Article

Landslide Susceptibility Mapping Using Machine Learning Algorithm Validated by Persistent Scatterer In-SAR Technique

Muhammad Afaq Hussain 1, Zhanlong Chen 1,*, Ying Zheng 1, Muhammad Shoaib 2, Safeer Ullah Shah 3, Nafees Ali 4 and Zeeshan Afzal 5

1 School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China; khanafaq121@cug.edu.cn (M.A.H.); 2016100229@icug.edu.cn (Y.Z.)
2 State Key Laboratory of Hydraulic Engineering, Simulation and Safety, School of Civil Engineering, Tianjin University, Tianjin 300072, China; xxs4shoaib@tju.edu.cn
3 Ministry of Climate Change, Islamabad 44000, Pakistan; safeershah@uop.edu.pk
4 Chinese Academy of Sciences, Beijing 100045, China; nafeesali@mails.ucas.ac.cn
5 Key State Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University; Wuhan 430079, China; zeeshanafzal@whu.edu.cn
* Correspondence: chenzl@cug.edu.cn

Abstract: Landslides are the most catastrophic geological hazard in hilly areas. The present work intends to identify landslide susceptibility along Karakorum Highway (KKH) in Northern Pakistan, using landslide susceptibility mapping (LSM). To compare and predict the connection between causative factors and landslides, the random forest (RF), extreme gradient boosting (XGBoost), k nearest neighbor (KNN) and naive Bayes (NB) models were used in this research. Interferometric synthetic aperture radar persistent scatterer interferometry (PS-InSAR) technology was used to explore the displacement movement of retrieved models. Initially, 332 landslide areas alongside the Karakorum Highway were found to generate the landslide inventory map using various data. The landslides were categorized into two sections for validation and training, of 30% and 70%. For susceptibility mapping, thirteen landslide-condition factors were created. The area under curve (AUC) of the receiver operating characteristic (ROC) curve technique was utilized for accuracy comparison, yielding 83.08, 82.15, 80.31, and 72.92% accuracy for RF, XGBoost, KNN, and NB, respectively. The PS-InSAR technique demonstrated a high deformation velocity along the line of sight (LOS) in modelsensitive areas. The PS-InSAR technique was used to evaluate the slope deformation velocity, which can be used to improve the LSM for the research region. The RF technique yielded superior findings, integrating with the PS-InSAR outcomes to provide the region with a new landslide susceptibility map. The enhanced model will help mitigate landslide catastrophes, and the outcomes may help ensure the roadway’s safe functioning in the study region.

Keywords: CPEC; random forest; landslides; susceptibility; PS-InSAR; ArcGIS

1. Introduction

The China–Pakistan Economic Corridor (CPEC) demonstrates the flagship project of the “One Belt, One Road” policy. It is also thought to hold Pakistan’s financial prospects, which are receiving a lot of interest. The Karakoram Highway was built in 1974–1978 and inaugurated in 1979 and runs alongside the CPEC in Northern Pakistan. It is regularly closed for a few months each year because of landslides.

Local topography, tectonic features, geomorphology, landcover, geology and human interference all have an influence on the spatial likelihood of landslides, which is then examined to determine landslide susceptibility (LS) [1]. Landslide vulnerability assessment models frequently assume that historical and current landslide conditions would be constant in the future [2]. LS methodologies can be quantitative or qualitative; quantitative methods evaluate the likelihood of landslide incidence in a susceptible zone, whereas
Sensors 2022, 22, 3119

qualitative methods introduce subjectivity into illustrative susceptibility zonation \[1,3\]. The analytical hierarchy model (AHP) \[4–6\], weight of evidence model \[7,8\], frequency ratio \[4,5,9\] and certainty factor \[4\] are all commonly used landslide susceptibility models. A growing trend is to compare the outcomes of implementing two or more models and the result is a landslide susceptibility model (LSM). However, most studies still use only one model for LSM \[1,10\]. Reichenbach et al. \[1\] suggest using numerous models to assess landslides and developing an “optimal” zonation map to reduce risk prediction errors and its integrity to be used for land-use planning. Our literature review of the investigated area demonstrates several statistical approaches for LS such as frequency ratio and weight of evidence \[11,12\], AHP, and Scoops3D \[13\], the weighted overlay technique, and the AHP \[14\] were used in the research region. Several investigations \[15,16\] provide bivariate analyses that measure the geographical links between particular variables and landslides that influence their occurrence. However, the key disadvantages of these models are that they change the ambiguity of risk processes, are typically static, incorporate geometrical assumptions, and are costly and difficult for the gathering of hydrological and geotechnical data, especially when examining vast and different locations.

In recent years, advances in ML algorithms, computing power, and geospatial innovations have made it easier to create landslide susceptibility (LS) maps \[17\]. The precision of LS maps can be improved using machine learning algorithms. Knowledge-based methods \[18\], multivariate logistic regression methods \[19–21\] and multivariate binary logistic regression \[22\] have all been presented in recent papers. General linear model \[23,24\], quadratic discriminant analysis \[10,24\], boosted regression tree \[23,25\], random forest \[26–29\], multivariate adaptive regression splines \[30,31\], classification and regression tree \[23,32\], support vector machine \[33–35\], naive Bayes \[36,37\], generalized additive model \[24,32\], neuro-fuzzy and adaptive neuro-fuzzy inference \[38–40\], fuzzy logic \[41\], artificial neural networks \[42–47\], classification and regression tree \[23,32\], support vector machine \[33–35\], naive Bayes \[36,37\], generalized additive model \[24,32\], neuro-fuzzy and adaptive neuro-fuzzy inference \[38–40\], fuzzy logic \[41\], artificial neural networks \[42–47\], maximum entropy \[48,49\] and decision tree \[19,50,51\] are some of the ML models used in LSM. Qing et al. \[52\] used various ML techniques for LSM alongside the China–Pakistan Karakoram Highway. In two South Korean catchments, Pradhan and Kim \[53\] compared the precision consequences of deep neural network (83.71%), and XGBoost (76.73%) approaches for LS mapping. Merghadi et al. \[54\] assessed the performance and competency of various ML techniques in the literature and discovered that tree-based ensemble optimization algorithms outcompete other ML algorithms. In a comparison analysis, Sahin \[55\] found that CatBoost had the best precision (85%), followed by XGBoost (83.36%) since the proportion of samples of the model was determined by Catboost was more precisely anticipated than other models. The primary advantages of ML and probabilistic processes are their objective statistical foundation, repeatability, capacity to quantitatively analyze the effect of variables on landslide evolution, and capacity to update them regularly. Machine learning models can be built using a variety of landslide-conditioning factors (slope, aspect, and elevation). Several studies on landslide susceptibility evaluation have been undertaken using remote sensing and GIS techniques \[56–58\].

Furthermore, remote sensing (RS) is an effective method for determining the motion of landslides \[59–61\]. It provides a solution in surveys or enhanced detection in places where catastrophic landslides occur frequently and quickly \[62–65\]. Furthermore, interferometric algorithms to radar images effectively map large-scale landslide mapping and detection. It may aid in the development of landslide inventory maps. In particular, decrypt ADInSAR and PS-InSAR \[66,67\], coherence pixel technique \[68\], SqueeSAR \[69\], small baseline subset \[70,71\], Stanford method for persistent scatterers \[72\], stable point network \[73,74\], and interferometric point target analysis \[75\] have created various useful case research. As noted in past studies \[75–77\], these approaches are involved with mapping and identifying landslide occurrences.

A diversity of researchers in northern Pakistan has analyzed landslides using historical records, field observations, tectonic characteristics, and geological data \[78–81\]. Previous studies \[14,82–85\] concentrated on probabilistic and statistical relationships and regression interpretation of landslides with parameters. For the first time, the PS-InSAR approach eval-
uated the surface displacement in the study area using RF, XGBoost KNN, and NB models, making it a distinctive method of identifying landslide movements. Persistent scatterer interferometry (PSI), interferometric synthetic aperture radar (InSAR), and area under curve (AUC) of ROC techniques were used along the KKH to assess displacements and the precision of the models used. Single landslides in hazardous areas can be identified and defined using PS-InSAR. Landslides can also be detected using a spatial statistical method based on a multitemporal assessment of SAR images that calculate slow landslide movements [71].

