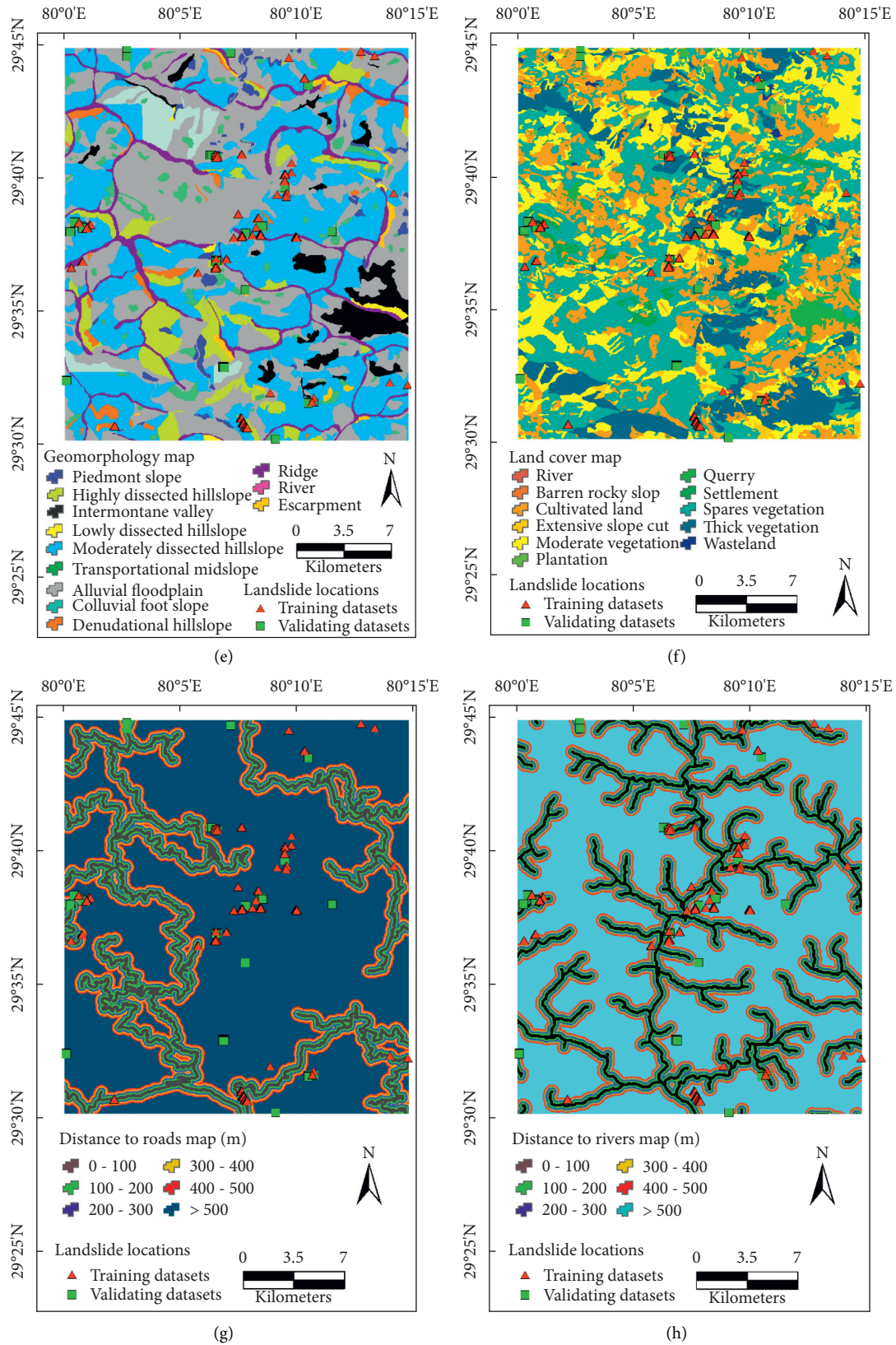(a)

(b)

(c)

(d)

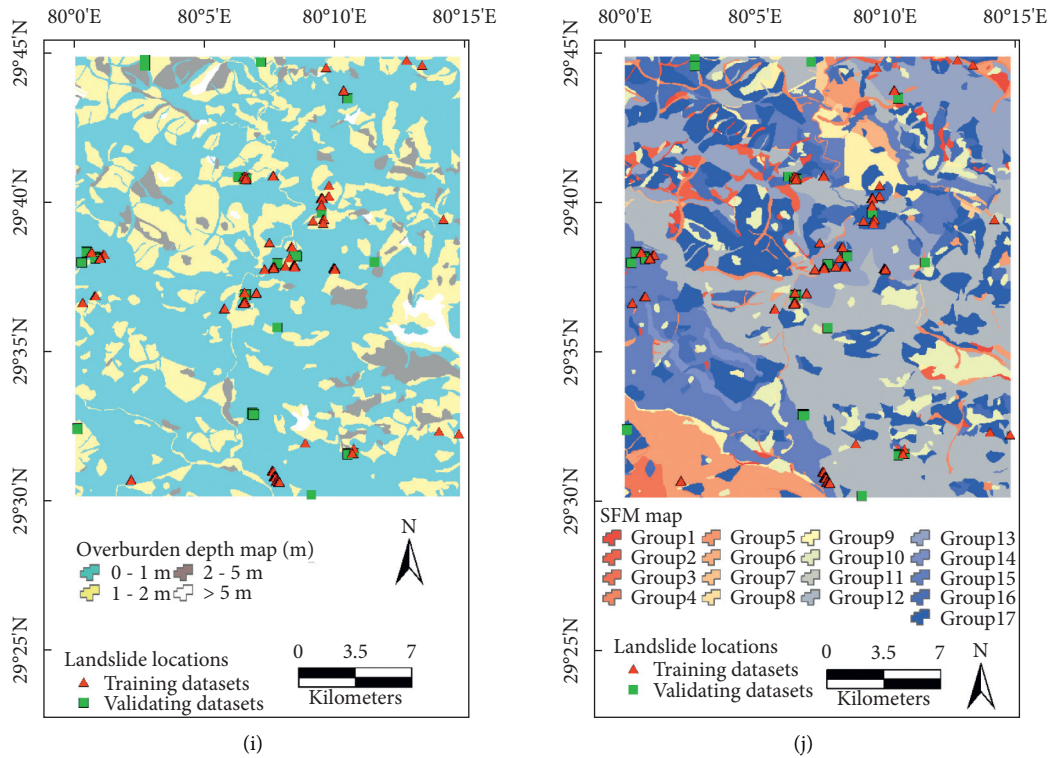Figure 4: Continued.

(e)

(f)

(g)

(h)

(i)

(j)

Figure 4: Thematic maps of the study area: (a) slope, (b) aspect, (c) curvature, (d) elevation, (e) geomorphology, (f) land cover, (g) distance to roads, (h) distance to rivers, (i) overburden depth, and (j) SFM.

impact on the runoff and infiltration, thus on landslide occurrences [56]. Here, the slope angle was derived from Aster DEM using the slope tool. Based on natural breaks classification, the slope was divided into five classes (Figure 4(a)).

(2) Slope Aspect. Slope aspect is also deemed as a significant parameter to effectively evaluate the landslide susceptibility that is associated with the direction of the sun, air/wind, and rain/precipitation in the specified area [57, 58]. This factor also indirectly impacts the vegetation and amount of water held in the soil. Here, the slope aspect was derived from Aster DEM. Based on natural breaks classification method, the slope was divided into nine different categories, flat, north, northeast, east, southeast, south, southwest, west, and northwest, using ArcGIS (Figure 4(b)).

(3) Curvature. It is another important factor in the surface runoff and ground infiltration which ultimately impacts the erosion, accumulation, and flow of surface and groundwater [57]. Therefore, the curvature tends to affect the occurrence of landslides to a larger extent. In the current study, the curvature map was generated from DEM and categorized as the concave, convex, and flat surfaces using ArcGIS software (Figure 4(c)).

(4) Elevation. Elevation is considered as an imperative parameter in the occurrence of landslides [56]. Generally, landslides occur in mountainous regions with deep valleys and rugged surfaces. In the hilly area, at higher elevations, usually the intensity of rainfall is less; however, glaciers' existence and prominence are relatively more [46]. At higher

elevations, landslides occur due to melting of ice causing erosion. Moraines also block the passage of streams and rivers at higher elevations forming lakes which sometimes burst and cause sudden floods and landslides in the area.

Majority of the rains and vegetation occur mostly in the range of lower and middle elevations in the Himalayas [45]. Furthermore, the landslide events are hugely associated with the elevation where slopes are moderate to high, moreover in the less vegetation area affected by heavy rainfall [59]. In this study, the elevation map was divided into different classes based on the natural break method for landslide susceptibility analysis (Figure 4(d)).

(5) Geomorphology. Geomorphology is one more vital factor for the evaluation of landslide susceptibility [60]. Geomorphological features encompass undulating grounds, mountains, river terraces, valleys, ridges, and escarpments, which affect the occurrence of landslides alongside other topographic and geo-environmental parameters. The geomorphology map (Figure 4(e)) was extracted from the GSI report (https://www.gsi.gov.in/webcenter/portal).

