Review

Landslide Susceptibility Mapping Using Machine Learning: A Literature Survey

Moziihrii Ado [1], Khwairakpam Amitab [1,*], Arnab Kumar Maji [1], Elżbieta Jasińska [2], Radomir Gono [3], Zbigniew Leonowicz [4] and Michał Jasiński [4]

1   Department of Information Technology, North-Eastern Hill University, Shillong 793022, India; moziihrii@nehu.ac.in (M.A.); akmaji@nehu.ac.in (A.K.M.)
2   Department of Operations Research and Business Intelligence, Wrocław University of Science and Technology, 50-370 Wrocław, Poland; elzbieta.jasinska@pwr.edu.pl
3   Department of Electrical Power Engineering, Faculty of Electrical Engineering and Computer Science, VSB—Technical University of Ostrava, 708 00 Ostrava, Czech Republic; radomir.gono@vsb.cz
4   Faculty of Electrical Engineering, Wrocław University of Science and Technology, 50-370 Wrocław, Poland; zbigniew.leonowicz@pwr.edu.pl (Z.L.); michal.jasinski@pwr.edu.pl (M.J.)
*   Correspondence: kamitab@nehu.ac.in

Abstract: Landslide is a devastating natural disaster, causing loss of life and property. It is likely to occur more frequently due to increasing urbanization, deforestation, and climate change. Landslide susceptibility mapping is vital to safeguard life and property. This article surveys machine learning (ML) models used for landslide susceptibility mapping to understand the current trend by analyzing published articles based on the ML models, landslide causative factors (LCFs), study location, datasets, evaluation methods, and model performance. Existing literature considered in this comprehensive survey is systematically selected using the ROSES protocol. The trend indicates a growing interest in the field. The choice of LCFs depends on data availability and case study location; China is the most studied location, and area under the receiver operating characteristic curve (AUC) is considered the best evaluation metric. Many ML models have achieved an AUC value > 0.90, indicating high reliability of the susceptibility map generated. This paper also discusses the recently developed hybrid, ensemble, and deep learning (DL) models in landslide susceptibility mapping. Generally, hybrid, ensemble, and DL models outperform conventional ML models. Based on the survey, a few recommendations and future works which may help the new researchers in the field are also presented.

Keywords: landslide; susceptibility mapping; dataset; causative factors; hybrid; ensemble; deep learning; machine learning; literature survey

1. Introduction

Landslide is the movement of rock, soil, and organic matter down the slope, influenced by gravitational force. Landslide is a devastating natural disaster, causing mass damages to vegetation and properties, as well as human fatalities. It is expected to occur more frequently due to three main reasons: increased urbanization, deforestation, and precipitation intensity due to global climate change [1]. That is why we are concerned with landslide. Prediction and risk assessment of landslide is key to reducing loss of lives and property damage. Landslide susceptibility map (LSM) indicates the landslide-prone areas, which can be used by policy-makers, scientists, engineers, and the general public to avoid catastrophic landslide. In recent years, many advanced machine learning (ML) techniques have been developed to model the complex relationships between the landslide and the causing factors. The success of ML in landslide susceptibility mapping is further enhanced by easily accessible satellite images, remote sensing data, historical records of landslide, and geographical information systems (GIS). This article presents a comprehensive survey of

existing ML algorithms for generating LSM, including the recently developed conventional, hybrid, ensemble, and deep learning techniques. This survey will help new researchers in choosing the ML algorithms, causative factors, performance evaluation methods, and landslide inventories, among other information.

Many literature surveys and critical review articles on statistical and knowledge-based methods for landslide susceptibility mapping are available [2–6]. However, only a few literature surveys or reviews are available for ML-based landslide susceptibility mapping. We could identify four [7–10] existing review articles on popular ML methods used for the generation of LSM, which are listed on the Web of Science (WoS) [11] as a review article. A summary and the limitation of the studies are presented in Table 1.

Table 1. Existing literature review on landslide susceptibility mapping using machine learning.

Year	Author	Objective	Summary	Limitation
2018	Yu Huang et al. [8]	To review SVM for landslide susceptibility mapping and compare it with four other ML techniques.	The basic theory of SVM is presented, followed by a discussion on the methodology involved in landslide susceptibility mapping using SVM. The SVM and other four techniques (AHP, LGR, ANN, and RF) are theoretically compared.	The work does not include the latest development on SVM. Only the basic theory of SVM is discussed. Many articles on landslide susceptibility mapping using SVM are not included. This review is limited to SVM and four other ML techniques.
2020	Merghadi et al. [7]	To present the popular ML techniques available for landslide susceptibility mapping.	Presented the basic architecture of popular ML techniques and highlighted the advantages and disadvantages of each model. Analyzed the performance of the ML techniques by considering a case study location in Algeria.	This study does not include many recently developed hybrid, ensemble, and deep learning techniques.
2021	Naemitabar et al. [9]	To do a comparative study of popular ML methods used for generating LSM.	The study's main focus was on the prioritization of effective LCFs to get better performance accuracy. Four ML models were reviewed: SVM, BRT, LMT, and RF.	The study discussed only four ML models.
2021	Zhang et al. [10]	To carry out a comparative study of four traditional ML models integrated with bagging strategy to improve the performance.	Four conventional ML models: BFTree, FT, SVM, and CART as a base model, were integrated with bagging-based ensemble method to improve the performance of the base models. The result shows that the bagging-based ensemble method outperformed the traditional ML models. Significant improvement was observed for CART.	The study was limited to bagging-based ensemble models of four conventional ML models.

Huang et al. [8] undertook a theoretical comparative study of support vector machine (SVM) and a few other ML models. The models include analytic hierarchy process (AHP), logistic regression (LGR), artificial neural network (ANN), and random forests (RF). The authors have discussed the similarities and differences of the models. The study suggests combining SVM with other methods and using high-quality landslide data to get an accurate susceptibility map. Merghadi et al. [7] compared linear regression (LR), ANN, SVM, decision tree (DT), RF, extremely randomized trees (ERT), Naive Bayes (NB), linear discriminant analysis (LDA), quadratic discriminant analysis, k-nearest neighbors (KNN), gradient boosting (GB), and neuro-fuzzy for landslide susceptibility mapping. The basic architectures of the models are presented, and an extensive comparative analysis between different ML techniques was performed by considering a case study location in Algeria. The authors discussed the algorithm's advantages, limitations, and accuracy. The study contribution includes; comprehensive reviews focusing exclusively on the use of ML in landslide susceptibility mapping to present the complexities, comparisons, challenges, and opportunities for future works. Naemitabar et al. [9] have also carried out a comparative

study of popular ML methods used in the generation of LSM. The study's primary focus is prioritizing effective landslide causative factor (LCF) to improve performance accuracy. Four ML models were reviewed: SVM, boosted regression trees (BRT) model, logistic model tree (LMT), and the RF. The study result showed that SVM with an area under the receiver operating characteristic curve (AUC) value of 0.86 and RF with an AUC value of 0.89 had a better performance than others. Multiple LCFs were used for the study, out of which the most effective ones identified were lithology, slope, slope aspect, distance to fault, and land use/land cover (LULC). The author also suggests that prioritization of effective LCFs for training the ML models produced better accuracy. Zhang et al. [10] used four conventional ML models namely, best-first decision tree (BFTree), functional tree (FT), SVM, and classification regression tree (CART). The bagging-based ensemble method was used to improve the performance of the ML models. The comparative study was carried out in Jiange County, Sichuan Province, China, to generate LSM for resource planning and landslide management. The result shows that the bagging-based ensemble method outperformed the conventional ML models. Significant improvement was observed for CART with an AUC value of 0.766 without bagging and an AUC value of 0.874 for the bagging-CART model.

There are few review papers or comparative studies [7–10] on ML-based landslide susceptibility mapping. Technology is constantly evolving rapidly, and better ML models are being introduced in landslide studies each passing day. Different studies [12–14] also observed ML methods such as the hybrid, ensemble, and deep learning-based models have good performance accuracy for generating LSM. Furthermore, a comprehensive literature survey covering all the elements in ML-based LSM generation is nonexistent. The study's primary objective is to conduct an extensive literature survey on the ML models for landslide susceptibility mapping. We will present the conventional, hybrid, ensemble, and deep learning models and discuss the advantages and limitations of these models. This comprehensive survey will give new researchers insight into the elements involved in landslide susceptibility mapping using ML, the current trend, and future scopes.

The rest of the paper is organized as follows: Section 2 discusses the survey methodology detailing the selection of articles considered in this survey. Section 3 presents the current trends showing the statistics of articles published per year, ML models used, top journals, and most studied case study locations. Section 4 discusses the essential elements of LSM using ML, including LCFs, landslide datasets and inventory, ML models, and evaluation methods. Section 5 surveys the ML models used in LSM. The ML models are categorized as conventional, hybrid, ensemble, and deep learning. Section 6 summarizes the current state of the research. Section 7 concludes this survey by suggesting a few recommendations and future research directions based on the knowledge gathered in this survey.

2. Survey Methodology

We have used reporting standards for systematic evidence synthesis in environmental research (ROSES) [15] for systematic literature survey. ROSES aims to improve the standard of systematic review in environmental research. It is an extension of PRISMA [16] which is widely used as a reporting standard in health care. ROSES ensure a reproducible and transparent literature review process. To find relevant articles for this survey, we have used different combinations of keywords in the WoS [11] platform to search for literature related to our study. Keyword combination includes; "landslide, susceptibility mapping, machine learning", "landslide, susceptibility mapping, deep learning", "landslide, susceptibility mapping, SVM, support vector machine", "landslide, susceptibility mapping, random forest, RF", "landslide, susceptibility mapping, hybrid, machine learning", "landslide, susceptibility mapping, ensemble, machine learning", "landslide, susceptibility mapping, ANN, artificial neural network", and "landslide, susceptibility mapping, DT, decision tree". We tried to include all the popular ML models used for landslide susceptibility mapping in the search to the best of our knowledge. Search containing landslide susceptibility mapping