The current work seeks to develop a susceptibility model and a complete visually interpreted landslide inventory utilizing recently developed ML models, including RF, XGBoost, KNN, and NB. The second goal is to quantify the deformation velocities of slow-moving landslides using PS-InSAR to identify high-susceptibility zones for future landslide disaster management. The third goal is to select the most susceptible model based on accuracy and AUC value and then combine it with PS-InSAR outcomes to produce a new landslide susceptibility map for the research area. These prediction approaches will help lead future development and land management efforts in the area. These susceptibility maps will aid in avoiding and limiting human and economic losses along this critical corridor.

2. Methods
2.1. Study Area

The research region is 178 km long and has a 5 km radius buffer zone along KKH (Figure 1). The KKH in northern Pakistan is a crucial component of the CPEC; nevertheless, it is frequently disrupted due to several hydro-climatological and geological risks along the route. Landslides are the most common and devastating to highways, human lives, and economic activity.

The research region experiences harsh winters and mild summers. The region’s annual rainfall ranges from 120 mm to 130 mm: the maximum and minimum temperatures vary from 16 °C to −21 °C (Meteorological Department of Pakistan). The lithology of various sources with thicknesses of up to 100 m is irregularly scattered [86,87]. The majority of these deposits are weakly consolidated, making them conducive to landslides in the form of rockfalls and debris flows [80]. The combination of complex topography, high erosion rates, human causes, and active tectonics makes this area one of the most susceptible to landslides.

2.2. Geological Setting of the Area

The rocks in the region are mostly Paleozoic, Proterozoic, and Mesozoic in age (Figure 2). According to the geological map prepared by Searle et al. [88], the study area is comprised of the following lithology.

The Chilas complex in the study area comprises mafic and ultramafic plutonic rocks and Kohistan batholiths composed of granodiorite, granite, and diorite. The Gilgit complex metasedimentary rocks are slates, minor phyllite, quartzite, and dolomite limestone. Komila amphibolite comprises of plutonic and meta plutonic rocks with intrusion of diorite granodiorite and granite. In Paleozoic metasedimentary rocks are marble, dolomite, and quartzite.

2.3. Landslide Susceptibility Mapping

Geological maps, remote sensing data, and meteorological data were gathered from various sources for the study (Table 1). The Alaska Satellite Facility dataset contained an Advanced Land Observing Satellite, Phased Array type L-band Synthetic Aperture Radar DEM (Digital Elevation Model) with a resolution of 12.5 m (https://search.asf.alaska.edu/ (accessed on 20 January 2022)). Sentinel-2 images with a resolution of 10 m were extracted from the USGS (https://earthexplorer.usgs.gov (accessed on 20 January 2022)) dataset to create a landcover map for the research area. The geological map for the area was digitized in the ArcGIS environment to comprehend surficial geological characteristics. PS-InSAR processing was used to compute the deformation velocity using Sentinel-1
Figure 1. The 541 band combination Landsat image showing the area under investigation. (a) Pakistan, (b) District boundaries, (c) Study area in red outline.
Figure 2. Regional geological map of the study area.

Table 1. Information on landslide conditioning factors.

| S.NO | Factors | Description/Extraction | Category |
|------|---------|-------------------------|----------|
| 1    | Elevation, aspect, curvature, slope, profile curvature, TWI, plan curvature, roughness | ALOS-PALSAR DEM ([https://search.asf.alaska.edu/](https://search.asf.alaska.edu/) (accessed on 20 January 2022)) | Topography |
| 2    | Geology, distance to fault | Geological Survey of Pakistan | Geology |
| 3    | Landcover | Land cover classes ([https://earthexplorer.usgs.gov/](https://earthexplorer.usgs.gov/) (accessed on 20 January 2022)) | Conditioning factor |
| 4    | NDVI | Normalized Different Vegetation Index ([Landsat-8, 2021](https://earthexplorer.usgs.gov/)) | Landcover |
| 5    | Precipitation | Annual rainfall ([Pakistan Meteorlogical Department](https://earthexplorer.usgs.gov/)) | Triggered factor |
2.4. Landslide Inventory

The landslide inventory is the stage in estimating susceptibility since it provides details on all sorts of past landslides in the research area. This is the most important stage since the required precision of the landslide inventory to fine tune the models influences the LSM accuracy [15,89,90]. As a result, the more accurate and high-quality the landslide inventory, the more improved the prediction execution of the SM [89]. The evaluation of landslide hazard begins with creating realistic and detailed landslide inventory maps that show the type of landslide, geographic extension, the date of the event, and location [89,91]. The produced landslide inventory maps are then associated with contributing geo-environmental parameters such as land cover, topography, geology, geomorphology, and other factors to assess the likelihood of terrain causing a landslide allocated to a susceptibility level [1,9,92–94].

Inventory maps contain information on all active and historical landslide distributions based on field surveying, aerial image interpretation, and previous report data [80]. In this research, we were using actual data of landslide occurrences obtained from the Geological Survey of Pakistan (GSP) publications [95–97], Frontier Works Organization road clearance logs, a research article [98], and Google Earth imagery to produce a multitemporal landslide inventory along the highway. On the other hand, the landslide inventory was created by the visual interpretation of Sentinel-2 photos with 10 m resolution (2020) and from Google Earth and was validated using earlier reports and a field assessment of the research area. Polygon shapes were constructed on satellite images for clearly visible landslides (based on GSP and FWO data). Debris flow (188 locations) and scree slopes (51 locations) are mostly found in the research area as a result of unconsolidated sediments on barren mountains and rainfall, although rock falls (93 locations) are also common as a result of seismic activity and toe cutting of steep slopes by anthropogenic activities for various causes (Figure 4). There were 332 landslides mapped, shown in Figure 5. Of these landslides, 30% (100 landslides) and 70% (232 landslides) were chosen for training for model validation [99].

---

Figure 3. Flow chart of research.
common as a result of seismic activity and toe cutting of steep slopes by anthropogenic activities for various causes (Figure 4). There were 332 landslides mapped, shown in Figure 5. Of these landslides, 30% (100 landslides) and 70% (232 landslides) were chosen for training for model validation [99].

Figure 4. Showing the debris flow, rockfall, and scree slopes in the research area.

Figure 5. Landslide inventory of study area. (a) Shows the rockfall and scree slopes near the Sazin area, (b) rockfall and debris flow near the Drang area.

2.5. Landslide Causative Factors

Landslides' spatial distribution is influenced by triggering, and conditioning variables were chosen based on the region's morphology, geology, hydrology, and anthropogenic activities. There are no general criteria for choosing independent factors for LSM [71]. The concepts
are that the factors must be non-redundant, non-uniform, operational, and measurable [100]. ArcGIS is commonly used to extract important susceptibility conditioning factors from digital elevation models, including elevation, slope, profile curvature, aspect, curvature, and topographic wetness index (TWI) [101]. Land cover, geology, precipitation, roughness, normalized difference vegetation index (NDVI), distance to faults, TWI, slope, plan curvature, curvature, elevation, profile curvature, and aspect were all utilized to estimate the landslides’ disaster susceptibility in the research area (Table 1). All of these maps were converted to a 12.5 × 12.5 m pixel raster format for the models, which was up to digital elevation model resolution. In the resampling method, the cell size of each factor was kept at 12.5 m so that the overlay assessment would obtain the pixels at the same scale, and the output was also the same scale. The maps were digitized at various scales, and the pixel resolution was kept at 12.5 m while converting them to raster format. The DEM with a pixel of 12.5 m was used to extract the majority of the factors, and all other factors were brought to a similar resolution. The thirteen landslide factors are depicted in Figures 6 and 7.

Figure 6. Landslides factors used in the research study.

The modeling procedure included machine learning model fitting, identification, and development.

I. The model unit in this investigation was the grid unit (12.5 m). The spatial resolution of DEM and RS data corresponds to 12.5 m, and all assessment variables have been recalculated at this level.

II. A condition property reached thirteen causative variables and a landslide decision attribute (1 indicates landslides, 0 indicates non-landslides), with each row creating an object.

III. Each column represents an object’s attribute and has been converted into training (70%) and testing the two-dimensional matrix (30%). Training data were used to assemble the models, and test data was used to make forecasts.
IV. The landslide susceptibility index maps were created using the forecast values of every model unit per group. The findings of the four algorithms were exported into GIS.

V. The Jenks natural breaks [102] classifications were used to categorize LS: very low, low, moderate, high, and very high. The ROC curve and the area under the ROC curve were used to test the four models.