(7) Land Cover. Generally, bare lands and nonvegetated areas are more susceptible to landslides [61]. In a densely vegetated area, roots of plants reinforce the ground, thereby preventing erosion and thus landslides. Incidences of landslides are also in the cultivated region because of saturation of soil as a result of irrigation and also due to removal of surface soil strata. The land cover map was extracted from the GSI report (https://www.gsi.gov.in/webcenter/portal) and verified from the

Google Earth images. The land cover map of the study area comprises of eleven major groups of land cover (Figure 4(f)).

(8) Distance to Roads. Presence of the road network in a region highly impacts the landslide occurrences [62]. Roadside excavation and vegetation removal are actions that lead to landslides during pavement/road construction [63]. Excavation of roads create instability in ground mass resulting in slope failure and landslides adjacent to roads up to certain distances based on the nature of soil and geology. Thus, the distance from the roads is an immensely imperative parameter in the landslide study [64]. In this study, roads were identified and delineated using Google Earth images. The road buffer map at 100 m intervals was generated using ArcGIS tool (Figure 4(g)).

(9) Distance to Rivers. Distance to rivers is of prime importance for the development of landslide susceptibility maps. The hydrological network regime, soil saturation of water sources, and groundwater recharge, alongside the mounting water pressure on the groundmass and slopes, lead to landslides in areas adjacent to water sources, rivers, and streams. Distance to rivers maps (Figure 4(h)) was created from the rivers extracted from DEM using Flow Accumulation tool in GIS and buffered at an interval of 100 m for the landslide susceptibility mapping.

(10) Overburden Depth. Chances of slope failure are more likely in thick overburden areas depending on the characteristics of the overburden material. Major part of the study area is covered thin over burden (1–3 m) with occasional relatively thick pockets of greater than 5 m. Therefore, the possibility of landslides due to failure of the overburden material is very less. However, nature and thickness of the material affect the infiltration, thus ground water conditions in the area which may affect moisture conditions of the underlying rock mass creating instability and thus landslides. The overburden material map was extracted from the report of GSI (https://www.gsi.gov.in/webcenter/portal) (Figure 4(i)).

(11) Slope for Mining Material (SFM). SFM map was extracted from the GSI report (https://www.gsi.gov.in/webcenter/portal). The type of the slope forming material is very important in the landslide. Characteristics of the groundmass depend on the SFM, namely, its permeability, porosity, and geotechnical properties. Types of landslides depends on the above characteristics of the material and also on the binding/looseness of soil and rock mass, and the SFM map of the present study area was classified into seventeen main groups (Figure 4(j)): Group 1 (alluvium); Group 2 (chlorite, schist, and massive amphibolite); Group 3 (dolomite); Group 4 (garnet, mica, schist, and micaceous quartzite); Group 5 (granite, gneiss, garnetiferous schist, and amphibolite); Group 6 (in situ soil); Group 7 (limestone with intercalations of shale); Group 8 (limestone, dolomite, shale, and cherty quartzite); Group 9 (metabasite); Group 10 (older well-compacted debris); Group 11 (phyllite, stromatoli-phyllite, and stromatolitic); Group 12 (phyllite, stromatolitic dolomite, limestone, and Cu and Mg min); Group 13 (schist, augen gneiss, quartzite, and amphibolites); Group 14 (slate, lenses of quartzite and dolomite, and epidorite dyke); Group 15

(slate, quartzite, and dolomite with epidiorite dyke); Group 16 (transported soil); Group 17 (transported soil and colluvium).

2.4. Methods Used.
In this study, we have used three ML models, namely, NB, ADT, and MLP. Brief description of these models is mentioned in the following sections:

2.4.1. Naïve Bayes (NB).
This model (NB) is a classification model based on Bayes theorem [65]. A variable value on a given class is assumed as independent of another variable by the NB classifier, which is also called class conditional independence. The NB algorithm is simpler to use on various types of datasets [66]. Consider X to be a dataset with no class label. Now, assume H to be a hypothesis that X belongs to a specified class C. We aim to ascertain $P(H|X)$, the probability that the hypothesis H relates with given the observed X. $P(H|X)$ is the posterior probability indicating our confidence in the hypothesis after X is given. The Bayesian theorem offers a way of estimating the posterior probability $P(H|X)$ using probability $P(H)$, $P(X)$, and $P(X|H)$ [67]. The equation is presented as follows:

$$P(H|X) = \frac{p(H|X). P(H)}{P(X)}. \tag{1}$$

Assume that there is a set comprising m samples $S = \{S1, S2, \ldots, Ms.\}$ where sample Si is denoting as an n-dimensional feature vector $\{X1, X2, \ldots, XNA\}$ as the training dataset. Values X_i relate to attributes A1, A2, $\ldots$, An, respectively. Also, k classes $c1, c2, \ldots, ck$, and every sample belong to one of these classes. In the current research approach, k value is set to two since only slide-prone and nonslide-prone classes exist. Given an additional data sample X whose class is unknown, it is possible to predict X using the highest conditional probability $P(Ci|X)$, where $i = 1, 2, \ldots, k$. Then, the NB theorem is formulated as

$$P(C_i|X) = \frac{p(X|C_i). P(C_i)}{P(X)}. \tag{2}$$

2.4.2. Multilayer Perceptron (MLP) Neural Network Classifier.
This MLP neural network is a classification technique also considered as one type of Artificial Neural network that has been widely in use, for instance, in remote sensing applications for the sake of land cover classification [59, 68]. It encompasses several merits, for instance, partitioning of training data points is done straightforwardly without any prior assumptions, no decision is deemed as fixed concerning the comparative significance of several input measurements, and most of these measurements are selected on the basis of weights adjustment during the training process [69].

The MLP neural network comprises three main layers: input layers, hidden layers, and output layers [70]. The input layers consist of factors that affect landslides, while the output layers yield the classified results that help in classifying landslide or nonlandslide classes. The hidden layers act as classifiers to convert the input into the output form.

Two main steps are involved in the MLP neural nets' training: initially, the inputs are processed through the hidden layers for obtaining the desired outcomes. Then, the output results are associated to the target values to calculate the specified error, and ultimately, the biases/weights are adjusted to improve the results [71].

Suppose that $x = x_i$, where $i = 1, 2, \ldots, n$ is a vector comprising n factors affecting the landslides, and $y = 1$ (representing landslide class) or 0 (non-andslide class). Afterwards, the classification function for the MLP neural network function is illustrated in the form of equation (1):

$$Y = f(x), \tag{3}$$

where $f(x)$ refers to the hidden function which is yielded by the adjustment of biases during the training process of the model for a given architecture of the network.

2.4.3. Alternating Decision Tree (ADT).

ADT refers to an advance version of Decision Trees (DT) having boosting which is capable of generating classifiers with higher accuracy [34, 72]. In literature, it has been shown that ADT-based landslide susceptibility modeling shows superior modeling performance [41, 73]. ADT encompasses dual alternating layers; the layer containing decision nodes, also called splitter nodes, help to specify the condition of predicate, while the second layer contains optional nodes that are also known as the prediction nodes which contain only a single number. The resulting tree always exhibits the prediction nodes in the form of both roots and leaves assuming no interconnection with extradecision nodes. ADT has been employed to segregate the landslide or nonlandslide classes while modeling the landslide susceptibility, in this current study. Also, this tree is capable to generate accurate classifiers which can be interpreted simply in contrast to the variety of DT algorithms, for example, Classification and Regression Trees (CART), and Random Forest (RF) [41, 73].

The classification rules of ADT are based on defining a set of paths at a particular time. When a certain path approaches the decision node, it is accompanied by the off spring that represents the results of the decision linked to the particular node. It is necessary to mention that the path remains with the entire offspring of the node upon approaching the prediction node, thus leading the instance to track a "multipath." The categorization for the recently produced instance refers to the landslide susceptibility of the individual pixel that can be determined from the sign of summation of the entire values present over predictor nodes along the aforementioned multipath [74, 75]. The basic algorithm of the ADT algorithm is stated below [72].

First of all, it is the weight of an individual training instance that produces a group of weak suppositions. Therefore, the initial estimation is done using equation (2), which also refers to entire weights comprising the positive and negative instances:

$$a = \frac{1}{2} \ln \frac{W_+(T)}{W_-(T)}, \tag{4}$$

where T represents the precondition and condition of an only base rule.