and ML keyword returned 424 papers. The combination of landslide susceptibility mapping with SVM (231), RF (209), deep learning (67), hybrid (97), ensemble (153), ANN (148), and DT (27) as keywords produced the search result values as indicated. The search yielded 1356 articles, and the articles were filtered using the ROSES reporting standard. The steps involved in screening the articles are presented in a flow diagram as shown in the Figure 1.
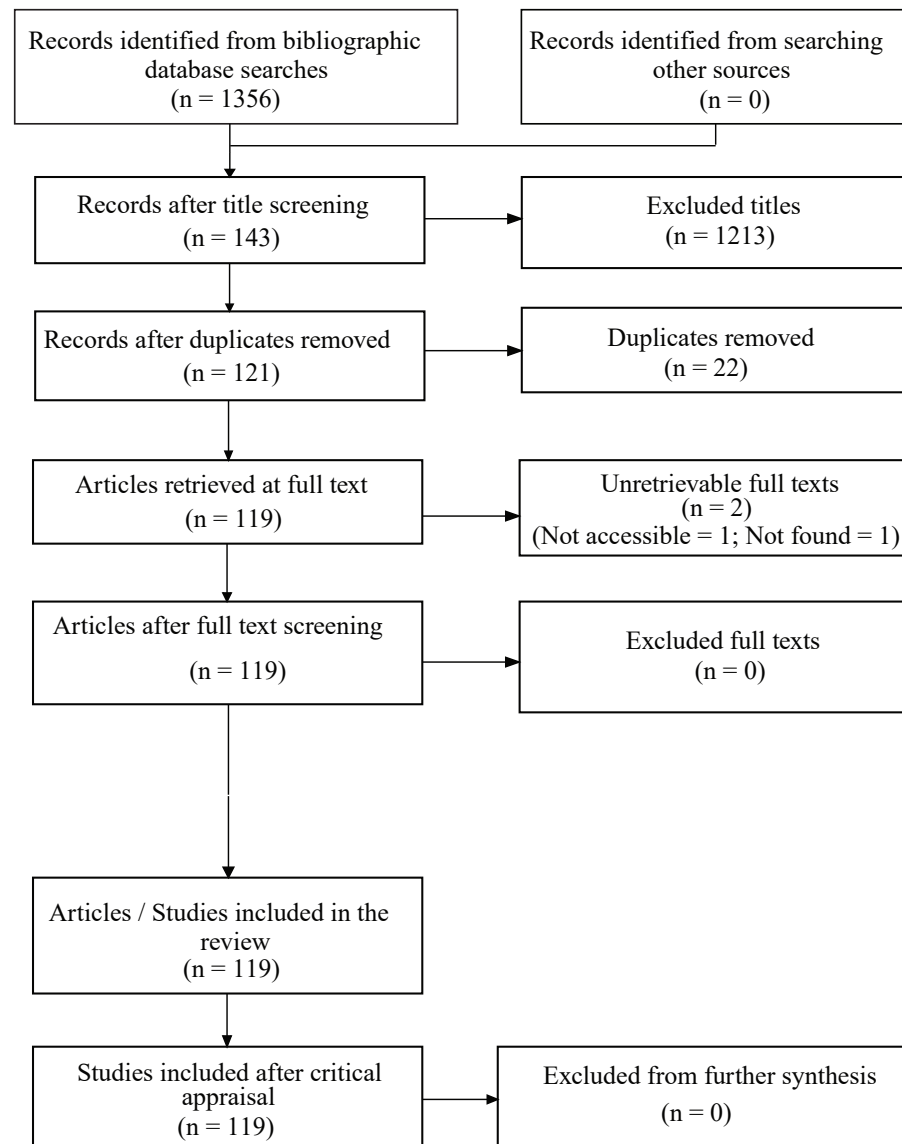
```
┌─────────────────────────────┐     ┌─────────────────────────────┐
│ Records identified from      │     │ Records identified from      │
│ bibliographic                │     │ searching                    │
│ database searches            │     │ other sources                │
│ (n = 1356)                   │     │ (n = 0)                      │
└─────────────────────────────┘     └─────────────────────────────┘

┌─────────────────────────────┐     ┌─────────────────────────────┐
│ Records after title screening│ →   │ Excluded titles              │
│ (n = 143)                    │     │ (n = 1213)                   │
└─────────────────────────────┘     └─────────────────────────────┘

┌─────────────────────────────┐     ┌─────────────────────────────┐
│ Records after duplicates     │ →   │ Duplicates removed           │
│ removed                      │     │ (n = 22)                     │
│ (n = 121)                    │     │                              │
└─────────────────────────────┘     └─────────────────────────────┘

┌─────────────────────────────┐     ┌─────────────────────────────────┐
│ Articles retrieved at full   │ →   │ Unretrievable full texts         │
│ text                         │     │ (n = 2)                          │
│ (n = 119)                    │     │ (Not accessible = 1; Not found=1)│
└─────────────────────────────┘     └─────────────────────────────────┘

┌─────────────────────────────┐     ┌─────────────────────────────┐
│ Articles after full text     │ →   │ Excluded full texts          │
│ screening                    │     │ (n = 0)                      │
│ (n = 119)                    │     │                              │
└─────────────────────────────┘     └─────────────────────────────┘

┌─────────────────────────────┐
│ Articles / Studies included  │
│ in the review                │
│ (n = 119)                    │
└─────────────────────────────┘

┌─────────────────────────────┐     ┌─────────────────────────────┐
│ Studies included after       │ →   │ Excluded from further        │
│ critical appraisal           │     │ synthesis                    │
│ (n = 119)                    │     │ (n = 0)                      │
└─────────────────────────────┘     └─────────────────────────────┘
```

Figure 1. Flow diagram for selection of articles for the survey based on ROSES protocol.

We use specific criteria to filter out literature not related to our study. These criteria ensure reproducible results for verification and validation of the survey. In the first filtration process, we analyzed the article's title to check if it contained the keywords specified in our search keywords. We were able to obtain 143 papers after the title screening process. In the second step, duplicate titles were removed, taking the number of articles down to 121. The next step involved retrieval of full text for 121 filtered articles. We could not retrieve one article, and one article was not found, taking the number of articles for the study to 119. The following two steps involved screening articles based on full-text screening and critical appraisal. All the 119 articles were considered valuable for landslide susceptibility mapping using ML. Hence none of the articles were excluded in these steps. After the filtration process, a total of 119 articles were selected for the study, as shown in the Figure 1.

3. Current Trend

This section discusses the current trend in ML-based landslide susceptibility mapping. Trends are good indicators of understanding the interest and progress in any field. In Figure 2a, we see there is a considerable jump in interest of ML-based landslide susceptibility mapping.
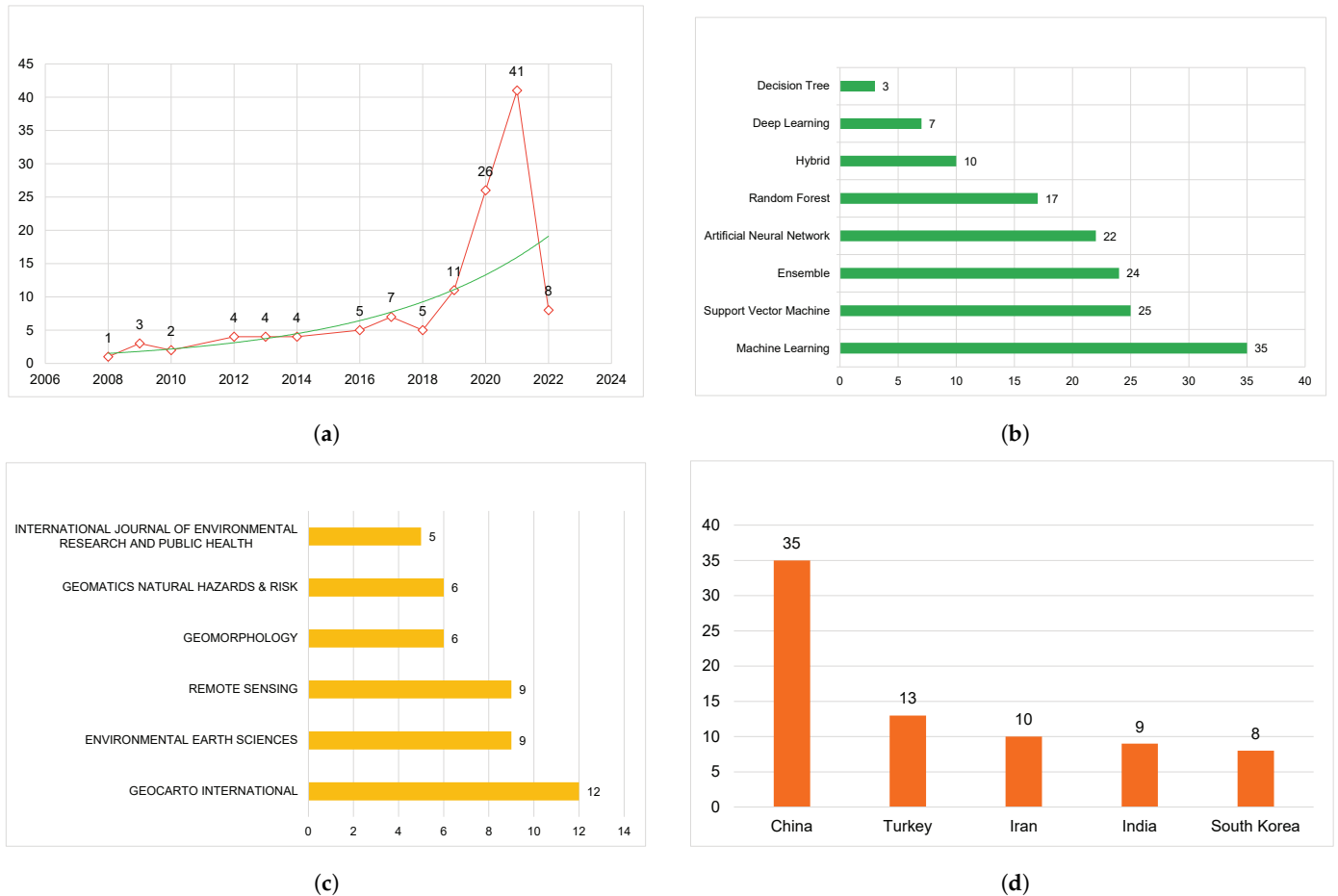
(a)

(b)

(c)

(d)

Figure 2. (**a**) The number of article/s published per year on machine learning based landslide susceptibility mapping. (**b**) The number of machine learning model names detected in article titles. (**c**) Top six journals on machine learning based landslide susceptibility mapping. (**d**) Top five countries as case study location.

The number of articles published per year increased from below ten to more than ten in 2019 and a substantial increase in 2020 (26) and 2021(41). The statistics show growing interest in LSM using ML. Figure 2b shows the trend of articles for landslide susceptibility mapping based on ML methods. The figure shows that articles with ML in the article title were the highest with 35. The remaining ML methods based on the article titles are SVM (25), ensemble (24), ANN (22), RF (17), hybrid (10), deep learning (seven), and DT (three). We found much diversity in the choice of journals and identified 44 different journals. The top six journals in terms of the number of papers published are shown in the Figure 2c. The journals *Geocarto International* (twelve), *Environmental Earth Science* (nine), *Remote Sensing* (nine), *Geomorphology* (six), *Geomatics Natural Hazards & Risk* (six), and *International Journal of Environmental Research and Public Health* (five) are the most popular choice of the authors. The trend on case study location shown in the Figure 2d indicates China (35) as the most popular case study location. Other countries, including Turkey (13), Iran (10), India (9), and South Korea (8), were the other popular choice as case study locations.

Table 2 presents the list of the top ten most-cited articles with the author name, the title of the article, and the citation count. We observed that the number of citation counts is different on Google scholar and WoS web portals. The google scholar citation count is always higher than WoS. The citation count in Table 2 is based on WoS, as we are using the WoS web portal as the primary source of information for our study. The citation count is accurate to the best of our knowledge when writing this article.

Table 2. Most cited articles.

Author	No. of Citation	Article Title
Pradhan et al. [17]	650	A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS
Yilmaz et al. [18]	504	Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey)
Yilmaz et al. [19]	346	Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine
Youssef et al. [20]	340	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
Yao et al. [21]	327	Landslide susceptibility mapping based on Support Vector Machine: A case study on natural slopes of Hong Kong, China
Zare et al.[22]	222	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
Park et al. [23]	217	Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea
Pourghasemi et al. [24]	188	Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran
Kalantar et al. [25]	182	Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between SVM, LGR and ANN
Huang et al. [8]	154	Review on landslide susceptibility mapping using support vector machines

4. Elements of Landslide Susceptibility Mapping Using Machine Learning

The study of landslide using ML involves many essential elements. They include LCFs, landslide datasets and inventory, ML models, and evaluation techniques. Each element influences the outcome of the landslide study differently.