Figure 7. Landslides factors used in the research area.

2.6. RF

One of the most widely used methods for regression and classification is the random forest, which was designed by [103]. RF has a lot of important features for classification tasks. Because RF is a non-parametric, non-linear approach, it can handle big datasets with numerical and category data and complicated nonlinearity and interactions between factors. Secondly, it can deal with the situation with more predictors than data and integrate the connection between different predictors. Third, random forest can manage missing values while maintaining precision for missing data.

Furthermore, unlike other ML approaches such as support vector machines and artificial neural networks, RF does not need extensive hyper-parameter tuning. In many instances, utilizing the default parameter values yields good results. When compared
to other tree-ensemble approaches (boosting), random forest is relatively fast. Decision trees are built during the training phase, and the output class is based on the classification or regression mode of each individual tree’s decision trees. In order to train random forests, the general method of bootstrap aggregation (also known as “bagging”) is used for tree learners. This bootstrapping approach improves model performance by reducing the model’s variance without raising the model’s bias [104]. Random forest has been extensively utilized for classification applications and large-scale mapping in LSM [26,28] ecology [105], flood mapping [106] and soil science [107].

R statistical software was used to develop the RF model [108]. Because the analysis in the RF model was grid-based, gridded cells (12.5 by 12.5 m) were derived from the randomly shaped sample spatial polygons of landslides and non-landslides, respectively.

The RF model operates by developing numerous decorrelated decision trees as a base learner, with replacement, utilizing a percentage of randomly chosen landslide-predicting variables and landslide observation. Every tree was trained using two-thirds of the randomly chosen training samples, while the remaining one-third of the training samples, called out-of-bag (OOB), was utilized to verify the prediction result. Finally, a pixel was assigned to a class using the majority vote or mode rule [109]. In this research, this model has employed the “randomForest” package in R-studio.

Table 2 lists the three significant parameters: the number of features that are appropriate for dividing (mtry), the minimum number of a sample can also be taken arbitrarily in each bootstrap sample to balance any tree with recursive portioning (ntree) [110].

| Parameters | Values |
|------------|--------|
| Node size  | 14     |
| mtry       | 10     |
| ntree      | 500    |

2.7. XGBoost

According to Stanford statistics professor Fridman, the gradient boosting algorithm was developed in 2001 to estimate gradient descent approaches [111]. As the supervised classification model in this work, the XGBoost approach [112] was applied. The approach was invented by the gradient tree boosting algorithm [113,114], which is a powerful machine learning approach. It employs the regularized boosting strategy to prevent overfitting and improve model precision. XGBoost provides scalability for various scenarios, sparse data handling, thorough documentation, minimal computing resource requirements, good performance (i.e., speed), and easy implementation [112]. The approach was chosen since it has won several data science contests [112]. Further adjustments to the approach are needed for extremely unbalanced datasets (e.g., [115]).

Algorithms that boost or lift data are known as “lifting tree models” or “XGBoost”. Their key innovations are summarized below [113].

I. They optimize their loss function.
II. The candidate split value may be quickly and accurately generated using the parallel approximation histogram method.
III. In addition to a novel sparsity-aware linear tree learning algorithm, they offer an efficient cache-aware block structure for out-of-core tree learning.

In this research, this model has employed the “XGBoost” package in R-studio. Several model preview parameters must be selected for the XGBoost model. User-friendly settings are needed for three of the most important ones: colsample_bytree (column ratio subsamples when each tree is constructed), nrounds (maximum number of iterations boosting), and subsamples (the training instance subsample ratio); (Table 3).
Table 3. Parameters used in XGBoost model.

| Parameters       | Values |
|------------------|--------|
| nround           | 210    |
| colsample_bytree | 1      |
| subsample        | 1      |
| max_depth        | 6      |
| eta              | 0.05   |
| gamma            | 0      |

2.8. KNN

The KNN algorithm is one of the most fundamental machine learning techniques. It has recently been used in several other disciplines, including LSM [116,117]. KNN uses the k nearest training examples in the components space as input. When it comes to classification difficulties, class membership possibilities describe the degree of uncertainty with which a particular given item may be assigned to any given class [118]. The attributes of the nearby data points are used to classify a data point using a KNN algorithm [119]. It is a more effective version of the ball tree idea [120] that may be used in bigger dimensions. The approach is commonly employed in SM applications [116], and the categorization of a data point’s nearest neighbors determines the chance of it being assigned to any class [118]. The data point chooses the categorization that classifies the greatest number of neighbors. The number of K will be determined through a tuning procedure to obtain better outcomes.

According to Chen et al. [121], they propose that in KNN, objects are evaluated based on the opinions of a majority of their immediate neighbors. The highest consistent closeness of its adjacent neighbors is used to assign the item. If k = 1, the object is solely transferred to the single contiguous neighbor’s class.

2.9. NB

NB is a statistical classifier predicated on the Bayesian principle [122]. The Bayes theorem enables this methodology, which is a classification method. The NB maintains that each attribute impacts classification outcomes individually to make estimating the posterior likelihood of observed instances in training data easier [123]. The conditional self-reliance assumption holds that all variables are completely self-sufficient of one another given the output class [124].

The NB technique’s most notable benefits include its robustness to noise and irrelevant variables, ease of use, and lack of reliance on time-consuming iterative procedures [125]. Numerous studies have used the NB approach for LSM [36,37,126]. The following equation can be used to estimate the spatial prediction of landslides using NB:

\[
y_{NB} = P(y_i) \prod_{i=1}^{n} P\left(\frac{x_i}{y_i}\right)
\]

where \( P(x_i/y_i) \) is the conditional probability of each attribute and \( P(y_i) \) is the prior probability of target class \( y_i \) (landslide).

2.10. PS-InSAR

PS-InSAR is an enhanced InSAR technology designed for gradual deformation monitoring or long-term displacement. InSAR is a time series-based method that is broadly classified into two classes: small baseline (SBAS) approaches that focus on spatial correlation and dispersed scattering and PS-InSAR techniques that work on the locations of persistent scatterers (PS) [127]. PSI is a multi-interferometric SAR technique that can estimate ground movement with millimeter precision [128]. The PS-InSAR process uses multitemporal SAR images wrapping the same region to analyze the consistency of the phase and amplitude, which identifies the pixels that are less influenced by spatiotemporal
decorrelation and then determines specific deformation details on the constituents of the phase which must be collectively evaluated and modeled to eliminate inconsistencies [129].

We employed Sentinel-1 C-band SAR pictures recorded along both ascending and descending orbit tracks in this investigation. To complete the analysis in C-band data, the PSI [68] requires at least 20 SAR pictures [130]. The PSI monitors surface displacement over months or years, accounting for signal noise, atmospheric, and topographic impacts. This sensor has a ground resolution of around 20 m in the azimuth direction and 5 m in the range direction [131]. This sensor has several acquisition modes, including interferometric wide (IW), wave (Wave), extra-wide swath (EW), and strip map (SM). This study gathered images from the Sentinel-1A IW sensor and analyzed them in SARPROZ software (12 days of temporal resolution). The line of sight (LOS) displacement velocity (V\text{LOS}) was determined using 0.7 as the coherence threshold in PS-InSAR processing, as shown in Table 4. The InSAR approach computes surface deformation values along the LOS; however, the deformation rate in the LOS direction is inadequate for representing the actual slope displacement [71]. The following equation was used to determine slope velocity (Vslope), which is actually deformation velocity [132]:

\[
Vslope = \frac{V\text{LOS}}{\cos \varnothing}
\]  

(2)

where VLOS is deformation and \( \varnothing \) is the incident angle.

Table 4. Details of PS-InSAR processing.

| Specification          | Ascending                  | Descending                 |
|------------------------|----------------------------|----------------------------|
| Temporal range         | 1 May 2020–20 May 2021     | 14 May 2020–9 May 2021     |
| No. of images          | 33                         | 31                         |
| No. of PS/DS           | 526,815                    | 450,990                    |
| Minimum VLOS (mm/year) | -98                        | -34                        |
| Maximum VLOS (mm/year) | 31                         | 73                         |

Finally, the calculated result was used for comparative analysis with susceptibility models generated by RF, XGBoost, KNN, and NB methods. The Vslope points were converted into 12.5 \times 12.5 grid cells to provide a more precise LSM-like ML model and integrated to enhance the susceptibility degree of those cells defined by ground deformation, minimizing missed alerts, while cells stable and consisting of high susceptibility degrees according to SAR interferometry were not altered [128].