After that, the precondition c_1 and condition c_2 are used to define $Z_t(c_1, c_2)$. In equation (3), $\hat{}$ and $\neg$ represent conjunction (AND gate) and negation (NOT gate), respectively:

$$Z_t(c_1, c_2) = 2\left(\sqrt{W_+(c_1^{c_2}) W_-(c_1^{c_2})} + \sqrt{W_+(c_1^{\neg c_2}) W_-(c_1^{\neg c_2})} \right) + W(\neg C_2), \tag{5}$$

where a and b are the estimated values which can be defined by equation (4), wherein the optimized c_1 and c_2 values are chosen by reducing $Z_t(c_1, c_2)$:

$$a = \frac{1}{2} \ln \frac{W_+(T) c_1^{c_2}}{W_-(T) c_1^{c_2}}, b = \frac{1}{2} \ln \frac{W_+(T) c_1^{\neg c_2}}{W_-(T) c_1^{\neg c_2}}. \tag{6}$$

The weights of individual training instances are updated in accordance with equation (5), where x defines a set of instances and y refers to a set of prediction layers:

$$\omega_{i,t+1} = \omega_{i,t} e^{-r_t(x_i) y_t}, \tag{7}$$

where r denotes a base rule, $r(x)$ denotes the real value which the rule relates with the instance x, t represents boosting iteration, and ω is the weight of examples [72]. Eventually, the classification rule can be achieved as illustrated in equation (8) that represents the sign of the summation of the entire base rules:

$$\text{class}(x) = \text{sign}\left(\sum_{i=1}^{T} r_t(x) \right). \tag{8}$$

2.4.4. Validation Methods

(1) Area under the Receiver Operating Characteristic Curve (AUC-ROC). The AUC-ROC curve is used as a prevalent validation technique for landslide susceptibility models [76, 77]. The current study also employed AUC-ROC for evaluating performance of the developed models. The values of sensitivity and 100% specificity were first plotted, and then, AUC was employed for quantitative evaluation [73]. The AUC value of 1 suggests a perfect model, whereas a value <0.5 indicates an incorrect model [78, 79].

(2) Statistical Indices. The common statistical indices for landslide susceptibility models, namely, Kappa (k), Specificity (SPF), Sensitivity (SST), Negative Predictive Value (NPV), Positive Predictive Value (PPV), Root Mean Square Error (RMSE), and Accuracy (ACC) were chosen for validation of the models [80, 81]. Out of these, the reliability of the models for landslide prediction was assessed using k value that varies between "0" and "1" [34]. If k is closer to "1," it indicates that the landslide models are more reliable, and if the values of Kappa are closer to 0, it indicates that the landslide models are not corrected. Furthermore, SPF and NPV were used to evaluate the goodness of landslide models for classification of nonlandslide pixels, SST and PPV were used for landslide pixels, RMSE indicates the error analysis of the models, and ACC was used to validate the general

accuracy of the model. Detailed description and calculation of these indices are presented in the published works [80, 82].

3. Results

3.1. Evaluation of Goodness of Fit of the Models. Using the training dataset, three landslide susceptibility models, namely, NB, MLP, and ADT were developed for spatial landslide prediction and evaluated for the goodness of fit of the models. The statistical evaluation of the developed models has been listed in Table 2 and Figures 5–8. The classification probability of the landslide class i.e., PPV for ADT and MLP is almost similar, which reflects the strong classifying capability of the landslide class as compared to NB. For the NPV, the MLP model indicates that the probability to correctly classify the pixels in the training data related to the nonlandslide class is 85.29%, followed by the ADT model (83.53%) and the NB model (80%).

The MLP model shows the highest SST (84.38%) demonstrating that the landslide points are accurately classified to the landslide class though in comparison to the ADT and NB models which are 82.93% and 78.88%, respectively. The maximum SPF is for the ADT model describing that 80.68% nonlandslide points are accurately classified to the nonlandslide class. MLP and NB models showed the SPF of 80.56% and 75.98%, respectively.

The MLP model manifests the highest values of the Kappa index of 0.647, followed by ADT (0.6353) and NB (0.547) models representing a substantial agreement between the predictions and reality (Figure 5). The minimum value of RMSE for the training data was observed as 0.364 for the MLP model, followed by 0.372 and 0.416 for ADT and NB models, respectively (Figure 6). The highest value of ACC is for the MLP model (82.35%), followed by ADT (81.76%), and NB (77.35%) (Table 2). From the ROC curve analysis, it can be seen that the ADT model received the highest value of AUC (0.902), followed by the MLP (0.886) and the NB (0.802) (Figure 8). From these results, it can be stated that all models have a high goodness of fit with the data used, but the MLP and ADT are slightly better than NB.

3.2. Validation of Predictive Capability of the Models. Using the testing dataset, the predictive capability of the landslide susceptibility models was tested. The results are illustrated in Tables 3 and Figures 5–8. The maximum ACC in the validation dataset is for the MLP model (83.55%), followed by the NB model (79.61%) and the ADT model (78.95%). The MLP model has the maximum PPV (81.58%), signifying the probability to accurately classify pixels to the landslide class at 81.58%. The value of PPV for ADT is similar to MLP, whereas it is 80.26% for the NB model (Table 2). The highest probability to classify pixels to the nonlandslide class is also for the MLP model (85.53%). The MLP model has the highest SST (84.93%) indicating that 84.93% of the landslide points are correctly classified into the landslide class whereas the NB and ADT can accurately classify 79.22% and 77.50%, respectively. SPF is 82.28% for

the MLP model supporting 82.28% nonlandslide points which are correctly categorized to the nonlandslide class whereas this value is 80% and 80.56% for NB and ADT, respectively. Figure 7 also shows the RMSE values of the models. The Kappa indices of 0.647, 0.635, and 0.547 for the MLP, the ADT, and the NB models, respectively, reveal a significant agreement between observed and anticipated landslides (Figure 5). The AUC-ROC results indicate that all three models have good prediction capabilities with the maximum for the ADT model (AUC = 0.902), followed by the MLP model (AUC = 0.886) and NB model (AUC = 0.829) (Figure 8). Overall, all three models have demonstrated close agreement between the predicted and actual values. However, the NB and MLP models showed better performance in the validation datasets comparatively.

3.3. Construction of Landslide Susceptibility Maps. Results of the studied landslide models were used to develop landslide susceptibility maps. Initially, landslide susceptibility indices (LSI) were created from the training of the models, and each pixel was assigned a distinctive LSI. In the subsequent step, the geometrical intervals' (GIs) method was used to reclassify LSI into different ranges [83]. On the basis of these intervals, five different susceptible classes were identified for developing landslide susceptibility maps, namely, very low, low, moderate, high, and very high (Figure 9).

Landslide susceptibility maps developed on the basis of the models were validated by Frequency Ratio (FR) analysis of each susceptible class for each model on the maps (Figure 10). FR for each susceptible class on the map was obtained from equation (9). Each class represents a ratio of the percentage of landslide pixels (P_{LS} %) to the percentage of all pixels of class (P_P %):

$$FR = \frac{P_{LS}}{P_p}. \tag{9}$$

It can be observed from the Figure 10 that the map produced by MLP and ADT models are the most reliable as they received the highest value of FR on the very high class compared with another model (NB).

4. Discussion

Landslide susceptibility maps are deemed as a final product that can assist authorities in land-use planning. The development of these maps is generally considered as the first important step in landslide hazard and risk evaluation for better landslide management [84, 85]. So, it is desirable to develop and use a number of models which can forecast correctly landslide susceptible areas. Since the precision of a variety of suggested methods for landslide susceptibility modeling is under discussion, investigations of existing and novel techniques for this purpose are very important to reach some satisfactory conclusions [86]. The recent advancement in the field of GIS and ML techniques has been helpful in performance of the models [87]. Some advanced innovative ML methods, for instance, ADT and KLR have been employed in several fields with high accuracy [88–91].

TABLE 2: Accuracy analysis of the models for the training dataset.

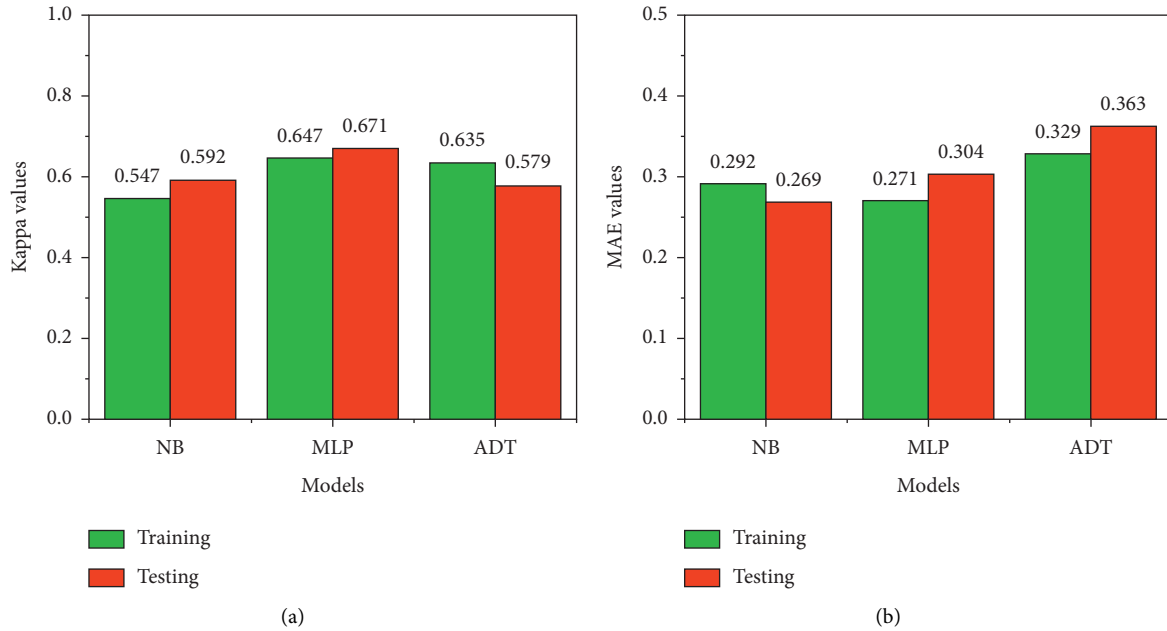| Parameters		PPV (%)	NPV (%)	SST (%)	SPF (%)	ACC (%)	Kappa (k)	MAE	RMSE
	NB	74.71	80	78.88	75.98	77.35	0.547	0.292	0.416
Training dataset	MLP	79.41	85.29	84.38	80.56	82.35	0.647	0.271	0.364
	ADT	80	83.53	82.93	80.68	81.76	0.635	0.329	0.372