4.1. Landslide Causative Factors

The LCF or landslide conditioning factors are the primary factors responsible for landslide. The objective of landslide susceptibility mapping using ML is to enable the ML model to find relationships between the occurrence of landslide and LCFs. Once the relationship is established, the ML models will be able to generate LSM using the LCFs. Researchers have identified nearly 60 LCFs [26]. Determining the exact number of LCFs to be considered in LSM is one of the most critical and challenging tasks [27]. Up till now, there is no universally agreed upon condition-specific determination of LCF. Studies like [13,28] use ML models to select LCFs for better accuracy. Popular LCFs can be broadly categorized as follow:

- Topography: Slope, aspect, elevation, plan curvature, profile curvature, and sediment transport index;
- Hydrology: Rainfall, solar radiation, stream power index, topographic wetness index (TWI), distance to rivers, and density of the river;
- Geological: Lithology, distance to faults, and density of fault;
- Land use/cover: LULC and normalized difference vegetation index (NDVI);
- Man-made: Distance to roads and density.

In the papers we have surveyed, the LCFs, namely slope, elevation, rainfall, distance to rivers, LULC, NDVI, and distance to roads, are commonly used in landslide susceptibility mapping. The LCFs are prepared as a thematic data layer using a GIS platform to study landslide. We shall briefly discuss some of the popular LCFs used in the generation of LSM. As a sample, different LCF layers were generated using Google earth engine (GEE), ArcGIS, and QGIS for Meghalaya, India.

Slope degree: The degree of a slope is considered an essential factor in determining the resilience of a slope [29]. It has a direct relation with the occurrence of landslide [30]. The degree of slope used for landslide susceptibility studies can be generated using the digital elevation model (DEM), and slopes are grouped into slope classes. The ML models are then used to determine the landslide susceptible class. Figure 3 shows the slope degree for the state of Meghalaya, India.
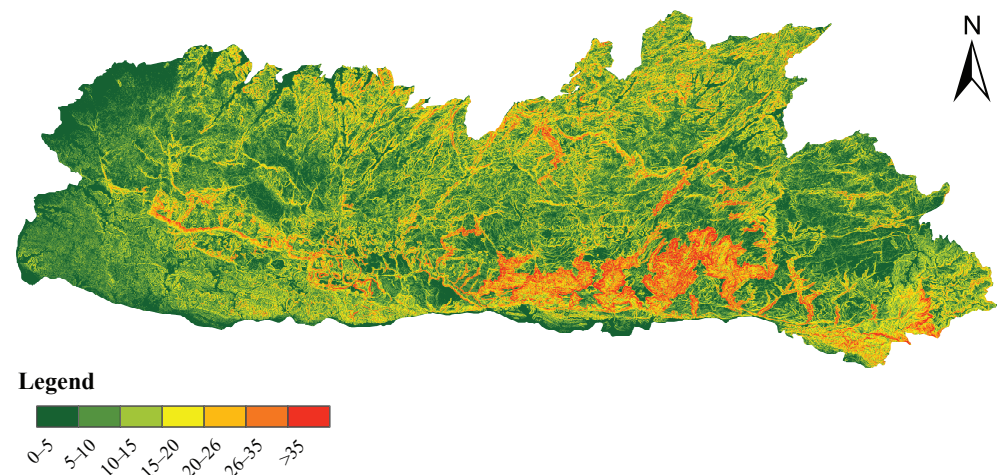
Figure 3. Slope degree for the state of Meghalaya, India.

The slope degree is classified into 7 class, where dark-green region indicates slope with 0–5°, green region 5–10°, light-green 10–15°, yellow region 15–20°, orange region 20–26°, brownish-red region 26–35°, and red region >35° degrees. The slope degree was generated using shuttle radar topography mission (SRTM) DEM data from GEE and ArcGIS.

Aspect: The slope aspect is one of the primary causative factors in many landslide susceptibility mapping studies. The paper [31] grouped aspects of a slope into nine classes. The classes were used to determine the direction of the slope receiving more rainfall and sunlight, to decide the side of the slope that is more susceptible to landslide. Figure 4 shows the aspect for the state of Meghalaya, India.
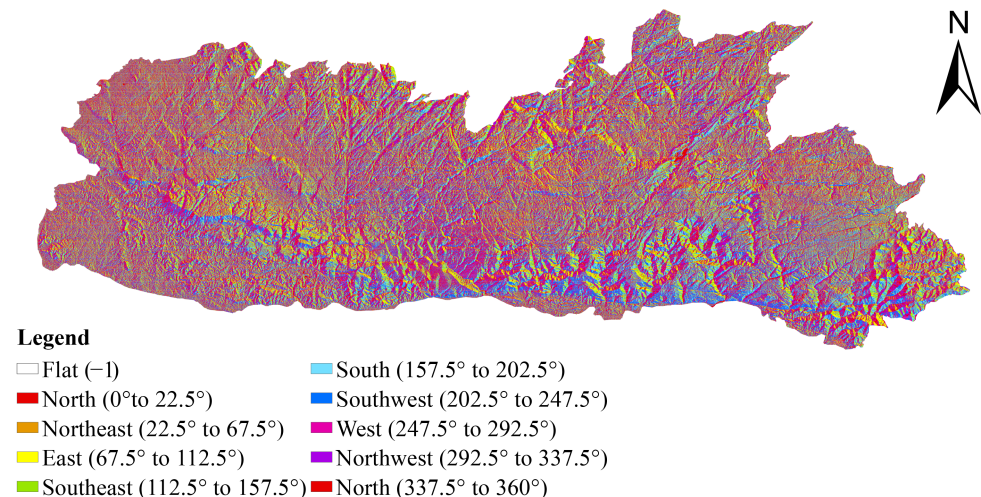

Figure 4. Slope aspect for the state of Meghalaya, India.

The aspect of the state is grouped into ten classes, Flat (−1), North (0° to 22.5°), Northeast (22.5° to 67.5°), East (67.5° to 112.5°), Southeast (112.5° to 157.5°), South (157.5° to 202.5°), Southwest (202.5° to 247.5°), West (247.5° to 292.5°), Northwest (292.5° to 337.5°), and North (337.5° to 360°) indicated by different colors as shown in the figure. The aspect was generated using SRTM DEM from GEE and ArcGIS.

Plan curvature: The plan curvature influences the characteristics of slope erosion or surface runoff. It describes the intersection or divergence of water flow in a downslope [32]. The study [31] uses an automated geoscientific analysis GIS for the generation of plan curvature map used for landslide studies. Multiple studies use the combination of profile and plan curvature as a causative factor. The plan curvature of Meghalaya, India is shown in the Figure 5, generated using SRTM DEM on GEE and QGIS software.

Figure 5. Plan curvature for the state of Meghalaya, India.

The curvature was classified into three classes represented by different colors: concave indicated by black, flat region by green, and convex region by red.

Lithology: The landslide susceptibility value varies based on the lithological units. Grouping lithological attributes in the correct group are crucial in landslide susceptibility mapping studies. The lithology map for susceptibility studies can be mapped and digitized using GIS. Figure 6 shows different lithological properties of Meghalaya, India.
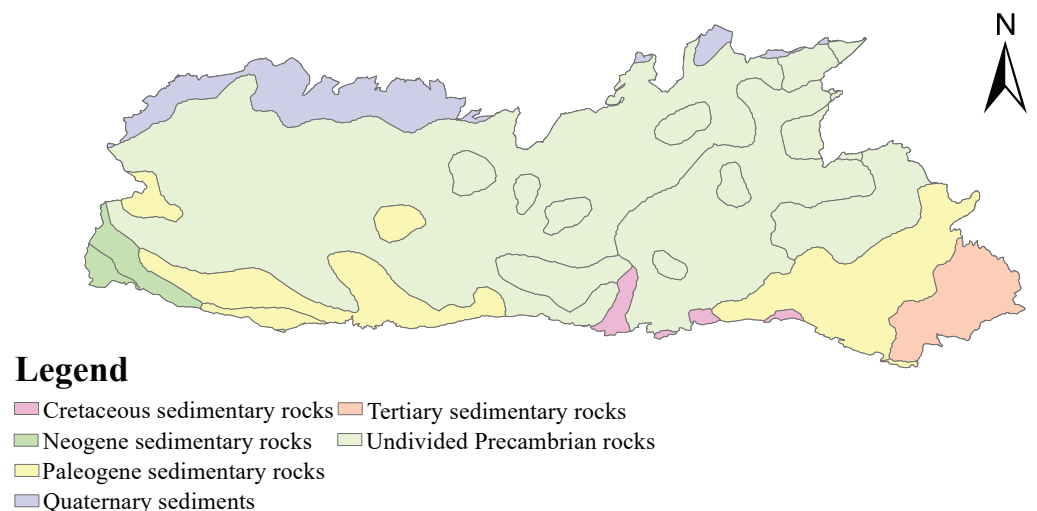

Figure 6. Lithology/Geology for the state of Meghalaya, India.

The lithological/geological data were downloaded from data.gov (accessed on 2 March 2022) [33] made available by the U.S. Geological Survey. The lithology attributes consist of

cretaceous sedimentary rocks, neogene sedimentary rocks, paleogene sedimentary rocks, quaternary sediments, tertiary sedimentary rocks, and undivided precambrian rocks.

Distance from river: A slope's closeness to a river or stream is also a significant factor in slope stability. Water drainage systems can seriously impact the stability by eroding the slope [34]. For landslide susceptibility mapping, the slopes can be classified into buffer areas to determine the drainage system's effect on slope stability. Figure 7 shows three buffer zones for the state of Meghalaya, India.
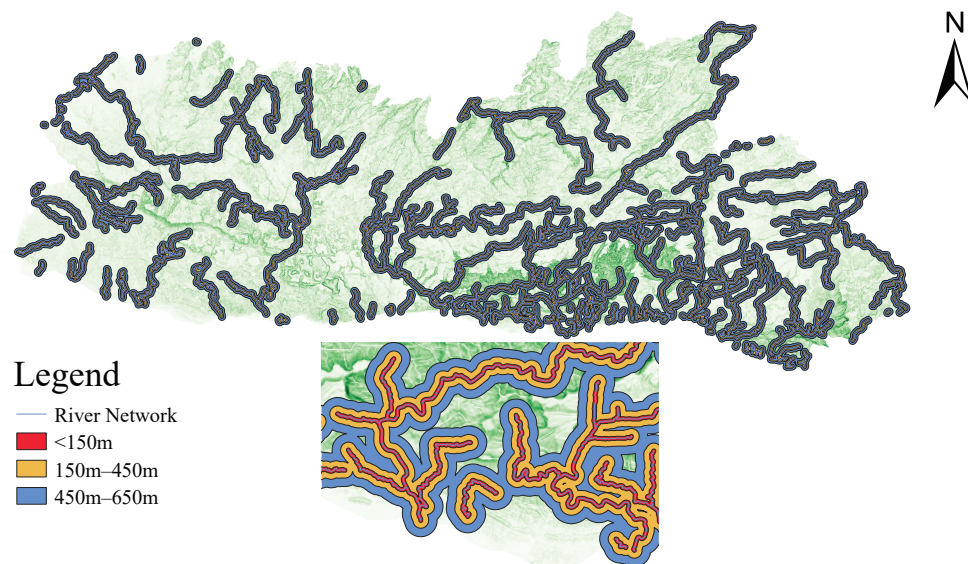
Figure 7. Distance from river for the state of Meghalaya, India.