3. Results

3.1. The Significance of Landslide Variables

To compute the significance of the landslide variables in this study, we utilized R-Studio Software. In comparison, the RF model performed better in estimating the relevance of each element in causing landslides.

Figure 8, using origin software, depicts the significance of the factors using the RF model. The slope and elevation had the greatest impact, according to Figure 8, and profile curvature, roughness, distance to fault, and NDVI were almost equal on landslides in the research region. The slope is critical for landslides in the region (Figure 8); it encourages landslides and makes an area susceptible to landslides. Weathered rocks and medium height frequently define high elevation zones, and slopes are usually overlaid by thin colluvium, making them more prone to landslides [112]. The barren ground is in close contact with climatological factors such as sunlight and precipitation, causing rock deterioration and increasing the likelihood of landslides [133]. Because shear zones and active faults strongly influence landslide activities in the region, the buffer class nearest to the fault line is more susceptible [14]. The bulk of the debris flow, rockfalls, and other slides in the area are caused by monsoon rains [134]. Annual average precipitation data
were utilized in this study, which found that while precipitation was not a significant causal factor, there were more landslides in locations with high precipitation (Figure 8). The aspect and plan curvature had a minor impact on the landslide in the studied region. The majority of landslides in the research area are northward facing and south-facing. Arabameri et al. [135] employed RF models for LSM in Iran and found that aspect has a minor impact on LSM. The Komila amphibolite and Gilgit complex metasedimentary rocks are the most vulnerable formations in the study area [11,12,14,101]. The rocks in research area are highly fractured and deformed.

Figure 8. Factors important in the research region using the RF model.

The outcomes of employing the four LSM models obtained using the LPI are depicted in Figure 9. The greater the LPI, the more probable it is that a landslide may happen [136]. The likelihood value of LS was categorized into five classes using the natural break (Jenks) [102] method: very high, high, moderate, low, and very low (Figure 10).
The precision of the maps was evaluated using a confusion matrix, as suggested by [137]. A confusion matrix illustrates the capabilities of the RF, XGBoost, KNN, and NB models during the training stage (Table 5). The RF model shows a high accuracy (0.830) in the research area. Validation was accomplished using the ROC approach [36]. A ROC curve is created in this approach by plotting “sensitivity” versus “specificity” on cut-off values, but it does not fully explain the model’s efficiency; so, the AUC of the ROC curve was utilized to analyze the quantitative functioning of the models [138]. A larger proportion of the area below the curve suggests that the model is more accurate. In contrast, a smaller
percentage of the area below the curve shows that the model is less accurate in predicting future occurrences of the phenomena [139]. The AUC of the prediction rate curve was determined to be 88.83, 87.44, 83.38, and 72.80% for the RF, XGBoost, KNN, and NB models, respectively (Figure 11).

Table 5. Confusion matrix of models.

| Models | Observation | Predicted | Accuracy |
|--------|-------------|-----------|----------|
|        | No  | Yes     |          |
| RF     | 35  | 12      | 0.830    |
|        | 43  | 235     |          |
| XGBoost| 33  | 13      | 0.821    |
|        | 45  | 234     |          |
| KNN    | 32  | 18      | 0.803    |
|        | 46  | 229     |          |
| NB     | 39  | 49      | 0.729    |
|        | 39  | 198     |          |

Figure 11. Receiver operating characteristic plots of models.

3.2. PS-InSAR Based Validation

PS-InSAR approaches were utilized to evaluate and verify the models by checking the displacement in the area. The Interferometric Synthetic Aperture Radar (InSAR) approach has been well documented for identifying and tracking mass movements during the previous decade due to its extensive high spatial–temporal resolution, spatial coverage, and operating
capacity in all-weather conditions [93]. Many PS-InSAR studies have been conducted to determine the temporal or spatial landslide deformation patterns or the kinematic resolution of slow-moving landslides to quantify the scale of slow-moving landslides [140]. The estimated result was compared to the RF model’s susceptibility model.

The line of sight (LOS) displacement velocity ($V_{LOS}$) was determined using 0.7 as the coherence threshold in PS-InSAR processing (Figure 12). PS-InSAR was also shown to be a useful technique for monitoring slow landslide movement in non-vegetation regions. The InSAR approach computes surface deformation values along the LOS; however, the deformation rate in the LOS direction is inadequate for representing the actual slope displacement [71].

**Figure 12.** Landslide deformation velocity map along LOS direction for ascending and descending paths using PS-InSAR.

PS-InSAR analysis was performed for both descending and ascending geometries, with VLOS approved deformation in the region. The total number of PS/DS target points acquired with LOS direction deformation results varying from $-98$ to $73$ mm/year was obtained. Using the transformation formula, the VLOS was changed to Vslope. The greatest slope deformation velocity was determined to be $-100$ mm/year. VLOS indicates just one direction’s deformation based on the satellite’s LOS, which is determined to evaluate slope orientation velocity (Vslope). Because most landslides or ground surface displacements occur along the direction of steep terrain in the event of landslide assertion, Vslope is the main ingredient employed to define landslide advancement. The calculated Vslope for ascending and descending pathways was added together (Figure 13). The only displacements in RF’s highly sensitive zone-produced susceptible model were depicted in an ultimate deformation map (Figure 14). The PSI findings revealed that most of the mapped landslides were in deforming zones, although slow-moving landslides were forecasted more precisely because of the Sentinel-1A sensor’s extended revisiting period.
Figure 13. Showing the deformation velocity along slope direction for both ascending and descending paths using PS-InSAR.

Figure 14. PSI distribution of LOS deformation velocity on Google Earth using RF landslide susceptibility index.
Finally, the RF-based LSM was combined with Vslope to improve the region’s precise susceptibility map. The Vslope points were converted into 12.5 × 12.5 grid cells to provide a more precise LSM-like RF model and integrated to enhance the susceptibility degree of those cells defined by ground deformation, minimizing missed alerts, while cells stable and consisting of high susceptibility degrees according to SAR interferometry were not altered [128]. The contingency matrix was used to improve an LSM for the region to a Vslope and RF-based susceptibility model (Figure 15). In other words, the degree of difference for each cell was evaluated using the newly created LSM, which was generated using the RF model.

![Figure 15. Final LSM via Vslope.](image)

4. Discussion

The findings demonstrate that the RF, XGBoost, KNN, and NB data-mining techniques have comparable precision for LSM along the KKH, with the RF outperforming the others in terms of AUC value and accuracy. Our results conform to the consequences outlined in other work [11,12,14,52,101].

The overall precision values found in this research (RF, 83.0%) were compared to Youssef et al. [23]; it was discovered that the precision values found in this study were higher than the RF (81.2%) in other ML models of the revealed research. Sevgen et al. [29] compared ANN, logistic regression, and RF for LSM and found that the RF model shows the best classification precision with respect to ANN and LR. Taalab et al. [26] evaluated the RF algorithm for landslide in northwest Italy and found that the RF model performed well compared to other tree-based models. Chen et al. [141] reassembled the random forest (RF), logistic model tree (LMT), and classification and regression tree (CART) models to map LS. The LMT (74%) and CART (73%) models showed slightly lower precision values than the RF model (77%); RF performed better in LSM. Zhang et al. [142] demonstrated that the random RF model outperformed the C5.0 decision tree model by comparing it to the C5.0 decision tree model. The RF technique has an advantage over other ML models. It can use multiple input parameters without removing them and provide a limited number of classes with good forecast precision [143]. This model’s categorization precision is determined by the training dataset’s type, scale, number, and accuracy. The combination of all appropriate parameters boosts the precision of this model. Furthermore, compared to other models, RF has a greater capacity to implement a large number of data [144]. Arabameri et al. [135]
employed RF models for spatial modeling of gully erosion in Iran and found that RF performed best. Zhang et al. [145], when compared to neural networks, decision trees, and the RF, obtained the best results for debris flow susceptibility with the RF method in Shigatse Area, China. In areas with debris flow and rockfalls, we discovered that the RF model is the best predictive technique for LSM.