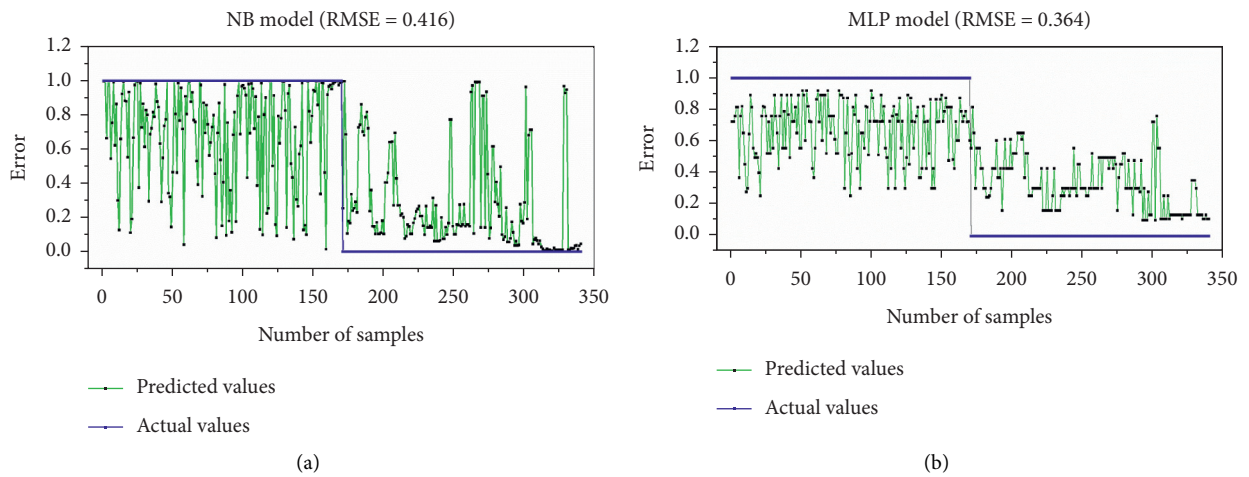FIGURE 5: Compare results of the models using the validating dataset: (a) Kappa values and (b) MAE values.
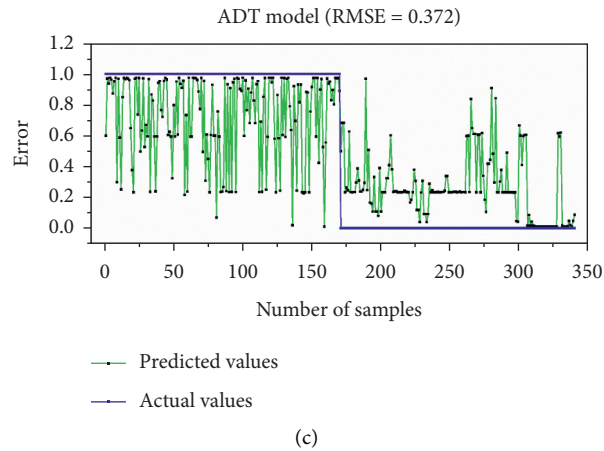

FIGURE 6: Continued.

(c)

FIGURE 6: RMSE analysis of the models using the training dataset: (a) NB model, (b) MLP model, and (c) ADT model.
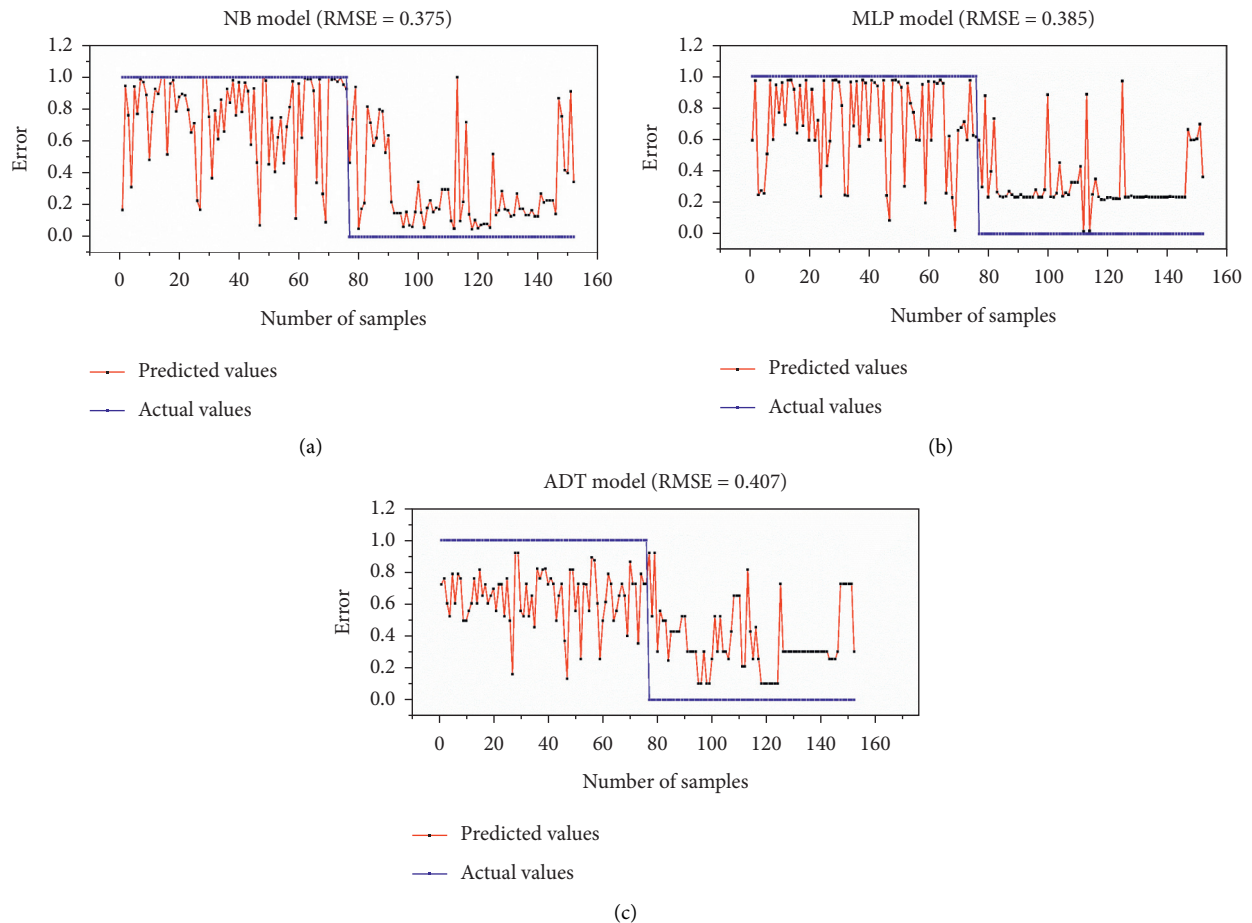

(a)

(b)

(c)

FIGURE 7: RMSE analysis of the models using validating dataset: (a) NB model, (b) MLP model, and (c) ADT model.

ML models such as NB, MLP, and ADT have also been employed in some landslide susceptibility studies [40, 73]. However, the literature lacks the comparison of these three models in the assessment of landslide susceptibility at the study area. This issue is addressed in the current research by analyzing and comparing the NB, ADT, and MLP techniques for landslide susceptibility modeling in the study area of Pithoragarh District of Uttarakhand State, India.