The three zones consist of areas with 150 m and less, 150–450 m, and 450–650 m distance from a river. The distance from the river map for this article was generated using the QuickOSM tool in QGIS with key-value pair set to 'waterways:river' and region set to 'Meghalaya'. The generated river network was exported as a shapefile and imported to ArcGIS to re-project the shapefile to the correct coordinate. The re-projected shapefile was used to generate the buffer zone in QGIS using the buffer tool.

Topographical wetness index: Among other factors, TWI is also a significant causative factor employed to measure the degree of water accumulation at a site. According to the studies [35,36], TWI is the proneness of water to accumulate at any point divided by the gravitational force moving the water downslope, given by Equation (1):

$$\text{TWI} = ln(a/tan\beta). \tag{1}$$

In the above equation, a is the total up-slope area, and $tan\beta$ is the angle of the slope at that particular point. The $ln(a/tan\beta)$ is the proneness of water to collect at any point (given by a) and the gravitational forces that move the water downslope (given by $tan\beta$). We have used QGIS to generate TWI for the state of Meghalaya, India. Figure 8 shows the wetness index of the location in the range of $4.582 \leq$ for lower wetness to $6.446 \geq$ for higher wetness region.

Land use/Land cover: The development of roads and urbanization have led to landslides. The process involves digging and excavation to construct houses or build roads destabilizing slopes. LULC can be classified into five classes which include water bodies, urban areas, thick vegetation, rock outcrop, and scant vegetation [37]. Figure 9 shows the distribution of vegetation areas represented in green, water bodies in white, and urban areas in red.

The LULC thematic layer was generated on GEE using satellite images from Landsat 8 Level 1 data. The satellite images were filtered using dates ranging from 2020 to 2022

and picking only images with less than 1% cloud cover. Further, CART a supervised ML method was used to classify the three classes in GEE.
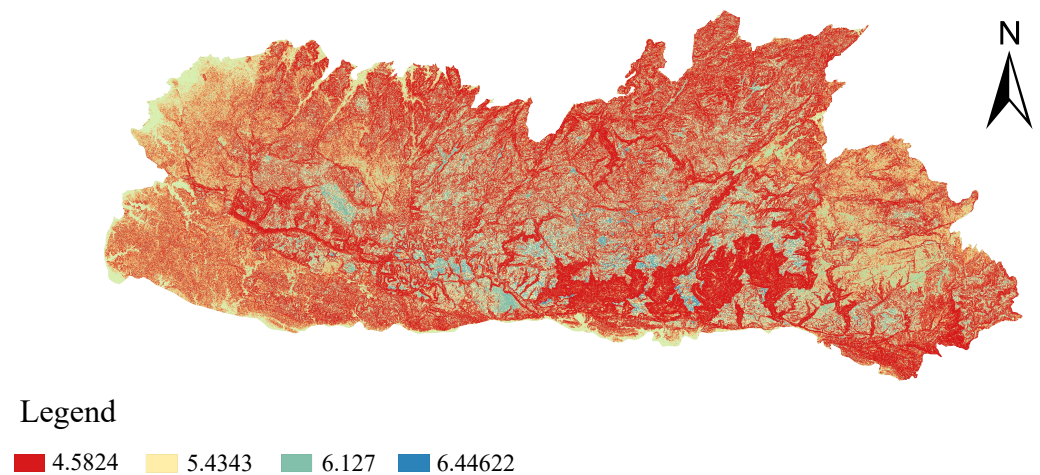

Figure 8. Topographic wetness index for the state of Meghalaya, India.
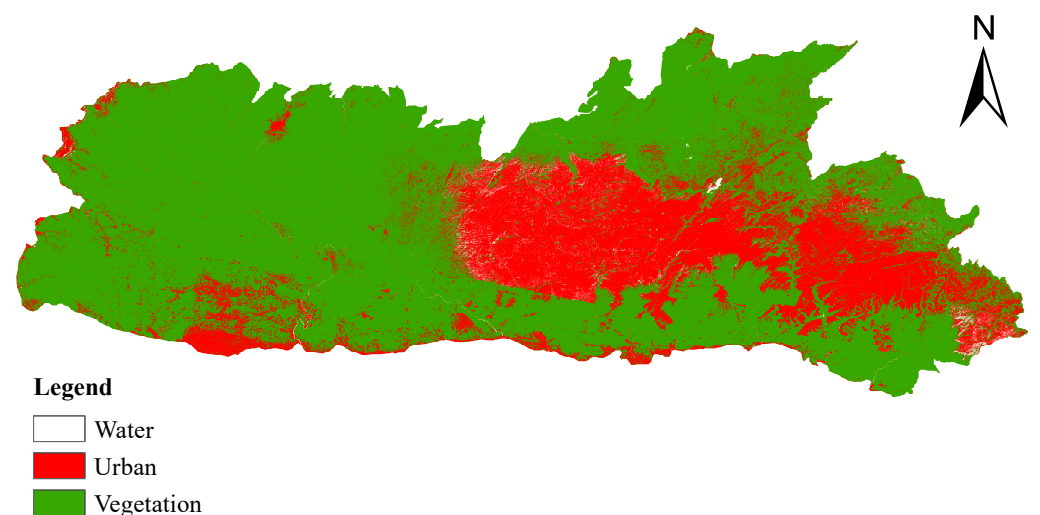

Figure 9. Land use/Land cover for the state of Meghalaya, India.

Generation of accurate LSM requires the requisite LCFs. There are no standard guidelines for selecting LCFs; data availability for the case study location is often the main factor in selecting LCFs. Many authors have considered 20 or more causative factors [38–43]. The importance of causative factors differs from one study location to another. Wang et al. [44] have proposed point of interest (POI) kernel density as an LCF which is an important factor in terms of human engineering and can influence the occurrence of landslide. Identification of the causative factors that directly influence the landslide is crucial. Including irrelevant features or causative factors can degrade the performance of the ML models. Several studies have employed feature selection and optimization methods to enhance the performance of the ML models for generating LSM. It is observed in many studies that this additional step positively impacts the accuracy and performance of ML models. For example, studies like [45–47] have used Chi-square attribute evaluation (CSAE) and multi-collinearity analysis by applying variance inflation factor (VIF) and tolerance to select the LCFs with the most influence on landslide. Researchers also have used the Ridge regression method for LCFs importance analysis [45]. Other popular methods for feature selection and analysis includes relative risk regression analyse [45], fractal analysis [48], resampling scheme analysis and Pearson's correlation analysis [49], correlation-based features selections (CFS) [50], frequency ratio (FR) [51], fuzzy and weights of LCFs using SVM [52],

principal component analysis (PCA) to select independent and significant LCFs [53], information gain method [54], GeoDetector and recursive feature elimination (RFE) method for LCFs optimization to reduce redundancy [51], interactive detector [51], one rule (one-R) [42], correlation attributes evaluation (CAE) where greater calculated average merit (AM) indicates more influence of the LCF [55], sensitivity analysis [56], Spearman's rank correlation coefficient [57], relief-F method [58], Fischer score analysis [47], and gain ratio method [59].

We found many methods employed for feature selection and analysis. It is difficult to pinpoint the best method for feature selection. All the techniques used had a positive influence on the generation of LSM by improving the accuracy and performance of the ML models. We also observed challenges concerning LCFs where some factors had different measurement scales from others, reducing the prediction accuracy of the ML models. For example, the study [51] used terrain factors generated from a DEM scale at 30 m resolution, the geological factors were vectored from a 1:200,000 geological map, and all factors were resampled at 30 m resolution for convenience. Liu et al. [60] employed resampling of the digital terrain model (DTM) with a 5 m × 5 m grid to 30 m × 30 m to make the scale of all factors consistent. Liu et al. [61] suggest exploring the effect of quantity and quality of LCFs on the prediction accuracy of ML models. Ngo et al. [62] state that the unavailability of essential LCFs data such as soil depth, soil texture, and distance from the water table was a limitation of their study, which could affect the accuracy of the LCM generated.

4.2. Datasets and Landslide Inventory

There are multiple ways to gather datasets and prepare landslide inventory data. Historical data, field surveys, aerial photograph interpretation, Google earth image interpretation, and satellite imagery are standard. The data collected are used for training and testing the ML models. Good quality and quantity of data are required to achieve high prediction accuracy. We shall discuss a few data collection methods briefly. As a sample, we have generated an landslide inventory map for the state of Meghalaya, India, shown in Figure 10 using data from the Geological Survey of India.
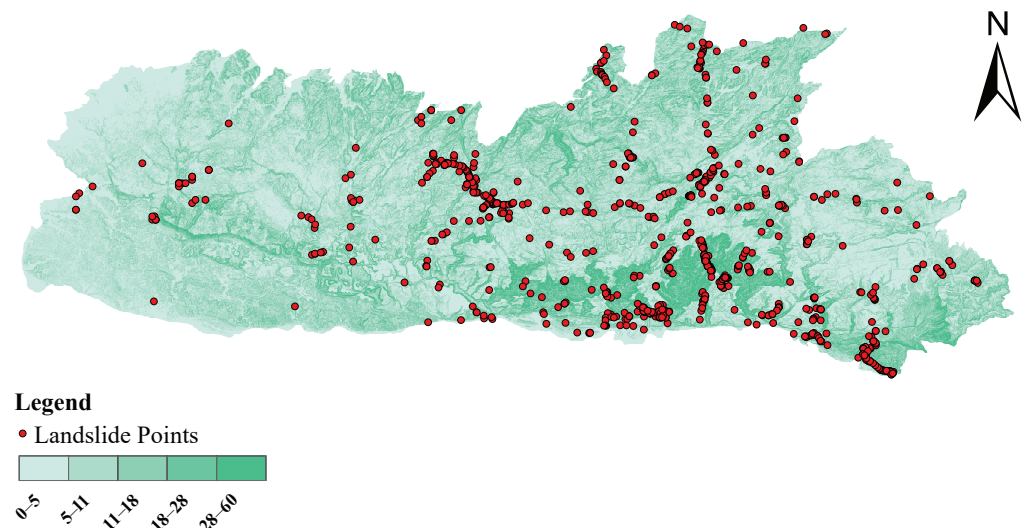

Figure 10. Landslide data points for the state of Meghalaya, India.

The existing landslide data points were downloaded using Job ID: d002505f-b10c-4182-aa05-71e009fff229, from Geological Survey of India, Government of India, Kolkata, India; Bhukosh website [63]. In Figure 10, existing landslide points are represented by the red dots, and the different shades of green represent the degree of slope for the state.

Historical data: Historical inventories have landslide records which consist of information about a location for periods ranging from decades to millennium [64]. In contrast, current inventories center around shorter periods, where data are collected in terms of

hours, days, or weeks. The historical data provides information about the nature of landslide in a particular area and the changes observed over time [65]. So, historical data are helpful in training ML models as they contain many landslide records, and the current inventories can be used for validation.