Recently, InSAR approaches for producing and updating landslide inventories have been created [146]. The findings of the InSAR techniques are thought to be more precise [147], yielding susceptibility maps with high production precision. The LOS velocity statistics only reveal the velocities along the slopes in the highly sensitive zones of both models. PS-InSAR was also shown to be a useful technique for monitoring slow landslide movement in non-vegetation regions. The ROC curve AUC was used to evaluate prediction capabilities, and it predicts 88.83, 87.44, 83.38, and 72.80% for RF, XGBoost, KNN, and NB, respectively, confirming the model’s accuracy. The collected susceptibility map was categorized into five groups using the Jenks natural breaks [102]: very low, low, moderate, high, and very high. In comparison, the RF model performed better in estimating the relevance of each element in causing landslides.

LSM was performed using the RF, XGBoost, KNN, and NB in this work. Nonetheless, several limitations caused misclassification in the results, such as (1) the accuracy of the landslide inventory and (2) the accuracy of data connected to each landslide variable. Because of the severe environment along KKH, only 332 landslides have been mapped for this research region. It resulted in considerable misclassification inaccuracies for the LSM, emphasizing the significance of upgrading the LSM utilizing PS-InSAR results. Surprisingly, when paired with the PS-InSAR data, the novel LSM reduced misclassification in which landscape altered by slope deformity was categorized as extremely low and very low.

Another problem is that the landslide susceptibility mapping merely represents anticipated landslide dispersion in regions rather than interactive displacement processes through time. Variations in landslide behavior over time, on the other hand, are a serious challenge for decision-makers [148]. In conjunction with the PS-InSAR outputs, a new landslide susceptibility map can depict the real conditions of landslides. It can be designed for quantitative hazard assessment and preliminary landslide mapping at the province level [149].

The LSM creates a susceptibility map for landslides, identifies the important variables that cause landslides, and evaluates the effect and their contribution [27,150]. Land cover, geology, slope, precipitation, NDVI, distance to faults, elevation, curvature, plan curvature, TWI, profile curvature, roughness, and aspect were all utilized to estimate the probability of landslides disaster in the research area. The main contributors of landslides in the area are slope, elevation. The slope is critical for landslides in the region (Figure 8); it encourages landslides and makes an area susceptible to landslides. Weathered rocks and medium height frequently define high elevation zones, and slopes are usually overlaid by thin colluvium, making them more prone to landslides [112]. Because shear zones and active faults strongly influence landslide activities in the region, the buffer class nearest to the fault line is more susceptible [14].

Previous studies such as [13] in this area relied mainly on statistical models, and a considerable number of landslides were missing in the inventory. As a result of the inadequate landslide inventory, the LSM is ineffective. This work focused on complete mapping of landslides to identify primary landslide triggers and define high susceptible zones using the PS-InSAR approach, which will be used in the future to mitigate landslide risks in the region.

5. Conclusions

Landslides are one of Pakistan’s most devastating natural disasters, generating major risks to lives and socioeconomic damage each year. So far, the process of landslide mapping has been highly difficult and volatile to perform correct and quick estimation of landslides in most places. Multiple attempts have been made to improve reliability based on many forecast models for mapping the landslide susceptibility, targeting different locations for
this goal. Decision-makers must construct more relevant landslide susceptibility maps to improve the prediction model’s performance. This study concentrated on complete landslide mapping to determine the fundamental causes of landslides and designate high-risk zones, which will be useful in the future to mitigate landslide threats in the region. The study’s distinctive feature is that it provides more accurate LSM by employing ML models verified by PS-InSAR processing.

This study used RF, XGBoost, KNN, and NB ML algorithms to enhance the LSM of the Karakoram Highway using the PS-InSAR approach. This study assessed the vulnerability using elevation, precipitation, slope, land cover, roughness, NDVI, curvature, distance to faults, plan curvature, aspect, profile curvature, geology, and TWI. Slope, elevation, and profile curvature are the primary causes of landslides in the region. The susceptibility model created will be used to identify zones for construction growth and improved management planning along the KKH. The LSM illustrates the just forecast landslide distribution in regions, not the dynamic displacement process over time. Variations in landslide activity eventually, on the other hand, are a main consideration for decision-makers. The newly developed LSM, when merged with the PS-InSAR results, may show the true situation of landslides and should be utilized for quantitative hazard analysis and preparatory landslide mapping at the regional level. Geotechnical and other slope stabilization procedures are necessary to minimize future landslide catastrophes in an environment. We conclude that our approach can give valuable insights into highway safety measures.

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

Funding: This research was funded by the National Natural Science Foundation of China (No. 41871305); National key R & D program of China (No.2017YFC0602204); Fundamental Research Funds for the Central Universities, China University of Geo-sciences (Wuhan) (No. CUGQY1945); Opening Fund of Key Laboratory of Geological Survey and Evaluation of Ministry of Education; and Fundamental Research Funds for the Central Universities (No. GLAB2019ZR02).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in the study are available on request from the first and corresponding author. The data are not publicly available due to the thesis that is being prepared from these data.

Acknowledgments: The author appreciates Javaid Iqbal for his recommendations and suggestions during the writing of the manuscript. Additionally, we are thankful to Muzamil Hassan for providing Sarproz Software for the analysis.

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

References

1. Reichenbach, P.; Rossi, M.; Malamud, B.D.; Mihir, M.; Guzzetti, F. A review of statistically-based landslide susceptibility models. Earth-Sci. Rev. 2018, 180, 60–91. [CrossRef]

2. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B.; Li, Y.; Catani, F.; Pourghasemi, H.R. Landslide susceptibility modeling applying machine learning methods: A case study from Longjui in the Three Gorges Reservoir area, China. Comput. Geosci. 2018, 112, 23–37. [CrossRef]

3. Guzzetti, F.; Carrara, A.; Cardinali, M.; Reichenbach, P. Landslide hazard evaluation: A review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 1999, 31, 181–216. [CrossRef]

4. Wu, Y.; Li, W.; Liu, P.; Bai, H.; Wang, Q.; He, J.; Liu, Y.; Sun, S. Application of analytic hierarchy process model for landslide susceptibility mapping in the Gangu County, Gansu Province, China. Environ. Earth Sci. 2016, 75, 422. [CrossRef]

5. Wang, Q.; Li, W. A GIS-based comparative evaluation of analytical hierarchy process and frequency ratio models for landslide susceptibility mapping. Phys. Geogr. 2017, 38, 318–337. [CrossRef]
6. Nguyen, T.T.N.; Liu, C.-C. A new approach using AHP to generate landslide susceptibility maps in the Chen-Yu-Lan Watershed, Taiwan. Sensors 2019, 19, 505. [CrossRef]

7. Neuhäuser, B.; Dann, B.; Terhorst, B. GIS-based assessment of landslide susceptibility on the base of the weights-of-evidence model. Landslides 2012, 9, 511–528. [CrossRef]

8. Razavizadeh, S.; Solaimani, K.; Massironi, M.; Kavian, A. Mapping landslide susceptibility with frequency ratio, statistical index, and weights of evidence models: A case study in northern Iran. Environ. Earth Sci. 2017, 76, 499. [CrossRef]

9. Khan, H.; Shafique, M.; Khan, M.A.; Bacha, M.A.; Shah, S.U.; Calligaris, C. Landslide susceptibility assessment using Frequency Ratio, a case study of northern Pakistan. Egypt. J. Remote Sens. Space Sci. 2019, 22, 11–24. [CrossRef]

10. Rossi, M.; Gazzetti, F.; Reichenbach, P.; Mondini, A.C.; Peruccacci, S. Optimal landslide susceptibility zonation based on multiple forecasts. Geomorphology 2010, 114, 129–142. [CrossRef]

11. Hussain, M.L.; Shafique, M.; Bacha, A.S.; Chen, X.-Q.; Chen, H.-Y. Landslide inventory and susceptibility assessment using multiple statistical approaches along the Karakoram highway, northern Pakistan. J. Mt. Sci. 2021, 18, 583–598. [CrossRef]

12. Iqbal, J.; Cui, P.; Hussain, M.L.; Pourghasemi, H.R.; Pradhan, B. Landslide susceptibility assessment along the dubair-dud ishal section of the karakoram highway, northwestern himalayas, pakistan. Acta Geodyn. Geomater 2021, 18, 137–155. [CrossRef]

13. Rashid, B.; Iqbal, J.; Su, L.-J. Landslide susceptibility analysis of Karakoram highway using analytical hierarchy process and scoops 3D. J. Mt. Sci. 2020, 17, 1596–1612. [CrossRef]