Overall, all three applied models performed well for landslide susceptibility mapping. However, the performance of the ADT and MLP models is slightly better than the NB model. It is reasonable as both of the ADT and MLP are

NB (AUC = 0.829)
MLP (AUC = 0.886)
ADT (AUC = 0.902)
--- Reference Line

(a)

NB (AUC = 0.873)
MLP (AUC = 0.864)
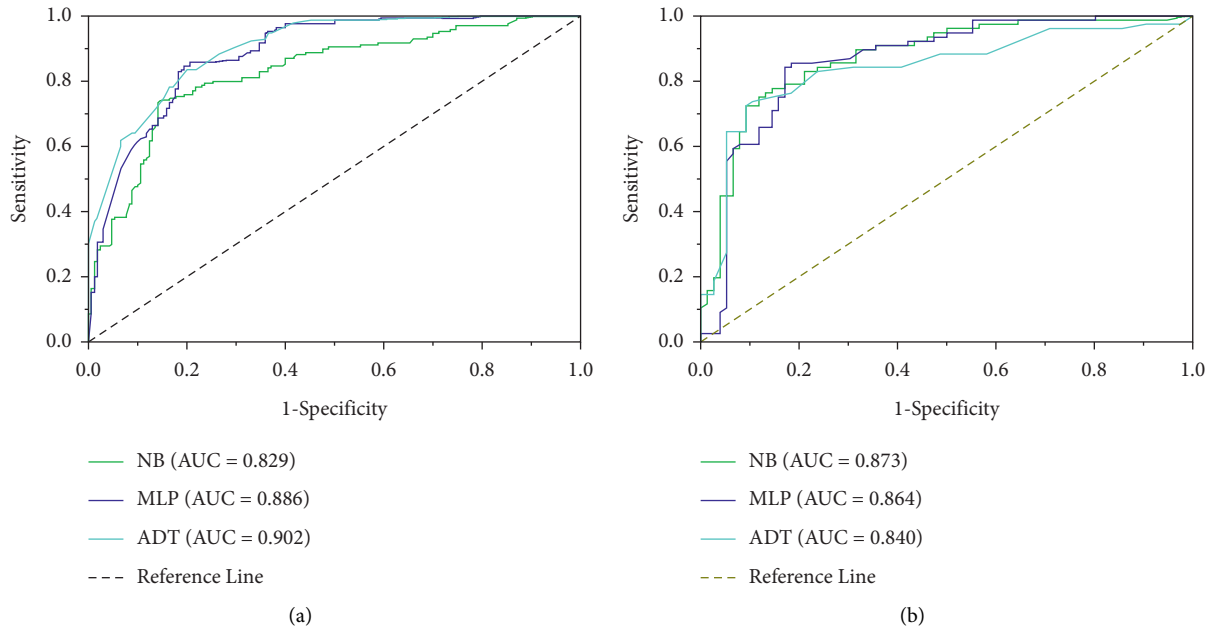ADT (AUC = 0.840)
--- Reference Line

(b)

FIGURE 8: AUC analysis of the models using (a) training dataset and (b) validating dataset.

TABLE 3: Accuracy analysis of the models for the testing dataset.

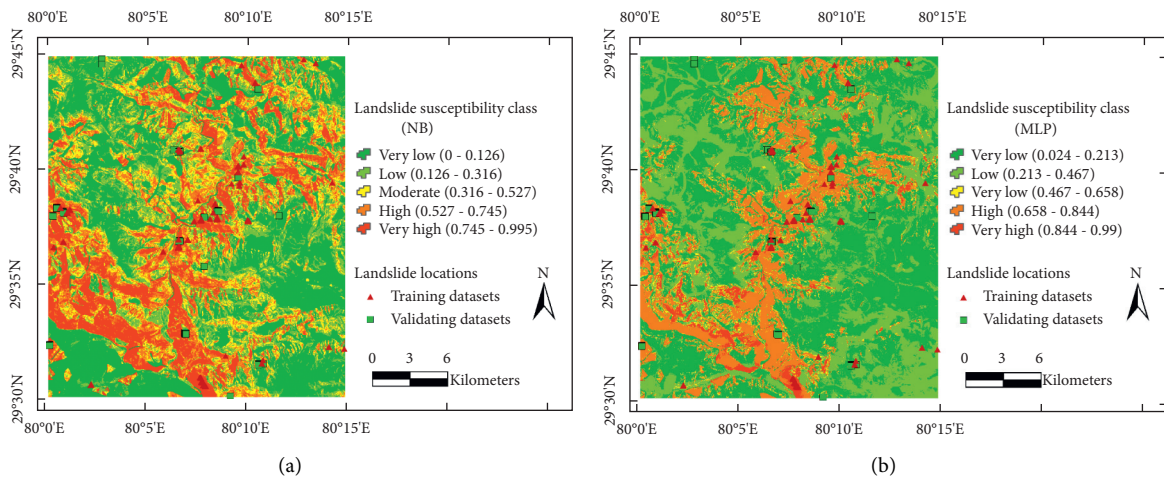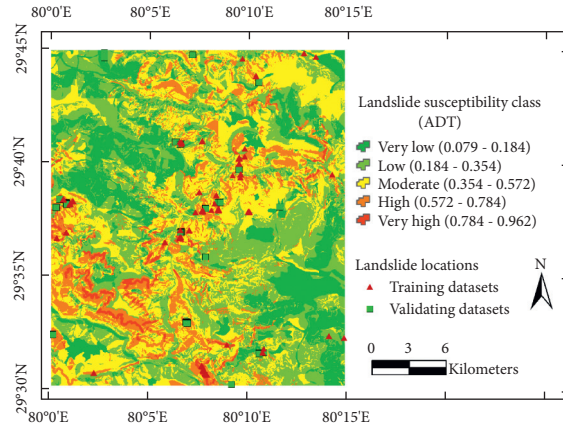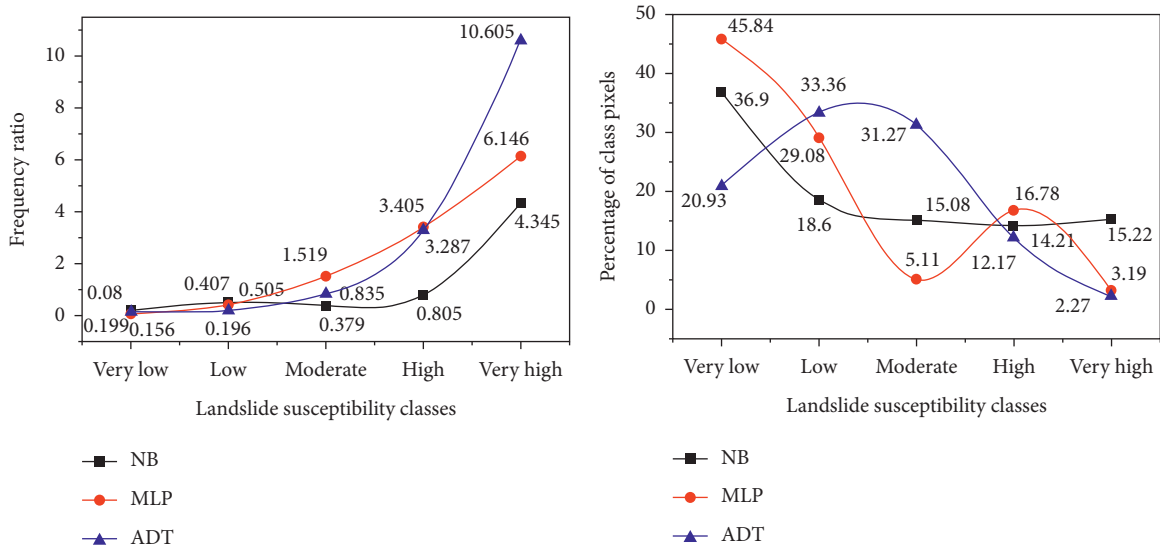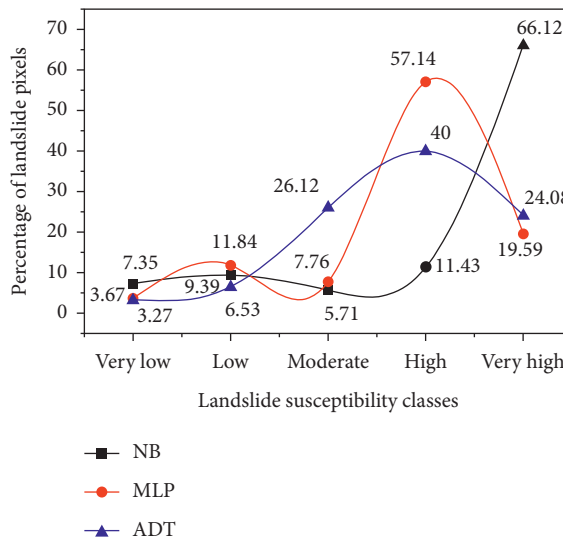| Parameters		PPV (%)	NPV (%)	SST (%)	SPF (%)	ACC (%)	Kappa (k)	MAE	RMSE
	NB	80.26	78.95	79.22	80	79.61	0.592	0.269	0.375
Testing dataset	MLP	81.58	85.53	84.93	82.28	83.55	0.671	0.304	0.385
	ADT	81.58	76.32	77.5	80.56	78.95	0.579	0.363	0.407

(a)

(b)

FIGURE 9: Continued.