Field survey: is considered the best method for generating landslide datasets or inventory. It involves a physical examination and detailed study of the affected location to identify and measure every distinguishable slope failure [66]. It is also used to validate landslide inventory or datasets generated by other methods such as satellite imagery, aerial photograph, and other forms of inventory generation. However, since field surveys require physical inspections of the study location, some of the landslide that are not physically accessible may be left out. Some measurements during field surveys may also be wrong due to human errors.

Aerial photograph interpretation: Is the compilation of landslide images captured from the air, the aerial photograph can be made with any camera, and the dataset can span over the years [66]. Aerial photograph provides datasets from locations that are not physically accessible and can generate more information about a particular study location due to its high mobility nature.

Google earth image interpretation: The number of satellites that can provide images of the earth's surface daily has increased. These satellite images can be used for landslide susceptibility mapping studies. However, the primary constraint is the cost, and the need for special software to process the satellite images makes it difficult to use in landslide susceptibility mapping. Many landslide susceptibility mapping studies like [67,68] have relied on Google earth image interpretation mainly for free of cost and easy accessibility. Google earth can be considered a reliable and valuable dataset source for generating LSM.

Satellite imagery: The number of satellites consistently obtaining the earth's surface has increased significantly in the last decades. The earth surface observation data can be obtained in multiple forms from multispectral, optical, radar-based to hyperspectral satellite image [69]. The data can have resolutions from medium to high (0.3 m to >50 m) and a revisiting time of a few days to a few weeks. The satellite imagery can also cover areas from small to large, making it suitable for monitoring earth surface deformation such as landslide. However, using real-time high-resolution satellite imagery for the LS susceptibility mapping may be cost-ineffective.

Accurate datasets and landslide inventory maps are essential for the correct prediction of landslide susceptibility. Many methods are employed for generating landslide dataset and landslide inventory. The most popular are historical data and field surveys. With satellite data and remote sensing technology advancements, abundant high-resolution datasets are available for landslide studies. Researchers have used landslide points ranging from less than 100 to greater than 1000. The data of the existing landslide points are used to generate the LSM, with the assumption that similar conditions of the existing landslide will cause a new or recurrence of landslide. A large dataset is required to train the ML models. Wang et al. [70] have used the synthetic minority oversampling technique (SMOTE) to explore the effect of sample size on the accuracy of ML models. By employing SMOTE, they artificially generated landslide points and could increase the size of the sample/dataset by two to thirty times the original size. Different sample sizes were generated to match the non-landslide locations and landslide sample locations for the study. The experimentation results show SMOTE can reasonably improve the performance of the ML models used for the landslide susceptibility study. Liu et al. [61] used Geo-detectors and Q-statics to generate four datasets. They are manual points with 13 features dataset (MPD13), random points with 13 features dataset (RPD13), MPD9, and RPD9 for training and validation. The study results show that MPD9 had the best performance, with an AUC value of 0.8990.

Landslide points of the study location can be collected by different methods such as historical data, field surveys, among other methods, and an equal amount of non-landslide points (balance) is required for the preparation of training and testing datasets. Several authors have selected the non-landslide points randomly. Nhu et al. [42] stated

choosing random non-landslide locations using a trial-and-error method has limitations and suggested a more standardized approach for selecting a non-landslide location for future studies. Liang et al. [71] used an unsupervised ML model to generate the non-landslide locations, and the generated dataset was used by supervised ML to generate LSM. The additional process of using an unsupervised ML model for data generation improved the accuracy of the LSM generated by the supervised ML model. Liang et al. [72] suggest the use of clustering analysis to improve the sampling of non-landslide locations. Zhang et al. [73] explored an imbalanced dataset by using a class-weighted algorithm to solve class imbalance with landslide and non-landslide samples. The study addressed the disadvantage of SMOTE and its limitation with over-fitting. A cost-sensitive matrix was used to represent the cost of miss classification. Based on the G-mean and balanced accuracy value, the class-weighted method produced a better performance for generating LSM.

Position accuracy of landslide inventory impacts the accuracy of LSM generated because LCFs are derived using these landslide data points. However, the data point may not be able to represent the whole area of landslide. Abrahan et al. [43] explored the positional accuracy and sampling strategy of landslide inventory. The study compared landslide polygon data and point data for generating LSM. The results indicate that the accuracy of KNN, SVM, and RF increased with polygon data, whereas NB and LR methods showed a slight decrease. They have used the k-fold validation method to determine the test and validation ratio. Rong et al. [74] explored borderline-SMOTE and random under-sample methods to handle imbalanced data. Furthermore, landslide data points located at the center of the landslide scarp were found to be the best landslide sampling strategy since the size and shape of landslide are always different. Yilmaz et al. [75] explored the effect of sampling techniques on the performance of ML-based LSM. The authors used three different datasets, map1 (landslide scarp), map2 (landslide seed-cell), and map3 (point). The study result shows that the dataset with the scarp technique had a more realistic LSM. Hu et al. [54] suggest the use of point landslide data as landslide inventory data to increase mapping efficiency, avoid uncertainty in landslide boundary, decrease spatial autocorrelation among case samples, and provide uniform treatment of diverse landslide sizes.

The ratio of dataset splitting for training and testing ML models is also important; an incorrect choice will lower the accuracy. Sahin et al. [76] explored the effect of sampling ratio on the performance of ML-based LSM to determine the best testing and validation ratio. The study used nine sampling ratios of 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20 and 90:10 and the performance of the ratio was evaluated using root-mean-square error (RMSE). The 90:10 had the highest prediction error using LGR, and 70:30 had the least RMSE. The 70:30 was considered the best sampling ratio based on the RMSE value. The authors also used a stratified random sampling technique to select pixels for non-landslide samples. Kalantar et al. [25] studied the effects of training data selection on LSM. They prepared five different sets of training and testing data generated randomly, maintaining the split ratio of 70:30. The study result shows random data selection has an influence on the LSM generated. Yao et al. [21] used 2031 landslide data that occurred before 1990 as training data and 229 landslide locations after 1990 as validation dataset. The authors used two types of SVM; a two-class SVM and a one-class SVM. The two-class SVM used a stable site class derived from random regions with a 40m buffer away from a known landslide location and the existing landslide area as the other class. The one-class SVM used only failed case data. Both the one-class and two-class SVM methods produced acceptable LSM with small training samples, and the two-class method had the better performance. Karakas et al. [77] classified the landslide inventory data into four classes; inactive mass movements (L1), active mass movements (L2), areas with new active location after 2018 inside the existing landslide (L3), and new landslide after 2018 (L4) triggered by an earthquake. The authors used an landslide inventory map generated before the earthquake as training data and new landslide locations after an earthquake as validation

dataset. Landslide data before and after an earthquake had considerable influence on the performance evaluation of ML models used for LSM.

4.3. Evaluation Methods

Many performance evaluation methods are used to assess the performance of the ML methods in landslide susceptibility mapping. Popular evaluation methods are AUC, Accuracy (ACC), Cohen's Kappa coefficient (κ), and RMSE, among other methods. The overwhelmingly common method is AUC which was used in most articles. We will briefly discuss some of these performance evaluation methods.

4.3.1. Area under the Receiver Operating Characteristic Curve

The receiver operating characteristic (ROC) curve graph consists of the sensitivity, or true positive rate (*TPR*) plotted along the y-axis against the specificity, or false positive rate (*FPR*) plotted along the x-axis. The AUC is the area under the ROC curve and is used to evaluate the performance of classification problems. It measures the capability of ML model in distinguishing the classes. A model is good at distinguishing between classes when the AUC value is high (close to 1) [78]. The terminologies used in AUC and ROC are defined below:

$$TPR/Recall/Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

$$FPR = 1 - Specificity \tag{4}$$

$$= \frac{FP}{TN + FP}. \tag{5}$$

In Equations (2), (3) and (5), *FP* is false positive, *TN* is true negative, *FN* is false negative, and *TP* is true positive. Substituting Equation (3) in Equation (4) we get Equation (5). When AUC is closer to 1, the model can distinguish between different classes better. When AUC is close to 0, the model has poor separability and gives the incorrect result, predicting false cases as true and true cases as false. When the AUC value is at 0.5, the model cannot distinguish between classes. The AUC value can represent the summary of overall performance and is considered the most useful measure in the evaluation of landslide susceptibility models [79].

4.3.2. Accuracy

Accuracy is a statistical metric to evaluate classification models; it is the fraction of the model's right predictions. ACC criteria are obtained using four possible consequences: *FP*, *TN*, *TP*, and *FN*. ACC can be defined as Equations (6) and (7).

$$Accuracy = \frac{No.\,of\,correct\,predictions}{Total\,no.\,of\,predictions}. \tag{6}$$

Positive or negative values are used in calculating the accuracy of binary classification [80].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{7}$$

Many researchers have used ACC to evaluate ML models for landslide susceptibility mapping.

4.3.3. Cohen's Kappa Coefficient

The Cohen's Kappa coefficient is used to measure inter-rater reliability [81], it considers random hits or success by chance [82]. The κ was initially designed to measure the difference in opinion between two people observing the same phenomenon. It can be employed in

classification problems and is recommended for use because it considers unexpected success. The κ is determined using the confusion matrix in classification problems. κ is defined in (8).

$$\kappa = \frac{n \sum_{i=1}^{C} x_{ii} - \sum_{i=1}^{C} x_{i \cdot} x_{\cdot i}}{n^2 - \sum_{i=1}^{C} x_{i \cdot} x_{\cdot i}}. \tag{8}$$

In Equation (8), x_{ii} represents the cell count, the number of examples is denoted by n, the number of class values is C, $x_{\cdot i}$ is the total column count, and $x_{i \cdot}$ is the total row count. κ has a range from -1 (total disagreement), 0 (some random classification), to 1 (perfect agreement). Compared with the ROC curve for binary classification problems, it is less expressive.

4.3.4. Root Mean Square Error

The root mean square error is considered a popular statistical method for assessing the overall performance of ML models. When the RMSE value is low (close to 0), the ML model has a good performance [83]. The RMSE value is calculated using the Equation (9).

$$\text{RMSE} = \left(\frac{1}{n} \sum_{i=1}^{n} (x_{pred.} - x_{act.})^2 \right)^{1/2}. \tag{9}$$

In Equation (9) the total samples of observations is given by n, the predicted value of observation is $x_{pred.}$, and $x_{act.}$ is the actual output of the observation [80].