14. Ali, S.; Biermanns, P.; Haider, R.; Reichert, K. Landslide susceptibility mapping by using a geographic information system (GIS) along the China–Pakistan Economic Corridor (Karakoram Highway), Pakistan. Nat. Hazards Earth Syst. Sci. 2019, 19, 999–1022. [CrossRef]

15. Regmi, A.D.; Devkota, K.C.; Yoshida, K.; Pradhan, B.; Pourghasemi, H.R.; Kumamoto, T.; Akgun, A. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. Arab. J. Geosci. 2014, 7, 725–742. [CrossRef]

16. Shirzadi, B.; Iqbal, J.; Cui, P.; Hussain, M.L.; Pourghasemi, H.R.; Pradhan, B.; Reicherter, K. How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? Geosci. Front. 2019, 10, 871–883. [CrossRef]

17. Kumar, R.; Anbalagan, R. Landslide susceptibility mapping using analytical hierarchy process (AHP) in Tehri reservoir rim region, Uttarakhand. J. Geol. Soc. India 2016, 87, 271–286. [CrossRef]

18. Tien Bui, D.; Pradhan, B.; Lofman, O.; Revhaug, I. Landslide susceptibility assessment in vietnam using support vector machines, adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study. Geosci. Front. 2019, 13, 159–178. [CrossRef]

19. Achour, Y.; Pourghasemi, H.R. How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? Geosci. Front. 2020, 11, 871–883. [CrossRef]

20. Kaviri, A.; Devkota, K.C.; Yoshida, K.; Pradhan, B.; Pourghasemi, H.R.; Kumamoto, T.; Akgun, A. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. Arab. J. Geosci. 2014, 7, 725–742. [CrossRef]

21. Mandal, S.; Mandal, K. Modeling and mapping landslide susceptibility zones using GIS based multivariate binary logistic regression, and artificial neural network methods at the Inje area, Korea. Environ. Earth Sci. 2015, 68, 1443–1464. [CrossRef]

22. Mandal, S.; Mandal, K. How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? Geosci. Front. 2019, 10, 871–883. [CrossRef]

23. Youssef, A.M.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Al-Katheeri, M.M. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. Landslides 2016, 13, 839–856. [CrossRef]

24. Pourghasemi, H.R.; Rahmati, O. Prediction of the landslide susceptibility: Which algorithm, which precision? Catena 2018, 162, 177–192. [CrossRef]

25. Park, S.; Kim, J. Landslide susceptibility mapping based on random forest and boosted regression tree models, and a comparison of their performance. Appl. Sci. 2019, 9, 942. [CrossRef]

26. Taalab, K.; Cheng, T.; Zhang, Y. Mapping landslide susceptibility and types using Random Forest. Big Earth Data 2018, 2, 159–178. [CrossRef]

27. Hussain, M.A.; Chen, Z.; Wang, R.; Shoaib, M. PS-InSAR-Based Validated Landslide Susceptibility Mapping along Karakorum Highway, Pakistan. Remote Sens. 2021, 13, 4129. [CrossRef]

28. Hussain, M.A.; Chen, Z.; Wang, R.; Shah, S.U.; Shoaib, M.; Ali, N.; Xu, D.; Ma, C. Landslide Susceptibility Mapping using Machine Learning Algorithm. Civ. Eng. J. 2022, 8, 209–224. [CrossRef]

29. Sevgen, E.; Kocaman, S.; Nefeslioglu, H.A.; Gokceoglu, C.J.S. A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression, ANN and random forest. Sensors 2019, 19, 3940. [CrossRef]

30. Felicisimo, Á.M.; Cuartero, A.; Remondo, J.; Quiros, E. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study. Landslides 2013, 10, 175–189. [CrossRef]
31. Conoscenti, C.; Ciaccio, M.; Carballo-Arias, N.A.; Gómez-Gutiérrez, Á.; Rotigliano, E.; Agnesi, V. Assessment of susceptibility to earth-flow landslide using logistic regression and multivariate adaptive regression splines: A case of the Belice River basin (western Sicily, Italy). *Geomorphology* 2015, 242, 49–64. [CrossRef]

32. Vorpahl, P.; Elsenbeer, H.; Märker, M.; Schröder, B. How can statistical models help to determine driving factors of landslides? *Ecol. Model.* 2012, 239, 27–39. [CrossRef]

33. Kalantar, B.; Pradhan, B.; Naghibi, S.A.; Motevalli, A.; Mansor, S. 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). *Geom. Nat. Hazards Risk* 2018, 9, 49–69. [CrossRef]

34. Ma, J.; Wang, Y.; Niu, X.; Jiang, S.; Liu, Z. A comparative study of mutual information-based input variable selection strategies for the displacement prediction of seepage-driven landslides using optimized support vector regression. *Stoch. Environ. Res. Risk Assess.* 2022, 1, 1–21. [CrossRef]

35. Ghasemian, B.; Shahabi, H.; Shirzadi, A.; Al-Ansari, N.; Jaafari, A.; Kress, V.R.; Geertsema, M.; Renoud, S.; Ahmad, A. A robust deep-learning model for landslide susceptibility mapping: A case study of Kurdistan Province, Iran. *Sensors* 2022, 22, 1573. [CrossRef]

36. Pham, B.T.; Pradhan, B.; Bui, D.T.; Prakash, I.; Dholakia, M. A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India). *Environ. Model. Softw.* 2016, 84, 240–250. [CrossRef]

37. Pham, B.T.; Tien Bui, D.; Pourghasemi, H.R.; Indra, P.; Dholakia, M. Landslide susceptibility assessment in the Uttarakhand area (India) using GIS: A comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. *Theor. Appl. Climatol.* 2017, 128, 255–273. [CrossRef]

38. Aghdam, I.N.; Varzandeh, M.H.M.; Pradhan, B. Landslide susceptibility mapping using an ensemble statistical index (Wi) and adaptive neuro-fuzzy inference system (ANFIS) model at Alborz Mountains (Iran). *Environ. Earth Sci.* 2016, 75, 553. [CrossRef]

39. Dehnavi, A.; Aghdam, I.N.; Pradhan, B.; Varzandeh, M.H.M. A new hybrid model using step-wise weight assessment ratio (SWARA) technique and adaptive neuro-fuzzy inference system (ANFIS) for regional landslide hazard assessment in Iran. *Catena* 2015, 135, 122–148. [CrossRef]

40. Mehrabi, M.; Pradhan, B.; Moayedi, H.; Alamri, A. Optimizing an adaptive neuro-fuzzy inference system for spatial prediction of landslide susceptibility using four state-of-the-art metaheuristic techniques. *Sensors* 2020, 20, 1723. [CrossRef]

41. Kumar, R.; Anbalagan, R. Landslide susceptibility zonation in part of Tehri reservoir region using frequency ratio, fuzzy logic and GIS. *J. Earth Syst. Sci.* 2015, 124, 431–448. [CrossRef]

42. Aditian, A.; Kubota, T.; Shinohara, Y. 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, 101–111. [CrossRef]

43. Arnone, E.; Francipane, A.; Scarbaci, A.; Puglisi, C.; Noto, L.V. Effect of raster resolution and polygon-conversion algorithm on landslide susceptibility mapping. *Environ. Model. Softw.* 2016, 84, 467–481. [CrossRef]

44. Bui, D.T.; Moayedi, H.; Kalantar, B.; Osouli, A.; Pradhan, B.; Nguyen, H.; Rashid, A.S.A. A novel swarm intelligence—Harris hawks optimization for spatial assessment of landslide susceptibility. *Sensors* 2019, 19, 3590. [CrossRef] [PubMed]

45. Moayedi, H.; Osouli, A.; Tien Bui, D.; Foong, L.K. Spatial landslide susceptibility assessment based on novel neural-metaheuristic geographic information system based ensembles. *Sensors* 2019, 19, 4698. [CrossRef]

46. Roshani, M.; Sattari, M.A.; Ali, P.J.M.; Roshani, G.H.; Nazemi, B.; Corniani, E.; Nazemi, E. Application of GMDH neural network technique to improve measuring precision of a simplified photon attenuation based two-phase flowmeter. *Flow Meas. Instrum.* 2020, 75, 101804. [CrossRef]

47. Charandabi, S.E.; Kamyar, K. Prediction of Cryptocurrency Price Index Using Artificial Neural Networks: A Survey of the Literature. *Eur. J. Bus. Manag.* Res. 2021, 6, 17–20. [CrossRef]