(c)

FIGURE 9: Landslide susceptibility maps using the models: (a) NB, (b) MLP Classifier, (c) ADT.

(a)

(b)

(c)

FIGURE 10: Analysis of FR on the susceptibility maps using the models: (a) FR values of classes of the maps, (b) percentage of class pixels of the maps, and (c) percentage of landslide pixels of the maps.

TABLE 4: Hyperparameters used in the applied models of this study.

Hyperparameters	Models		
	NB	MLP	ADT
Batch size	100	100	100
Debug	False	False	False
Display model in old format	False	—	—
Do not check capabilities	False	False	False
Num decimal places	2	2	2
Use kernel estimator	False	—	—
Use supervised discretization	False	—	—
Activation function	—	Approximate sigmoid	—
Loss function	—	Squared error	—
Num function	—	2	—
Num threads	—	1	—
Pool size	—	1	—
Ridge	—	0.01	—
Seed	—	1	1
Tolerance	—	1.0E-6	—
Use CGD	—	False	—
Num of booting iterations	—	—	10
Resume	—	—	False
Save instance data	—	—	False
Search path	—	—	Expand all paths

advanced and popular ML models with many advantages. In the case of the ADT, it has the handling of large input data with the proper assessment of the accuracy of the output [92]. In addition, in the ADT, each decision rule can be analyzed in isolation as it is known explicitly [93]. With the case of MLP, it has an adapting learning, which is able to learn how to do the tasks on the given data. However, in this case, no assumption is made about the pattern classes for the underlying probability density functions [94].

Comparison with the published works shows that the modeling results of this study are comparative with other published works. In the study of Thai Pham et al. [76] for landslide susceptibility of Luc Yen district of northern part of Viet Nam, it obtained the AUC value of 0.827 for the ADT model. Similarly, Hong et al. [41] obtained the AUCs' equals of 0.957, 0.925, and 0.888 for the ADT, the KLR, and the SVM models. Hong et al. [41] evaluated the values of Kappa indices as 0.741, 0.702, and 0.626 for the ADT, the KLG, and the SVM models, respectively, indicating a substantial agreement between the models and the reality. Pham et al. [95] also compared MLP with NB for landslide susceptibility mapping and stated that the MLP (AUC = 0.850) is slightly better than the NB (AUC = 0.838). In [73], it was also stated that the ADT is slightly better than NBT, which is a hybrid model of NB and DT, for landslide susceptibility modeling. From these comparisons, we can state that although the ML models have good accuracy prediction for landslide susceptibility, however, their predictive capability is different from case to case, which depends on the data used for each model. In addition, we insisted that the performance of the ML models depends on the selection of the hyperparameters used to train the models. In this work, the hyperparameters of the applied models are selected by the trial-error process as presented in Table 4.

5. Concluding Remarks

In this study, we compared three popular ML models, namely, NB, MLP, and ADT for landslide susceptibility modeling and mapping at the Pithoragarh District, Uttarakhand State, India. Several validation indices including the AUC-ROC curve were used for validation and comparison of the models. The results infer that all the three models performed well for landslide susceptibility mapping. However, the ADT and MLP are slightly better than NB. Therefore, it can be concluded that the ADT and MLP models are promising techniques that could be used for landslide susceptibility modeling and mapping. The developed maps would be useful in identifying future landslide-prone areas for the better landslide management, prevention of causalities and damages to infrastructure, and proper land-use planning of the area. Model development is a continuous process. Therefore, it is proposed to carryout similar studies by comparing these models with other soft computing models in other areas also for its wider application.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] T. Li, *Landslide Management in the Mountain Areas of China*, ICIMOD, Lalitpur, Nepal, 1990.

[2] D. P. Kanungo, M. K. Arora, R. P. Gupta, and S. Sarkar, "Landslide risk assessment using concepts of danger pixels and fuzzy set theory in Darjeeling Himalayas," *Landslides*, vol. 5, no. 4, pp. 407–416, 2008.

[3] P. De Vita, P. Reichenbach, J. C. Bathurst et al., "Rainfall-triggered landslides: a reference list," *Environmental Geology*, vol. 35, no. 2-3, pp. 219–233, 1998.

[4] S. L. Gariano and F. Guzzetti, "Landslides in a changing climate," *Earth-Science Reviews*, vol. 162, pp. 227–252, 2016.

[5] J. Corominas and J. Moya, "A review of assessing landslide frequency for hazard zoning purposes," *Engineering Geology*, vol. 102, no. 3-4, pp. 193–213, 2008.

[6] Y. Huang and L. Zhao, "Review on landslide susceptibility mapping using support vector machines," *Catena*, vol. 165, pp. 520–529, 2018.

[7] J. Roy, S. Saha, A. Arabameri, T. Blaschke, and D. T. Bui, "A novel ensemble approach for landslide susceptibility mapping (LSM) in Darjeeling and Kalimpong districts, West Bengal, India," *Remote Sensing*, vol. 11, no. 23, Article ID 2866, 2019.

[8] B. C. Ozer, B. Mutlu, H. A. Nefeslioglu et al., "On the use of hierarchical fuzzy inference systems (HFIS) in expert-based landslide susceptibility mapping: the central part of the Rif Mountains (Morocco)," *Bulletin of Engineering Geology and the Environment*, vol. 79, no. 1, pp. 551–568, 2020.

[9] R. Roslee, A. C. Mickey, N. Simon, and M. N. Norhisham, "Landslide susceptibility analysis LSA using weighted overlay method wom along the genting sempah to bentong highway pahang," *Malaysian Journal Geosciences*, vol. 1, no. 2, pp. 13–19, 2017.

[10] P. Tsangaratos, I. Ilia, H. Hong, W. Chen, and C. Xu, "Applying Information Theory and GIS-based quantitative methods to produce landslide susceptibility maps in Nancheng County, China," *Landslides*, vol. 14, no. 3, pp. 1091–1111, 2017.

[11] B. Ahmed, "Landslide susceptibility mapping using multi-criteria evaluation techniques in Chittagong Metropolitan Area, Bangladesh," *Landslides*, vol. 12, no. 6, pp. 1077–1095, 2015.

[12] P. Wongsasuluk, S. Chotpantarat, W. Siriwong, and M. Robson, "Using hair and fingernails in binary logistic regression for bio-monitoring of heavy metals/metalloid in groundwater in intensively agricultural areas, Thailand," *Environmental Research*, vol. 162, pp. 106–118, 2018.

[13] Y. Achour, A. Boumezbeur, R. Hadji, A. Chouabbi, V. Cavaleiro, and E. A. Bendaoud, "Landslide susceptibility mapping using analytic hierarchy process and information value methods along a highway road section in Constantine, Algeria," *Arabian Journal of Geosciences*, vol. 10, no. 8, p. 194, 2017.

[14] K. Kavoura and N. Sabatakakis, "Investigating landslide susceptibility procedures in Greece," *Landslides*, vol. 17, no. 1, pp. 127–145, 2020.

[15] A. Arabameri, B. Pradhan, K. Rezaei, M. Sohrabi, and Z. Kalantari, "GIS-based landslide susceptibility mapping using numerical risk factor bivariate model and its ensemble with linear multivariate regression and boosted regression tree algorithms," *Journal of Mountain Science*, vol. 16, no. 3, pp. 595–618, 2019.

[16] B. T. Pham and I. Prakash, "Evaluation and comparison of logitboost ensemble, fisher's linear discriminant analysis, logistic regression and support vector machines methods for landslide susceptibility mapping," *Geocarto International*, vol. 34, no. 3, pp. 316–333, 2019.

[17] L. Dahoua, S. V. Yakovitch, R. Hadji, and Z. Farid, "Landslide susceptibility mapping using analytic hierarchy process method in BBA-bouira region, case study of East-West Highway, NE Algeria," in *Euro-Mediterranean Conference for Environmental Integration*, Springer, 2017.

[18] P. Kadavi, C.-W. Lee, and S. Lee, "Application of ensemble-based machine learning models to landslide susceptibility mapping," *Remote Sensing*, vol. 10, no. 8, p. 1252, 2018.

[19] A. M. S. Pradhan and Y.-T. Kim, "Spatial data analysis and application of evidential belief functions to shallow landslide susceptibility mapping at Mt. Umyeon, Seoul, Korea," *Bulletin of Engineering Geology and the Environment*, vol. 76, no. 4, pp. 1263–1279, 2017.

[20] Z. Fang, Y. Wang, L. Peng, and H. Hong, "A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping," *International Journal of Geographical Information Science*, vol. 35, no. 2, pp. 321–347, 2021.

[21] Q. Wang, Y. Guo, W. Li, J. He, and Z. Wu, "Predictive modeling of landslide hazards in Wen County, northwestern China based on information value, weights-of-evidence, and certainty factor," *Geomatics, Natural Hazards and Risk*, vol. 10, no. 1, pp. 820–835, 2019.