Evaluation of ML models is important because it indicates the prediction accuracy or the reliability of the generated LSM. The AUC is the most widely used accuracy indicator for landslide susceptibility modeling and to quantitatively compare the performance of the ML models. The AUC value close to 1 is considered good [38]. In general, models with an AUC value between 0.5 to 0.6 are considered to have poor accuracy. Values between 0.6 to 0.7 have average or acceptable accuracy. Values between 0.7 to 0.9 have good accuracy. Values above 0.9 to 1.0 have excellent accuracy [84]. The AUC value is also often used to determine the effect of LCFs on the LSM generated. The inclusion or omission of one LCF can increase or decrease the AUC value indicating the influence of the LCF in generating an accurate LSM [85]. The κ index is another widely used index to determine the classification performance of the ML model in landslide susceptibility mapping. According to the study carried out by Wang et al. [70], the RF was considered a perfect classifier with κ index value > 0.8. The ANN and SVM models achieved a good level of performance with κ index value > 0.6. However, LGR had a moderate performance with κ index value > 0.5. The study [70] also employed ACC as a performance indicator. Based on the study results, RF achieved the highest accuracy with 97.8%, followed by ANN with 95.8%, SVM with 85.1%, and LGR with the lowest accuracy value of 74.1%. The RMSE value is used to quantify how accurately the ML model can classify the landslide susceptibility. The lower the RMSE value indicates better prediction accuracy of the model [86]. In the study carried out by Fang et al. [13], the RMSE value was used to indicate the improvement made by the ML model during the training and validation process. According to the study, the RMSE value of the model decreased as the number of epochs gradually increased, indicating gradual improvement. The training stops when the model converges to a small RMSE value, indicating a good fit and generalizability of the trained model.

4.4. Machine Learning

Landslide susceptibility mapping is a function of the occurrence of landslide and LCFs. LSM is used to depict regions likely to experience landslide and group them into different susceptibility classes ranging from low to very high susceptibility. The ML method uses advanced algorithms to model complex relationships by analyzing factors influencing landslide and non-landslide locations. ML can produce repeatable and highly accurate results through continuous learning. Many ML methods have been employed for landslide

susceptibility mapping, and the most popular ones are RF, SVM, LGR, and ANN. Well-known statistical methods such as LR, the weight of evidence (WoE), and the AHP are generally used as a benchmark for evaluating new or proposed ML models. Recent studies focus on combining conventional ML methods to produce ensemble or hybrid forms and deep learning to improve the performance of ML in landslide susceptibility mapping.

5. Machine Learning in Landslide Susceptibility Mapping

Many ML-based landslide susceptibility mapping has been explored by researchers. The ML models can be grouped into conventional, hybrid, ensemble, and deep learning methods. We shall now discuss these methods.

5.1. Conventional Machine Learning Method

Conventional ML models are standalone ML models and have shown good prediction accuracy, hence popular in the generation of LSM and many other applications as well. Standalone conventional models are also often used as a benchmark for evaluating newer models and in combination with other models in hybrid and ensemble setups. The following conventional models are popular in landslide susceptibility mapping.

5.1.1. Random Forest

Random forest is an ensemble of DTs [87,88]. During training, it constructs many trees based on random subsets of input data. RF predicts by taking the average or votes of different decision trees. Increasing the number of trees reduces overfitting at the expense of computational complexity. RF can be used for classification and regression. For classification, each decision tree predicts a class and the final result is the most voted class. In regression, the dependent variable is estimated by taking the average results of individual decision trees. The selection of input variables (causative factors) is a critical step in RF because the predictive ability of the model depends on the inputs [89]. The optimal condition for input data is given by $log2(M+1)$ [87], where M is the number of inputs, and the mean-squared error (ϵ) for the RF is defined in Equation (10).

$$\epsilon = \left(v_{observed} - v_{response}\right)^2,\tag{10}$$

where $v_{observed}$ is the observed data, and $v_{response}$ is the result. The average prediction of all the trees is given by Equation (11):

$$S = 1/K \sum K^{th} v_{response},\tag{11}$$

where S represents the forest prediction while K is applied to each tree.

Many studies used RF for landslide susceptibility mapping. Akinci et al. [84] used an RF classification model. Two important parameters in RF are $ntree$, which signifies the count of DTs, and $mtry$, the number of causative factors used in each DT. There is no specific rule or universally accepted value of $ntree$ and $mtry$ in RF [90]. After many trials, the study [84] used 50 $ntrees$ and $mtry$ set to 8. The success rate of RF model was 98.3%, and the AUC value of prediction was 97.7%.

Another study [91] used an RF classification model defined in Equation (12), trained on 700 landslide pixels.

$$o = (i) = max\left[\sum_k f(d)\right],\tag{12}$$

where i is the input data, o is the computed output for max ensemble, and the indicator function $f(d)$ is defined as:

$$f(d) = \begin{cases} 1, & if \ d \ is \ 'YES' \\ 0, & Otherwise. \end{cases}\tag{13}$$

In Equation (13), *YES* and *Otherwise* represents landslide and non-landslide locations. The value *k*, selected after parameters tuning, was 1000, and 3 features were tested at each node. The accuracy of the RF was measured using out-of-bag (OOB), which estimates the prediction error. OOB was <2%, indicating accuracy is >98%.

Zhou et al. [92] did a comparative study of a traditional statistical method certainty factor (CF), conventional ML models SVM and RF, and hybrids of CF-SVM and CF-RF. The study implemented RF consisting of 500 DTs, and the number of randomly selected features was set to three. A total of 464 landslide and non-landslide samples were used to train the RF model. The trained model was tested on 200 test samples, out of which 156 samples were accurately predicted. The RF model had a prediction accuracy of 78%, indicating a good predictive performance.

Micheletti et al. [93] used RF for feature selection. According to the study, removing irrelevant features or variables improves the accuracy of the susceptibility map. The user provided the parameters *ntree* and *mtry*. As previously discussed, more trees reduce overfitting. After detailed analysis, 500 trees and four features were selected to study the contribution of each subset variable and, at the same time, keep the convergence fast throughout the iteration. According to the experimentation result, the RF method had an AUC performance of 0.93. The study focused on selecting causative features and removing irrelevant features to improve the performance of other classification models used in landslide susceptibility mapping.

The study [87] used STATISTICA to implement RF for LSM. ArcGIS was used to transform the causative factors into ASCII data format. The study used 2100 landslide points randomly split into two data groups, the first data group of 1050 for training and the remaining 1050 for validation. They could achieve an AUC value of 0.7934 for RF.

Based on this survey, we find that RF is a popular method for implementing LSM. Reasons for its popularity include; ease of implementation, high prediction accuracy, and less overfitting. It can also rank the importance of LCFs and use the ranked list to remove irrelevant ones, giving better prediction accuracy. Determining the optimum number of trees and random input for each tree can be future scope for landslide susceptibility mapping using RF.

5.1.2. Support Vector Machine

Support vector machine is a supervised ML model. It can be used to solve classification and regression problems. It was initially designed for binary classification but can be extended to solve n-class problems [94]. The basic concepts of SVM are as follows [95]:

- The separating hyper-plane: The idea of SVM is to determine a line or hyper-plane in higher-dimensional space that can separate the input data into classes;
- Maximum-margin hyper-plane: We can have many hyper-planes for separating different classes. Training the SVM aims to determine the hyper-plane with maximum margin where the margin represents the distance from the hyper-pane to the nearby data points (also known as support vectors). The hyper-plane with maximum margin from the data points is the best class-separator selected for the classification;
- The soft margin: If the data points are linearly separable, the maximum-margin hyperplane can separate the classes, but data in real-life applications might not be linearly separable. Soft-margin tolerates few miss-classification or anomalies, allowing the model to generalize better in the case of linearly inseparable data;
- The kernel function: Kernel in SVM is a mathematical function that accepts the input data and transforms the data into the desired form. Generally, the data are transformed to a higher dimension space, where non-linear data becomes separable.

The primary SVM classification function is defined in Equation (14), where number of coordinates is given by n, parameters in vector x in original space is given by x_i, class label by y, parameters of the hyper-plane are weight w and bias b, and the sign function by sgn.

$$y = sgn(f(x)) = sgn\left(\sum_{i=1}^{n} w_i x_i + b\right) = sgn(w \cdot x + b). \tag{14}$$

To deal with real time noisy and non-linearly separable classification dataset, soft-margin is introduced to accommodate empirical errors of the classification model. Equation (14) now becomes Equation (15). The weight is given as $w = \sum_{i=1}^{m} \alpha_i y_i x_i, i = 1, ..., m$ where α_i denotes the weight of the ith training example as a support vector.

$$f(x) = sgn \sum_{i=1}^{m} \alpha_i y_i (x_i \cdot x) + b. \tag{15}$$

Again to deal with non-linear decision boundary, SVM can map data points (x_i) into a higher-dimensional space where the data can be linearly separated, and x_i is replaced by $\phi(x_i)$ where ϕ produce the higher-dimensional mapping. Furthermore, by using kernel function which works on lower-dimension vectors (x_i, x_j) it is possible to produce a value identical to the dot product of higher-dimensional vectors ($\phi(x_i) \cdot \phi(x_j)$). The decision function using kernel trick is given as Equation (16):

$$f(x) = sgn \sum_{i=1}^{m} \alpha_i y_i k(x_i, x) + b. \tag{16}$$

SVM's final output is a probabilistic binary landslide membership of 0 and 1 [93]. Many studies have used SVM for the generation of LSM. Chen et al. Chen et al. [96] explored SVM with four different kernel functions, namely linear, polynomial degree of 2, sigmoid, and RBF, to generate LSM. The SVM models were trained using the ENVI5.1 software, and the result was exported to ArcGIS software for visualization. The landslide susceptibility index was grouped into five classes using the natural break method. Based on the study's outcome, RBF-SVM had the best success rate value of 83.15%, followed by polynomial-SVM with 82.72%, linear-SVM with 81.77%, and sigmoid-SVM with 79.99%. The prediction capability of the four kernel functions was tested using the validation dataset. The RBF-SVM method had the best prediction rate with 77.98%, 77.07% for linear-SVM, 77.50% for polynomial-SVM, and 76.08% for sigmoid-SVM.

The article [97] used Gaussian radial function as its kernel function, where its parameter $\gamma = 2$ and penalty factor C = 100 were found to be the optimum value for model fitting. The datasets were divided into two random sets where 50% of the data were used for training and the other half to test the model's performance. They have an additional step where the training datasets were swapped with the testing datasets, and the model's accuracy is the arithmetic mean of the two steps. All experiments were performed using an open-source package *LIBSVM*. This study also had multiple iterations for its training data points, 1%, 5%, and 50%. The accuracy of each iterations were 1% = 71.0, 5% = 71.0, and 50% = 88. The experiment found that small training datasets are enough to get high accuracy of susceptibility prediction.

Oh et al. [98] evaluated the performance of three ML methods, namely evidence belief function (EBF), LGR, and SVM, to generate LSM. The study used the ENVI 5.0 Exelis visual information solutions software for the SVM model application. The RBF, which is the default kernel function in ENVI, was used because of its ability to accurately predict in non-linear environments. The RBF function used for the study is given in Equation (17).

$$K(X_i, X_j) = exp(\gamma X_i - X_j^2), \tag{17}$$

where γ is the kernel parameter and $K = (X_i, X_j)$ is the kernel function. The sensitivity analysis of each LCF was conducted by excluding each factor from the SVM landslide susceptibility index map. The normalized weight for each LCF was then calculated based on the differences in the AUC values between the SVM landslide susceptibility index, including all LCFs, and the landslide susceptibility index excluding each LCF. The weight ranges from 0, indicating low susceptibility, to 1 indicating high susceptibility. Comparing the AUC values of all the three models used for the study, SVM had the least AUC value with 0.8178. The lower AUC may be due to the choice of SVM optimized kernel function and γ parameter. Even though SVM had the least AUC value, its prediction performance is acceptable for landslide susceptibility mapping.