48. Park, N.-W. Using maximum entropy modeling for landslide susceptibility mapping with multiple geoenvironmental data sets. *Environ. Earth Sci.* 2015, 73, 937–949. [CrossRef]

49. Kornejady, A.; Ownegh, M.; Bahremand, A. Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. *Catena* 2017, 152, 144–162. [CrossRef]

50. Wu, X.; Ren, F.; Niu, R. Landslide susceptibility assessment using object mapping units, decision tree, and support vector machine models in the Three Gorges of China. *Environ. Earth Sci.* 2014, 71, 4725–4738. [CrossRef]

51. Shirzadi, A.; Soliamani, K.; Habibnejhad, M.; Kavian, A.; Chapi, K.; Shahabi, H.; Chen, W.; Khosravi, K.; Thai Pham, B.; Pradhan, B. Novel GIS based machine learning algorithms for shallow landslide susceptibility mapping. *Sensors* 2018, 18, 3777. [CrossRef] [PubMed]

52. Qing, F.; Zhao, Y.; Meng, X.; Su, X.; Qi, T.; Yue, D. Application of Machine Learning to Debris Flow Susceptibility Mapping along the China–Pakistan Karakoram Highway. *Remote Sens.* 2020, 12, 2933. [CrossRef]

53. Pradhan, A.M.S.; Kim, Y.-T. Rainfall-induced shallow landslide susceptibility mapping at two adjacent catchments using advanced machine learning algorithms. *ISPRS Int. J. Geo-Inf.* 2020, 9, 569. [CrossRef]

54. Merghari, A.; Yunus, A.P.; Dou, J.; Whiteley, J.; ThaiPham, B.; Bui, D.T.; Avtar, R.; Abderrahmane, B. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Sci. Rev.* 2020, 207, 103225. [CrossRef]

55. Sahin, E.K. Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. *Geocarto Int.* 2020, 35, 1–25. [CrossRef]
84. Ahmed, M.F.; Rogers, J.D.; Ismail, E.H. A regional level preliminary landslide susceptibility study of the upper Indus river basin. *Eur. J. Remote Sens.* **2014**, *47*, 343–373. [CrossRef]

85. Basharat, M.; Shah, H.R.; Hameed, N. Landslide susceptibility mapping using GIS and weighted overlay method: A case study from NW Himalayas, Pakistan. *Arab. J. Geosci.* **2016**, *9*, 292. [CrossRef]

86. Owen, L. Wet-sediment deformation of Quaternary and recent sediments in the Skardu Basin, Karakoram Mountains, Pakistan. In *Proceedings of the Glaciotectonics: Forms and Processes; Various Meetings of the Glaciotectonics Work Group*. Routledge: Norfolk, UK, 1988; pp. 123–147.

87. Hewitt, K. Catastrophic landslides and their effects on the Upper Indus streams, Karakoram Himalaya, northern Pakistan. *Quat. Res.* **1999**, *51*, 220–237. [CrossRef]

88. Searle, M.; Khan, M.A.; Fraser, J.; Gough, S.; Jan, M.Q. The tectonic evolution of the Kohistan-Karakoram collision belt along the Karakoram Highway transect, north Pakistan. *Tectonics* **1999**, *18*, 929–949. [CrossRef]

89. Guzzetti, F.; Mondini, A.C.; Cardinale, M.; Fiorucci, F.; Santangelo, M.; Chang, K.-T. Landslide inventory maps: New tools for an old problem. *Earth-Sci. Rev.* **2012**, *112*, 42–66. [CrossRef]

90. Shafique, M.; van der Meijde, M.; Khan, M.A. A review of the 2005 Kashmir earthquake-induced landslides; from a remote sensing perspective. *J. Asian Earth Sci.* **2016**, *118*, 68–80. [CrossRef]

91. Ilia, I.; Tsangaratos, P. Applying weight of evidence method and sensitivity analysis to produce a landslide susceptibility map. *Landslides* **2016**, *13*, 379–397. [CrossRef]

92. Roy, J.; Saha, S.; Arabameri, A.; Blaschke, T.; Bui, D.T. A novel ensemble approach for landslide susceptibility mapping (LSM) in Darjeeling and Kalimpong districts, West Bengal, India. *Remote Sens.* **2015**, *7*, 292. [CrossRef]

93. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

94. Hameed, N.; Shah, H.R.; Basharat, M. Landslide Hazard Zonation in Pakistan Using the Logistic Regression Model. *J. Mt. Sci.* **2017**, *14*, 1816–1829. [CrossRef]

95. Fayaz, A.; Latif, M.; Khan, K. Landslide Evaluation and Stabilization Between Gilgit and Thakot along the Karakoram Highway; Geological Survey of Pakistan: Islamabad, Pakistan, 1985.

96. Khan, K.; Fayaz, A.; Latif, M.; Hussain, M.; Latif, M. Landslides Problems and Their Mitigation along the Karakoram Highway; Geological Survey of Pakistan: Islamabad, Pakistan, 2003.

97. Hewitt, K. Catastrophic landslides and their effects on the Upper Indus streams, Karakoram Himalaya, northern Pakistan. *Geomorphology* **1998**, *26*, 47–80. [CrossRef]

98. Nayak, S.C.; Bhattacharyya, B.; Singh, A.; Jana, S.K. Landslide Susceptibility Mapping Using Machine Learning—A Case Study Along Karakoram Highway transect, north Pakistan. *Nat. Hazards Res.* **2018**, *19*, 05018007. [CrossRef]

99. Fujinawa, N.; Tani, H.; Narita, M.; Nakamura, T.; Torii, K.; Kato, Y. Landslide susceptibility mapping of mountainous terrain: Case study from the Northwest Himalayas, Pakistan. *J. Indian Soc. Remote Sens.* **2022**, *50*, 239. [CrossRef]

100. Ahmed, M.F.; Rogers, J.D.; Ismail, E.H. A regional level preliminary landslide susceptibility study of the upper Indus river basin. *Eur. J. Remote Sens.* **2014**, *47*, 343–373. [CrossRef]

101. Kim, J.-C.; Lee, S.; Jung, H.-S.; Lee, S. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. *Geocarto Int.* **2018**, *33*, 1000–1015. [CrossRef]

102. Roy, J.; Saha, S.; Arabameri, A.; Blaschke, T.; Bui, D.T. A novel ensemble approach for landslide susceptibility mapping (LSM) in Darjeeling and Kalimpong districts, West Bengal, India. *Remote Sens.* **2015**, *7*, 292. [CrossRef]

103. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

104. Hameed, N.; Shah, H.R.; Basharat, M. Landslide Hazard Zonation in Pakistan Using the Logistic Regression Model. *J. Mt. Sci.* **2017**, *14*, 1816–1829. [CrossRef]
112. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd Acvm Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

113. Friedman, J.H. Stochastic gradient boosting. Comput. Stat. Data Anal. 2002, 38, 367–378. [CrossRef]

114. Friedman, J.H. Greedy function approximation: A gradient boosting machine. Ann. Stat. 2001, 29, 1189–1232. [CrossRef]

115. Wang, C.; Deng, C.; Wang, S. Imbalance-XGBoost: Leveraging weighted and focal losses for binary label-imbalanced classification with XGBoost. Pattern Recognit. Lett. 2020, 136, 190–197. [CrossRef]

116. Marjanovic, M.; Bajat, B.; Kovacevic, M. Landslide susceptibility assessment with machine learning algorithms. In Proceedings of the 2009 International Conference on Intelligent Networking and Collaborative Systems, Barcelona, Spain, 4–6 November 2009; pp. 273–278.

117. Miner, A.; Vamplew, P.; Windle, D.; Flentje, P.; Warner, P. A Comparative Study of Various Data Mining Techniques as Applied to the Modeling of Landslide Susceptibility on the Bellarine Peninsula, Victoria, Australia. In Proceedings of the 11th Iaeg Congress of the International Association of Engineering Geology and the Environment, Auckland, New Zealand, 5–10 September 2010.

118. Bröcker, J.; Smith, L.A. Increasing the reliability of reliability diagrams. Weather Forecast. 2007, 22, 651–661. [CrossRef]

119. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 2011, 12, 2825–2830.