[22] B. T. Pham, C. Luu, T. V. Phong et al., "Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in Quang nam Province, Vietnam," *Journal of Hydrology*, vol. 592, Article ID 125815, 2021.

[23] E. J. M. Carranza and M. Hale, "Geologically constrained probabilistic mapping of gold potential, Baguio district, Philippines," *Natural Resources Research*, vol. 9, no. 3, pp. 237–253, 2000.

[24] F. Guzzetti, R. Paola, A. Francesca, C. Mauro, and G. Mirco, "Estimating the quality of landslide susceptibility models," *Geomorphology*, vol. 81, no. 1-2, pp. 166–184, 2006.

[25] B. Pradhan, A. Chaudhari, J. Adinarayana, and M. F. Buchroithner, "Soil erosion assessment and its correlation with landslide events using remote sensing data and GIS: a case study at Penang Island, Malaysia," *Environmental Monitoring and Assessment*, vol. 184, no. 2, pp. 715–727, 2012.

[26] S. Wan, T. C. Lei, and T. Y. Chou, "Optimized object-based image classification: development of landslide knowledge decision support system," *Arabian Journal of Geosciences*, vol. 7, no. 5, pp. 2059–2070, 2014.

[27] M. H. Tangestani, "A comparative study of dempster-shafer and fuzzy models for landslide susceptibility mapping using a GIS: an experience from Zagros Mountains, SW Iran," *Journal of Asian Earth Sciences*, vol. 35, no. 1, pp. 66–73, 2009.

[28] B. Neuhäuser and B. Terhorst, "Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic escarpment (SW-Germany)," *Geomorphology*, vol. 86, no. 1-2, pp. 12–24, 2007.

[29] L. Chu, L.-J. Wang, J. Jiang, X. Liu, K. Sawada, and J. Zhang, "Comparison of landslide susceptibility maps using random forest and multivariate adaptive regression spline models in combination with catchment map units," *Geosciences Journal*, vol. 23, no. 2, pp. 341–355, 2019.

[30] S. Lee, S.-M. Hong, and H.-S. Jung, "A support vector machine for landslide susceptibility mapping in Gangwon Province, Korea," *Sustainability*, vol. 9, no. 1, p. 48, 2017.

[31] Á. M. Felicísimo, A. Cuartero, J. Remondo, and E. Quirós, "Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study," *Landslides*, vol. 10, no. 2, pp. 175–189, 2013.

[32] S. Park and J. Kim, "Landslide susceptibility mapping based on random forest and boosted regression tree models, and a comparison of their performance," *Applied Sciences*, vol. 9, no. 5, p. 942, 2019.

[33] G. Wang, X. Chen, and W. Chen, "Spatial prediction of landslide susceptibility based on gis and discriminant functions," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, p. 144, 2020.

[34] D. Tien Bui, P. Biswajeet, L. Owe, and R. Inge, "Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and Naive Bayes Models," *Mathematical Problems in Engineering*, vol. 2012, Article ID 974638, 26 pages, 2012.

[35] A. Kornejady, M. Ownegh, and A. Bahremand, "Landslide susceptibility assessment using maximum entropy model with two different data sampling methods," *Catena*, vol. 152, pp. 144–162, 2017.

[36] H. Hong, Y. Miao, J. Liu, and A.-X. Zhu, "Exploring the effects of the design and quantity of absence data on the performance of random forest-based landslide susceptibility mapping," *Catena*, vol. 176, pp. 45–64, 2019.

[37] H. Moayedi, M. Mehrabi, M. Mosallanezhad, A. S. A. Rashid, and B. Pradhan, "Modification of landslide susceptibility mapping using optimized PSO-ANN technique," *Engineering with Computers*, vol. 35, no. 3, pp. 967–984, 2019.

[38] D. C. Camilo, L. Lombardo, P. M. Mai, J. Dou, and R. Huser, "Handling high predictor dimensionality in slope-unit-based landslide susceptibility models through LASSO-penalized generalized linear model," *Environmental Modelling & Software*, vol. 97, pp. 145–156, 2017.

[39] V.-H. Nhu, A. Shirzadi, H. Shahabi et al., "Shallow landslide susceptibility mapping: a comparison between logistic model tree, logistic regression, naïve bayes tree, artificial neural network, and support vector machine algorithms," *International Journal of Environmental Research and Public Health*, vol. 17, no. 8, p. 2749, 2020.

[40] M. Zare, H. R. Pourghasemi, M. Vafakhah, and B. Pradhan, "Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms," *Arabian Journal of Geosciences*, vol. 6, no. 8, pp. 2873–2888, 2013.

[41] H. Hong, B. Pradhan, C. Xu, and D. Tien Bui, "Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines," *Catena*, vol. 133, pp. 266–281, 2015.

[42] R. Senouci, N.-E. Taibi, A. C. Teodoro, L. Duarte, H. Mansour, and R. Yahia Meddah, "GIS-based expert knowledge for landslide susceptibility mapping (LSM): case of mostaganem coast district, west of Algeria," *Sustainability*, vol. 13, no. 2, p. 630, 2021.

[43] M. Sahana, B. T. Pham, M. Shukla et al., "Rainfall induced landslide susceptibility mapping using novel hybrid soft computing methods based on multi-layer perceptron neural network classifier," *Geocarto International*, pp. 1–25, 2020.

[44] Z. Chen and J. Wang, "Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada," *Natural Hazards*, vol. 42, no. 1, pp. 75–89, 2007.

[45] A. K. Saha, R. P. Gupta, I. Sarkar, M. K. Arora, and E. Csaplovics, "An approach for GIS-based statistical landslide susceptibility zonation?with a case study in the Himalayas," *Landslides*, vol. 2, no. 1, pp. 61–69, 2005.

[46] I. Das, S. Sahoo, C. van Westen, A. Stein, and R. Hack, "Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system, along a road section in the Northern Himalayas (India)," *Geomorphology*, vol. 114, no. 4, pp. 627–637, 2010.

[47] S. Sarkar, A. K. Roy, and T. R. Martha, "Landslide susceptibility assessment using information value method in parts of the Darjeeling Himalayas," *Journal of the Geological Society of India*, vol. 82, no. 4, pp. 351–362, 2013.

[48] A. Dikshit, R. Sarkar, B. Pradhan, S. Segoni, and A. M. Alamri, "Rainfall induced landslide studies in Indian Himalayan region: a critical review," *Applied Sciences*, vol. 10, no. 7, p. 2466, 2020.

[49] S. Das, "Uttarakhand tragedy," *Journal of the Geological Society of India*, vol. 82, no. 2, p. 201, 2013.

[50] C. Xu, X. Xu, F. Dai et al., "Application of an incomplete landslide inventory, logistic regression model and its validation for landslide susceptibility mapping related to the May 12, 2008 Wenchuan earthquake of China," *Natural Hazards*, vol. 68, no. 2, pp. 883–900, 2013.

[51] T. Chen, L. Zhu, R.-Q. Niu, C. J. Trinder, L. Peng, and T. Lei, "Mapping landslide susceptibility at the Three Gorges Reservoir, China, using gradient boosting decision tree, random forest and information value models," *Journal of Mountain Science*, vol. 17, no. 3, pp. 670–685, 2020.

[52] G. Wang, X. Lei, W. Chen, H. Shahabi, and A. Shirzadi, "Hybrid computational intelligence methods for landslide susceptibility mapping," *Symmetry*, vol. 12, no. 3, p. 325, 2020.

[53] S. V. Razavi-Termeh, A. Sadeghi-Niaraki, and S.-M. Choi, "Ubiquitous GIS-based forest fire susceptibility mapping using artificial intelligence methods," *Remote Sensing*, vol. 12, no. 10, p. 1689, 2020.

[54] A. M. Youssef and H. R. Pourghasemi, "Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at abha basin, asir region, Saudi Arabia," *Geoscience Frontiers*, vol. 12, no. 2, pp. 639–655, 2021.

[55] M. Kasai, M. Ikeda, T. Asahina, and K. Fujisawa, "LiDAR-derived DEM evaluation of deep-seated landslides in a steep and rocky region of Japan," *Geomorphology*, vol. 113, no. 1-2, pp. 57–69, 2009.

[56] J.-H. Lee, M. I. Sameen, B. Pradhan, and H.-J. Park, "Modeling landslide susceptibility in data-scarce environments using optimized data mining and statistical methods," *Geomorphology*, vol. 303, pp. 284–298, 2018.

[57] H. Hong, B. Pradhan, M. I. Sameen, W. Chen, and C. Xu, "Spatial prediction of rotational landslide using geographically weighted regression, logistic regression, and support vector machine models in Xing Guo area (China)," *Geomatics, Natural Hazards and Risk*, vol. 8, no. 2, pp. 1997–2022, 2017.

[58] X. Yu and H. Gao, "A landslide susceptibility map based on spatial scale segmentation: a case study at Zigui-Badong in the Three Gorges Reservoir Area, China," *PLoS One*, vol. 15, no. 3, Article ID e0229818, 2020.