Yu et al. [40] explored the combination of geographically weighted regression (GWR), SVM, and particle swarm optimization (PSO) to improve the accuracy of LSM. Firstly, the GWR segments study areas into several prediction regions with appropriate sizes. An SVM method is used to classify the landslide susceptibility of each region. To further improve the model's accuracy, the PSO algorithm is used to obtain the best parameters for the SVM classifier. The GWR-PSO-SVM hybrid method produced a highly accurate prediction of landslide susceptibility with an AUC value of 0.978.

Micheletti et al. [93] explored three different ML methods: SVM, RF, and AdaBoost, for the selection of causative features to understand its relation with landslide. They used a Gaussian radial basis function as the kernel function of SVM. To select the causative features by SVM, the weights of the features had to be determined first, which was impossible with the isotropic Gaussian kernel. The anisotropic kernel would also require too much computational power. To overcome such problems, scaling input feature properties was performed before applying an isotropic Gaussian kernel. This approach could get the required output with a small number of input features. Based on the papers we surveyed, SVM is an effective ML method for generating LSM. SVM has the least false-positive rate (using a balanced training set) for a standard ground class which is considered essential in risk assessment application [99]. Further studies can be carried out on the application of automated landslide susceptibility mapping using SVM and exploring different or new SVM kernels to get higher accuracy.

5.1.3. Logistic Regression

Logistic regression is considered the most popular statistical method used in earth science [19]. LGR has its roots in statistical pattern recognition and is regarded as the precursor of ANN, as ANN is a generalization of LGR [100]. It is a supervised ML method mainly used for binary classification problems. Despite its name, it is a classification model rather than a regression model [101]. The relationship between the input and output variables does not need to be linear in LGR [102].

Statistically LGR uses the concept of odds. It is the probability of events occurring divided by the events not occurring. LGR has weights associated with the input. The relationship between weights and output is exponential [103]. LGR is a log transformation of the odds ratio given by Equation (18).

$$Logistic function = \frac{1}{1 + e^{-x}}. \tag{18}$$

In Equation (18), x represents the input variable, and the logistic function is the natural log of the odds. So, 0.5 probability translates to logistic function 0. The value of the logistic function have a range of 0 to 1 given by $P \in [0, 1]$ [102]. LGR is used to predict the binary outcome or state, such as yes/no, success/failure, will occur/will not occur. Using LGR for LSM, we can predict the chances for landslide occurrence in the Region of interest (ROI), given the causative features as input variables.

According to Kavzoglu et al. [27], LGR is one of the most commonly used multivariate methods to produce LSM. The study aimed to find how independent variables (causative

factors) were related to the dependent variable (existence or lack of landslide). Logistic function $Logit(p)$ for LSM can be defined as:

$$Logit(p) = \frac{p}{1-p}. \tag{19}$$

In Equation (19), the dependent variable is p, and $(p/1-p)$ is the odds ratio. Multiple linear regression function can also be implemented using *Logit* transformations and the equation is given by:

$$Logit(pi) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i. \tag{20}$$

In Equation (20), β_0 is intercept, β_i is the coefficients that measure impact of independent variable x_i, and the number of independent variables is n. Kavzoglu et al. [27] have used eight independent variables (causative factors), and the dependent variable was the existence or lack of landslide. LGR was used to determine the spatial relationship between independent and dependent variables. The slope had the most influence on the dependent variable because of its high coefficient value. On the other hand, NDVI and TWI had negligible coefficient values close to 0, indicating little influence or relationship to the occurrence of landslide. According to the experimentation results, LGR had an accuracy of 78.46% and AUC = 0.868, which is the least compared to other ML methods, namely RF, bagging, rotation forest (RotFor), and support vector regression considered in their study.

The study [13] uses a hybrid method for the generation of LSM. They used standard ML methods such as RF, SVM, and LGR with a convolutional neural network (CNN). CNN can learn and extract features, which improves the performance of the ML methods. LGR with CNN had the best predictive capability compared to other hybrid methods with the overall accuracy (OA) of 79.82%, κ value of 0.5963, and Matthews correlation coefficient (MCC) of 0.6013. Furthermore, LGR had the most visible improvement in the hybrid system with an increase of 8.72% for OA, 0.1743 for κ, and 0.1727 for MCC. The frequency ratio of landslide using CNN-LGR was 42.57%, higher than LGR of 40.93%. The experimentation results indicate that hybrid methods had a better performance than the traditional ML methods, with CNN-LGR achieving the most significant AUC improvement of 0.063.

Yilmaz et al. [19] did a comparative study between SVM, ANN, LGR, and conditional probability. The initial process of landslide susceptibility mapping involves a map showing the area affected by landslide and raster format LCFs converted into ASCII format. A statistical software SPSS 10.0, was then used to calculate the correlation between landslide events and the LCFs. In order to predict the possible occurrence of landslide in each grid, the probability was calculated. The result LSM was generated by converting the file into a raster format.

Li et al. [104] used Random SubSpace (RSS)-based classification and regression tree (RSSCART) as the primary model to study LSM with LGR, CART as the benchmark models. They have used 203 landslide locations and 14 causative factors. The result shows that RSSCART had the best training and validation AUC of 0.852 and 0.827, while LGR came in second with training AUC at 0.797 and validation AUC at 0.758. This result further indicates the superiority of hybrid models. We found that LGR is an effective model for LSM with the AUC performance measures consistently above 70%, meaning good prediction performance. However, it underperformed compared to other ML methods in most papers. The LGR used in combination with other ML models generally gives higher performance.

5.1.4. Artificial Neural Network

The artificial neural network gets its inspiration from biological neurons. It replicates the functionality of neurons and their interconnections to process information in parallel. The basic architecture of ANN consists of interconnected neurons organized in different layers; it has an input layer, a single or multiple hidden layers, and an output layer.

A numerical value called weight is associated with every connection between neurons. Equation (21) defines the output h_i of the hidden layer neuron i.

$$h_i = \sigma\left(\sum_{j=1}^{N} V_{ij}x_j + T_i^{hid}\right). \tag{21}$$

In Equation (21), $\sigma()$ is the activation function, N the number of input neurons, V_{ij} represents weights, x_j represents inputs to the input neurons, and the bias or threshold of the hidden neurons is given by T_i^{hid} [105]. The activation function is introduced to add non-linearity into the hidden layer and bind the neuron value. An activation function that is popularly used is the sigmoid function, given in Equation (22).

$$\sigma(u) = \frac{1}{1 + exp(-u)}. \tag{22}$$

Other common activation functions used in ANN include arc tangent and hyperbolic tangent. Both functions respond to the input similarly to the sigmoid function, but the output range is different. ANN is supervised and requires a training dataset. The training objective is to determine the optimal weights that can predict the output correctly. Once trained, it can predict the output for a given input.

Several researchers have used ANN for generating LSM. Choi et al. [106] developed a $14 \times 30 \times 2$ ANN. Fourteen causative factors such as the thickness of soil, the density of the forest, slope, land cover, and other factors extracted from Landsat TM satellite images were used as the input to the network. The ANN was trained using a backpropagation algorithm. The study's objective was to determine the weights of the causative factors. The trained network was validated with the data collected from Korea's Youngin, Janghung, and Boeun regions. The factors with higher weights were land cover, slope, and distance from the lineament. In contrast, forest density and soil type had lesser weights. All the study areas had similar causative factor weights. The trained ANN was evaluated in nine study locations, achieving the best accuracy of 81.36% and the worst of 71.72%.

Lucchese et al. [107] explored the use of the Mamdani fuzzy inference system (FIS) for LSM and compared it with ANN. FIS has not been explored in previous LSM papers considered in this survey; therefore, the article adds significant value to the field. All the metrics indicate that ANN has less uncertainty as compared to FIS. Both the methods had good overall prediction accuracy. The AUC values of the models were FIS (0.8886) and ANN (0.9409).

Yilmaz et al. [108] carried out a case study using ANN to generate LSM. The study used a backpropagation-based ANN model. The backpropagation algorithm used was based on the generalized Widrow–Hoff learning rule. The network consists of input layers, multiple hidden layers, and output layers. The learning involves selecting small random-sized numbers (0 to 1) for the weights and the biases, calculating the network's output, and comparing it with the expected output. If the network is close to the expected output, then continue. Otherwise, weights are corrected using the correction rule. The ANN had an AUC value of 0.847, indicating an good prediction accuracy.

Quan et al. [109] explored the use of ANN for landslide susceptibility mapping and did a comparative study with a knowledge-based method AHP. The sample size of the network was chosen from the range $30 \cdot N_i \cdot (N_i + 1)$ and $60 \cdot N_i \cdot (N_i + 1)$, where N_i is the number of the input layer. After extracting the samples from the study location, the neural network analysis was carried out. The number of neurons was 7, 14, and 2 in the input, hidden, and output layers, respectively. The number of weights between the input and the hidden layer was 98, and 14 between the hidden layer and output layer. The network was trained in 50,000 learning cycles. The learning rate and momentum are essential parameters that influence the convergence of the ANN model. The study explored sixteen pairs of learning rate and momentum values. They have tested the performance of ANN and AHP

models in twelve different areas. The AHP could identify seven areas of the field data correctly, and the ANN model could correctly identify nine areas of the field data.

Wang et al. [110] compared ANN and WoE. The proposed ANN has 12 nodes in the input layer, 25 nodes in the hidden layer, and two nodes in the output layer. The input data were normalized in the range of 0.1 to 0.9. The initial weight was randomly selected, and the learning rate was set to 0.01. The stopping criterion RMSE was set to 0.01. The 12 LCFs and the final weights between layers of the neural network were used to generate the LSM. The success rate in AUC of the models was ANN (0.8251) and WoE (0.7982). The prediction rate was ANN (0.7731) and WoE (0.7459).

Zare et al. [22] explored two ANN models, MLP and RBF, for landslide susceptibility prediction. The results showed that MLP with the Broyden-Fletcher-Goldfarb-Shanno learning algorithm produced better predictions than the RBF algorithm. The success rate of the algorithm was MLP (0.9193) and RBF (0.9085), and the prediction rate was MLP (0.881) and RBF (0.8724).

Can et al. [111] explored two different ANN structures, which are composed of single and double hidden layers. Furthermore, four different training algorithms, namely quick propagation, batch backpropagation, Levenberg–Marquardt, and conjugate gradient descent algorithms, were used to train the ANN models. ANN-CGD had the best overall prediction accuracy with an AUC value of 0.817. The models trained by CGD had the highest prediction accuracy but were slower than the other algorithms.

Harmouzi [112] use MLP-ANN to generate LSM for the coastal region of Morocco. The study classified landslide into four types: complex, slide, rockfall, and flow. The MLP-ANN method generated LSM for all four landslide types. ArcGIS software was used for preparing the data and MATLAB for supervised learning. MLP-ANN model had an AUC of greater than 0.90 for all the landslide types. We found that the ANN can generate LSM with very high prediction accuracy.