120. Omohundro, S.M. Five Balltree Construction Algorithms; International Computer Science Institute Berkeley: Berkeley, CA, USA, 1989.

121. Chen, J.S.; Huang, H.Y.; Hsu, C.Y. A kNN based position prediction method for SNS places. In Proceedings of the Asian Conference on Intelligent Information and Database Systems, Cham, Germany, 23 March 2020; pp. 266–273.

122. Soria, D.; Garibaldi, J.M.; Ambrogi, F.; Biganzoli, E.M.; Ellis, I.O. A ‘non-parametric’ version of the naive Bayes classifier. Knowl.-Based Syst. 2011, 24, 775–784. [CrossRef]

123. Domingos, P.; Pazzani, M. Beyond independence: Conditions for the optimality of the simple Bayesian classifier. In Proceedings of the Thirteenth International Conference on Machine Learning, Miami, FL, USA, 4–7 December 2013; pp. 105–112.

124. Soria, D.; Garibaldi, J.M.; Biganzoli, E.; Ellis, I.O. A comparison of three different methods for classification of breast cancer data. In Proceedings of the 2008 Seventh International Conference on Machine Learning and Applications, San Diego, CA, USA, 11–13 November 2008; pp. 619–624.

125. Wu, X.; Kumar, V.; Ross Quinlan, J.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Yu, P.S. Top 10 algorithms in data mining. Knowl. Inf. Syst. 2008, 14, 1–37. [CrossRef]

126. Chen, Y.-R.; Chen, J.-W.; Hsieh, S.-C.; Ni, P.-N. The application of remote sensing technology to the interpretation of land use for rainfall-induced landslides based on genetic algorithms and artificial neural networks. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2009, 2, 87–95. [CrossRef]

127. Singh, H.; Pandey, A. Land deformation monitoring using optical remote sensing and PS-InSAR technique nearby Gangotri glacier in higher Himalayas. Modeling Earth Syst. Environ. 2021, 7, 221–233. [CrossRef]

128. Ciampalini, A.; Rasolini, F.; Lagomarsino, D.; Catani, F.; Casagli, N. Landslide susceptibility map refinement using PSI-PSAR data. Remote Sens. Environ. 2016, 184, 302–315. [CrossRef]

129. Zhou, C.; Cao, Y.; Yin, K.; Wang, Y.; Shi, X.; Catani, F.; Ahmed, B. Landslide characterization using sentinel-1 images and InSAR technique: The muyuabao landslide in the three gorges reservoir area, China. Remote Sens. 2020, 12, 3385. [CrossRef]

130. Crosetto, M.; Devanthéry, N.; Cuevas-González, M.; Monserrat, O.; Crippa, B. Exploitation of the full potential of PSI data for subsidence monitoring. Proc. Int. Assoc. Hydrol. Sci. 2015, 372, 311–314. [CrossRef]

131. Yagüe-Martínez, N.; Prats-Iraola, P.; Gonzalez, F.R.; Bric, R.; Shau, R.; Geudtner, D.; Eineder, M.; Bämler, R. Interferometric processing of Sentinel-1 TOPS data. IEEE Trans. Geosci. Remote Sens. 2016, 54, 2220–2234. [CrossRef]

132. Yastika, P.; Shimizu, A.; Abidin, H. Monitoring of long-term land subsidence from 2003 to 2017 in coastal area of Semarang, Indonesia by SBAS DInSAR analyses using Envisat-ASAR, ALOS-PAL SAR, and Sentinel-1A SAR data. Adv. Space Res. 2019, 63, 1719–1736. [CrossRef]

133. Xu, C.; Xu, X.; Dai, F.; Xiao, J.; Tan, X.; Yuan, R. Landslide hazard mapping using GIS and weight of evidence model in Qingshui river watershed of 2008 Wenchuan earthquake struck region. J. Earth Sci. 2012, 23, 97–120. [CrossRef]

134. Malek, Ž.; Zumpano, V.; Schröter, D.; Glade, T.; Balteanu, D.; Micu, M. Scenarios of land cover change and landslide susceptibility: An example from the buzau subcarpathians, romania. In Engineering Geology for Society and Territory; Springer: Berlin/Heidelberg, Germany, 2015; Volume 5, pp. 753–745.

135. Arabameri, A.; Pradhan, B.; Pourghasemi, H.R.; Rezaei, K.; Kerle, N. Spatial modelling of gully erosion using GIS and R programming: A comparison among three data mining algorithms. Appl. Sci. 2018, 8, 1369. [CrossRef]

136. Rahim, I.; Ali, S.M.; Aslam, M. GIS Based landslide susceptibility mapping with application of analytical hierarchy process in District Ghizer, Gilgit Baltistan Pakistan. J. Geosci. Environ. Prot. 2018, 6, 34–49. [CrossRef]

137. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. Landslides 2014, 11, 425–439. [CrossRef]

138. Song, Y.; Niu, R.; Xu, S.; Ye, R.; Peng, L.; Guo, T.; Li, S.; Chen, T. Landslide susceptibility mapping based on weighted gradient boosting decision tree in Wanzhou section of the Three Gorges Reservoir Area (China). ISPRS Int. J. Geo-Inf. 2019, 8, 4. [CrossRef]
139. Rehman, A.; Song, J.; Haq, F.; Mahmood, S.; Ahamad, M.I.; Basharat, M.; Sajid, M.; Mehmood, M.S. Multi-Hazard Susceptibility Assessment Using the Analytical Hierarchy Process and Frequency Ratio Techniques in the Northwest Himalayas, Pakistan. Remote Sens. 2022, 14, 554. [CrossRef]

140. Aslan, G.; Foumelis, M.; Raucoules, D.; De Michele, M.; Bernardie, S.; Cakir, Z. Landslide mapping and monitoring using persistent scatterer interferometry (PSI) technique in the French Alps. Remote Sens. 2020, 12, 1305. [CrossRef]

141. Chen, W.; Xie, X.; Wang, J.; Pradhan, B.; Hong, H.; Bui, D.T.; Duan, Z.; Ma, J. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena 2017, 151, 147–160. [CrossRef]

142. Zhang, K.; Wu, X.; Niu, R.; Yang, K.; Zhao, L. The assessment of landslide susceptibility mapping using random forest and decision tree methods in the Three Gorges Reservoir area, China. Environ. Earth Sci. 2017, 76, 405. [CrossRef]

143. Immitzer, M.; Atzberger, C.; Koukal, T. Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. Remote Sens. 2012, 4, 2661–2693. [CrossRef]

144. Yu, K.; Yao, X.; Qiu, Q.; Liu, J. Landslide spatial prediction based on random forest model. Trans. CSAM 2016, 47, 338–345.

145. Zhang, Y.; Ge, T.; Tian, W.; Liou, Y.-A. Debris flow susceptibility mapping using machine-learning techniques in Shigatse area, China. Remote Sens. 2019, 11, 2801. [CrossRef]

146. Piacentini, D.; Devoto, S.; Mantovani, M.; Pasuto, A.; Prampolini, M.; Soldati, M. Landslide susceptibility modeling assisted by Persistent Scatterers Interferometry (PSI): An example from the northwestern coast of Malta. Nat. Hazards 2015, 78, 681–697. [CrossRef]

147. Hakim, W.L.; Achmad, A.R.; Lee, C.-W. Land subsidence susceptibility mapping in jakarta using functional and meta-ensemble machine learning algorithm based on time-series InSAR data. Remote Sens. 2020, 12, 3627. [CrossRef]

148. Xie, Z.; Chen, G.; Meng, X.; Zhang, Y.; Qiao, L.; Tan, L. A comparative study of landslide susceptibility mapping using weight of evidence, logistic regression and support vector machine and evaluated by SBAS-InSAR monitoring: Zhouqu to Wudu segment in Bailong River Basin, China. Environ. Earth Sci. 2017, 76, 313. [CrossRef]

149. Chalkias, C.; Ferentinou, M.; Polykretis, C. GIS-based landslide susceptibility mapping on the Peloponnese Peninsula, Greece. Geosciences 2014, 4, 176–190. [CrossRef]

150. Hussain, S.; Hongxing, S.; Ali, M.; Sajjad, M.M.; Ali, M.; Afzal, Z.; Ali, S. Optimized landslide susceptibility mapping and modelling using PS-InSAR technique: A case study of Chitral valley, Northern Pakistan. Geocarto Int. 2021, 36, 1–22. [CrossRef]