[59] B. T. Pham, D. Tien Bui, I. Prakash, and M. B. Dholakia, "Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS," *Catena*, vol. 149, pp. 52–63, 2017.

[60] K. Shirani, M. Pasandi, and A. Arabameri, "Landslide susceptibility assessment by dempster-shafer and index of entropy models, sarkhoun basin, southwestern Iran," *Natural Hazards*, vol. 93, no. 3, pp. 1379–1418, 2018.

[61] S. Beguería, "Changes in land cover and shallow landslide activity: a case study in the Spanish Pyrenees," *Geomorphology*, vol. 74, no. 1-4, pp. 196–206, 2006.

[62] Z. Liu, G. Gilbert, J. M. Cepeda et al., "Modelling of shallow landslides with machine learning algorithms," *Geoscience Frontiers*, vol. 12, no. 1, pp. 385–393, 2021.

[63] M. I. Sameen, R. Sarkar, B. Pradhan, D. Drukpa, A. M. Alamri, and H.-J. Park, "Landslide spatial modelling using unsupervised factor optimisation and regularised greedy forests," *Computers & Geosciences*, vol. 134, Article ID 104336, 2020.

[64] H. R. Pourghasemi, B. Pradhan, C. Gokceoglu, M. Mohammadi, and H. R. Moradi, "Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran," *Arabian Journal of Geosciences*, vol. 6, no. 7, pp. 2351–2365, 2013.

[65] C. Borgelt, M. Steinbrecher, and R. R. Kruse, *Graphical Models: Representations for Learning, Reasoning and Data Mining*, John Wiley & Sons, Hoboken, NJ, USA, 2009.

[66] C. Catal, U. Sevim, and B. Diri, "Practical development of an Eclipse-based software fault prediction tool using Naive Bayes algorithm," *Expert Systems with Applications*, vol. 38, no. 3, pp. 2347–2353, 2011.

[67] T. Shatovskaya, V. Repka, and A. Good, "Application of the bayesian networks in the informational modeling," in *Proceedings of the 2006 International Conference-Modern Problems of Radio Engineering, Telecommunications, and Computer Science*, IEEE, Lviv, UKraine, March 2006.

[68] C. Zhang, X. Pan, H. Li et al., "A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 140, pp. 133–144, 2018.

[69] M. W. Gardner and S. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," *Atmospheric Environment*, vol. 32, no. 14-15, pp. 2627–2636, 1998.

[70] G. Panchal, A. Ganatra, Y. P. Kosta, and D. Panchal, "Behaviour analysis of multilayer perceptronswith multiple hidden neurons and hidden layers," *International Journal of Computer Theory and Engineering*, vol. 3, no. 2, pp. 332–337, 2011.

[71] D. T. Bui, B. Pradhan, I. Revhaug, D. B. Nguyen, H. V. Pham, and Q. N. Bui, "A novel hybrid evidential belief function-based fuzzy logic model in spatial prediction of rainfall-induced shallow landslides in the Lang Son city area (Vietnam)," *Geomatics, Natural Hazards and Risk*, vol. 6, no. 3, pp. 243–271, 2015.

[72] Y. Freund and L. Mason, *The Alternating Decision Tree Learning Algorithm*, ICML, Velur, India, 1999.

[73] W. Chen, X. Xie, J. Peng, J. Wang, Z. Duan, and H. Hong, "GIS-based landslide susceptibility modelling: a comparative assessment of kernel logistic regression, Naïve-Bayes tree, and alternating decision tree models," *Geomatics, Natural Hazards and Risk*, vol. 8, no. 2, pp. 950–973, 2017.

[74] T. Van Gestel, B. Baesens, J. Suykens et al., "Bankruptcy prediction with least squares support vector machine classifiers," in *Proceedings of the 2003 IEEE International Conference on Computational Intelligence for Financial Engineering*, IEEE, Hong Kong, China, March 2003.

[75] B. T. Pham, D. Tien Bui, M. B. Dholakia, I. Prakash, and H. V. Pham, "A comparative study of least square support vector machines and multiclass alternating decision trees for spatial prediction of rainfall-induced landslides in a tropical cyclones area," *Geotechnical & Geological Engineering*, vol. 34, no. 6, pp. 1807–1824, 2016.

[76] B. Thai Pham, D. Tien Bui, and I. Prakash, "Landslide susceptibility modelling using different advanced decision trees methods," *Civil Engineering and Environmental Systems*, vol. 35, no. 1-4, pp. 139–157, 2018.

[77] S. Beguería, "Validation and evaluation of predictive models in hazard assessment and risk management," *Natural Hazards*, vol. 37, no. 3, pp. 315–329, 2006.

[78] H. Hong, J. Liu, D. T. Bui et al., "Landslide susceptibility mapping using J48 decision tree with AdaBoost, bagging and rotation forest ensembles in the guangchang area (China)," *Catena*, vol. 163, pp. 399–413, 2018.

[79] D. T. Bui, T.-C. Ho, B. Pradhan, B.-T. Pham, V.-H. Nhu, and I. Revhaug, "GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with adaboost, bagging, and multiboost ensemble frameworks," *Environmental Earth Sciences*, vol. 75, no. 14, pp. 1–22, 2016.

[80] V.-H. Nhu, A. Mohammadi, H. Shahabi et al., "Landslide detection and susceptibility modeling on cameron highlands (Malaysia): a comparison between random forest, logistic regression and logistic model tree algorithms," *Forests*, vol. 11, no. 8, p. 830, 2020.

[81] D.-H. Lee, Y.-T. Kim, and S.-R. Lee, "Shallow landslide susceptibility models based on artificial neural networks considering the factor selection method and various nonlinear activation functions," *Remote Sensing*, vol. 12, no. 7, p. 1194, 2020.

[82] P. Yariyan, S. Janizadeh, T. Van Phong et al., "Improvement of best first decision trees using bagging and dagging ensembles for flood probability mapping," *Water Resources Management*, vol. 34, no. 9, pp. 3037–3053, 2020.

[83] N. Kazakis and K. S. Voudouris, "Groundwater vulnerability and pollution risk assessment of porous aquifers to nitrate: modifying the DRASTIC method using quantitative parameters," *Journal of Hydrology*, vol. 525, pp. 13–25, 2015.

[84] R. Fell, C. Jordi, B. Christophe, C. Leonardo, L. Eric, and W. Z. Savage, "Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning," *Engineering Geology*, vol. 102, no. 3-4, pp. 99–111, 2008.

[85] G. Manzo, V. Tofani, S. Segoni, A. Battistini, and F. Catani, "GIS techniques for regional-scale landslide susceptibility assessment: the Sicily (Italy) case study," *International Journal of Geographical Information Science*, vol. 27, no. 7, pp. 1433–1452, 2013.

[86] D. Tien Bui, T. A. Tuan, H. Klempe, B. Pradhan, and I. Revhaug, "Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree," *Landslides*, vol. 13, no. 2, pp. 361–378, 2016.

[87] I. H. Witten and E. Frank, "Data mining," *Acm Sigmod Record*, vol. 31, no. 1, pp. 76–77, 2002.

[88] C.-J. F. Chung, A. G. Fabbri, and C. J. Van Westen, "Multivariate regression analysis for landslide hazard zonation," in *Geographical Information Systems in Assessing Natural Hazards*, pp. 107–133, Springer, Berlin, Germany, 1995.

[89] R. T. Guy, P. Santago, and C. D. Langefeld, "Bootstrap aggregating of alternating decision trees to detect sets of SNPs

that associate with disease," *Genetic Epidemiology*, vol. 36, no. 2, pp. 99–106, 2012.

[90] K.-Y. Liu, J. Lin, X. Zhou, and S. T. C. Wong, "Boosting alternating decision trees modeling of disease trait information," *BMC Genetics*, vol. 6, no. Suppl_1, 2005.

[91] M. Maalouf and T. B. Trafalis, "Robust weighted kernel logistic regression in imbalanced and rare events data," *Computational Statistics & Data Analysis*, vol. 55, no. 1, pp. 168–183, 2011.

[92] A. Arabameri, E. Karimi-Sangchini, S. C. Pal et al., "Novel credal decision tree-based ensemble approaches for predicting the landslide susceptibility," *Remote Sensing*, vol. 12, no. 20, p. 3389, 2020.

[93] M. P.-L. Ooi, H. K. Sok, Y. C. Kuang, and S. Demidenko, "Alternating decision trees," in *Handbook of Neural Computation*, pp. 345–371, Elsevier, Amsterdam, Netherlands, 2017.

[94] D. S. Jayas, J. Paliwal, and N. S. Visen, "Review paper (AE-Automation and emerging technologies)," *Journal of Agricultural Engineering Research*, vol. 77, no. 2, pp. 119–128, 2000.

[95] B. T. Pham, T. B. Dieu, P. Hamid Reza, I. Prakash, and M. B. Dholakia, "Landslide susceptibility assesssment in the Uttarakhand area (India) using GIS: a comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods," *Theoretical and Applied Climatology*, vol. 128, no. 1-2, pp. 255–273, 2017.