5.1.5. Naive Bayes

Naive Bayes is a learning algorithm based on the Bayes' rule. NB assumes that the features considered for classification are independent of other features given the class [113]. However, the assumption of independence is often violated in practice, and NB provides a competitive classification accuracy. It provides a means to use the information provided by the sample data to estimate posterior probability $P(y|x)$ of each class y given an object x. The followings are the features of NB:

- Efficient computation: Training and classification time are linear to the number of features. While the number of training examples does not affect classification, training time is linear to the number of training examples;
- Less variance: NB does not utilize search and hence has less variance. However, this may also result in high bias;
- Cumulative learning: NB works on the learning of lower-order probabilities using the available training data, and the probability can be updated once new training data are acquired;
- Posterior probabilities can be directly predicted using NB;
- Robustness: All the features are used for its prediction. Hence it is not affected by noise.;
- Handling missing values: Missing features do not affect NB because it always uses all its features for all its predictions, so the missing attribute effect is not noticeable.

NB is a Bayesian network (BN) classifier and is defined in Equation (23).

$$P(y|x) = \frac{P(y)P(x|y)}{P(x)}.$$

(23)

Taking into consideration that the features are conditionally independent. The feature-value data, $P(y|x)$ can be expressed as Equation (24).

$$P(y|x) = \prod_{i=0}^{n} P(x_i|y). \tag{24}$$

In Equation (24) the ith feature in x value is x_i, and the number of features is given by n. Equation (23) can be calculated by normalizing the numerators of the right-hand-side of Equation (25).

$$P(x) = \prod_{i=1}^{k} P(c_i)P(x|c_i). \tag{25}$$

In Equation (25), k is the number of classes and ith class is given by c_i. The NB classifier is similar to LGR, the only difference being the manner by which features are selected.

Lee et al. [114] demonstrated how the NB model can be used for LSM. The study used 17 LCFs vectors x_i, and y is the occurrence or non-occurrence vector. The NB classifier is defined in Equation (26).

$$y_{\text{NB}} = \underset{y_i = [event, non-event]}{\overset{argmaxP(y_i)}{}} \prod_{i=1}^{17} P\left(\frac{x_i}{y_i}\right) \tag{26}$$

$$P\left(\frac{x_i}{y_i}\right) = \frac{1}{\sqrt{2\pi\alpha}} e^{\frac{-(x_i - \eta)^2}{2\alpha^2}}. \tag{27}$$

In Equation (27) the conditional probability is $P(x_i/y_i)$, where $P(y_i)$ represents prior probability of y_i, the standard deviation α for x_i, and the mean η. The result of the study indicates a satisfactory prediction accuracy with 78.3% accuracy and a success rate of 79.2%. Since the accurate prediction of landslide is difficult because it depends on many factors, the study's main objective was to use NB prediction as a preventive support measure to control the damage caused by landslide.

Imtiaz et al. [115] aim to compare LSM generated using ML models. The study used three ML, namely RF, SVM, and NB. Eleven LCFs were used to determine the susceptibility index of the case study location. The study implemented NB using the e1071 library in R. The NB model was used to classify the pixels into landslide and non-landslide classes. The accuracy of NB was in the range of 96% to 97%.

Hu et al. [48] conducted a case study using ML models with fractal theory (FRT) for landslide susceptibility mapping. The study used two ML models namely SVM and NB. The prediction using NB is given by the Equation (28)

$$y_j = \underset{y_j = \{ls, non-ls\}}{\overset{argmaxP(y_i)}{}} \prod_{i=1}^{10} P(x_i|y_j) \tag{28}$$

In Equation (28), y_j represents the landslide and non-landslide category to be predicted, $P(y_j)$ is the prior probability of y_j, $P(x_i|y_j)$ is the conditional probability. The study uses 10 LCFs. The accuracy of the ML model with FRT used for selecting non-landslide locations produced a better prediction accuracy with an AUC value of 0.969 for SVM and 0.989 for NB. The AUC value for random selection of non-landslide locations was 0.708 for SVM and 0.727 for NB. The AUC value of the ML models indicates that the quality of non-landslide location has a considerable influence on the accuracy of LSM generated.

Youssef et al. [37] implemented NB algorithm using R 3.0.2 and the rminer package for modeling LSM. AUC and RMSE were used to evaluate the model performance. The performance of NB was AUC = 0.916, and RMSE = 0.464. Based on the survey of NB classification model, we found the NB model has a good prediction accuracy and success rate with an average AUC above 80%. NB classifier can be an effective model for generating LSM and predicting the occurrence of landslide.

5.2. Hybrid Techniques

The selection of causative factors greatly influences ML models' accuracy for generating LSM. The use of redundant LCFs or the lack of causative factors with close co-relation to the occurrence of landslide can reduce the prediction accuracy of an ML model. Therefore, many studies have employed different optimization and feature selection methods to improve the overall accuracy of the LSM generated using ML models. Combining a feature selection and optimization technique with ML models or a combination of multiple ML models is called a hybrid method. The conventional techniques for selecting causative factors includes multi-collinearity analysis by calculating VIF and tolerance (TOL) [46,47,116], and co-relation methods such as Pearson's correlation analysis [49], CFS [50], CAE [55], and Spearman's rank correlation coefficient [57]. ML models were also used for feature selection, some of the models includes SVM [52] and RF [44]. Other methods of factor selection and analysis include information gain, CSAE, GeoDetectors, FR, Fisher score analysis, Relief-F, and One-R, among other methods. The inclusion of factors selection and analysis has a considerable positive influence on the accuracy and performance of ML-based LSM.

5.2.1. Literature on Hybrid Techniques

A total of 23 articles were identified on hybrid-based LSM generation. Indicating a favorable research trend on hybrid methods for landslide susceptibility mapping. Table 3 presents the overview of studies on hybrid-based LSM generation.

Peng et al. [41] proposed a novel hybrid model by combining the rough set (RS) theory and SVM model. The study's main objective was to assess landslide susceptibility at a regional scale using multi-source data to generate LSM. The RS method was used for feature selection to identify the essential LCFs and SVM to predict landslide susceptibility. The RS-SVM had better overall goodness of fit than the general SVM model, which increases the performance of the ML model for the generation of LSM.

Yu et al. [40] proposed a novel hybrid method. The study aimed to consider environmental factors on a local scale by using the GWR method to provide better prediction accuracy. A hybrid technique consisting of GWR-PSO-SVM was used to generate a reliable LSM. The importance of each LCF was obtained using SPSS Clementine 12 software. GWR-PSO-SVM was able to provide accurate predictions with an AUC value of 0.978.

Pham et al. [86] proposed a hybrid method using sequential minimal optimization (SMO) and SVM (SMOSVM). The hybrid method was proposed to overcome the limitation of SVM. Several advantages were observed in combining SMO with SVM. The main benefits of using SMO were simple, fast, and easy algorithm implementation, better results while using extensive data but less input, and reduced problems' complexity. The proposed hybrid model SMOSVM was compared with another hybrid model cascade generalization optimization-based SVM and other conventional ML models such as SVM and NB-tree. The SMOSVM produced the highest AUC value of 0.824, indicating that the proposed hybrid model can be considered an effective method for generating LSM.

Zhang et al. [117] aim to quantitatively predict the extent of landslide. A hybrid method using fractal dimension with index of entropy (IoE) and SVM was developed for the study. The 10-fold cross-validation and information gain ratio (IGR) was used for feature selection. The fractal dimension provided a better spatial distribution of landslide. The authors stated combining fractal IoE and SVM in developing the hybrid technique for LSM is new in the literature. The hybrid methods fractal IoE (AUC = 0.8591) and fractal SVM (AUC = 0.9761) produced better LSM than the standalone IoE (0.7434) and SVM (0.7946).

Table 3. Literature on hybrid-based landslide susceptibility mapping.

Author	Year	Hybrid Method
Peng et al. [41]	2014	Novel hybrid method combining rough set theory and SVM
Yu et al. [40]	2016	SVM with geographical weighted regression and PSO
Pham et al. [86]	2019	Novel hybrid method using sequential minimal optimization and SVM
Zhang et al. [117]	2019	Fractal dimension with index of entropy and SVM
Adnan et al. [118]	2020	LSM generated by combining the LSM produced by four ML models KNN, MLP, RF, and SVM
Wang et al. [119]	2020	GeoSOM with RF and ensemble ML model consisting of ANN-SVM-GBDT
Fang et al. [13]	2020	Proposed three hybrid models CNN-SVM, CNN-RF, and CNN-LGR
Hu et al. [48]	2020	Combining fractal theory with SVM and NB
Rong et al. [74]	2020	Combination of Bayesian optimization with RF and GBDT
Wang et al. [55]	2020	Integration of MultiBoost with RBFN and CDT
Sahana et al. [120]	2020	Multi-layer perceptron neural network classifier with ensemble ML models like Bagging, Dagging, and DECORATE
Xie et al. [79]	2021	GeoDetector using factor detectors and interaction detectors with four ML models ANN, BN, LGR, and SVM
Alqadhi et al. [121]	2021	Four optimized ML model namely PSO-ANN, PSO-RF, PSO-M5P, and PSO-SVM with LGR
Arabameri et al. [122]	2021	Credal decision tree based hybrid models namely CDT-bagging, CDT-MultiBoost, and CDT-SubSpace
Saha et al. [123]	2021	Hybrid ensemble method using RF as a base classifier and ensemble methods, namely RotFor-RF, RSS-RF, and bagging-RF
Xing et al. [124]	2021	The output of ML models namely back propagation, RF, and SVM are combined using weight factors
Hu et al. [125]	2021	Fuzzy c-means clustering and factor analysis with LGR
Zhou et al. [51]	2021	RF with GeoDetector and recursive feature elimination
Sun et al. [126]	2021	GeoDetector and RF
Lui et al. [61]	2021	GeoDetector with RF
Liang et al. [71]	2021	Combination of unsupervised and supervised ML method
Dung et al. [127]	2021	Novel hybrid method consisting bagging-based rough set and AdaBoost-based rough set
Wei et al. [128]	2022	Spatial response feature with ML classifiers

The objective of Adnan et al. [118] was to reduce uncertainties in landslide studies and determine the spatial agreement of ML-based LSM generation. LSM was generated by the combination of four ML models, namely KNN, MLP, RF, and SVM. The LSM generated by the hybrid/combined method had a better spatial agreement with correlation coefficients ranging between 0.88 and 0.92 and had better prediction accuracy than individual ML models. Some limitations of the study include the use of landslide inventory data collected from secondary sources, lack of data, and the DEM used for the analysis had low RMSE accuracy.

Wang et al. [119] implemented a hybrid method using GeoSOM, RF, and ensemble ML model consisting of ANN-SVM-Gradient boosting decision tree (GBDT). RF and Pearson correlation coefficient was used for feature selection. GeoSOM was used to cluster the study location into clusters of the homogeneous region to solve the heterogeneity problem. The ensemble ML model consisting of ANN-SVM-GBDT was used to generate the LSM. A few limitations of the study include random non-landslide samples selected from landslide-free locations. The process was time-consuming and error-prone, which can affect the performance of ML models. The study used Thiessen polygons to ensure each pixel is given a cluster attribute for landslide modeling. However, this clustering approach may not be the best way to assign cluster attributes to all grid cells.

Fang et al. [13] assessed the generation of LSM by integrating CNN with three conventional ML models, namely SVM, RF, and LGR. The proposed hybrid models were CNN-SVM, CNN-RF, and CNN-LGR. CNN was used for feature selection and effectively improved the performance of the conventional ML models. The authors noted overfitting was an issue with the study.

Hu et al. [48] improved the predictive performance of ML models by including FRT to select the non-landslide locations in a hybrid setup. Three datasets samples for landslide

and non-landslide were obtained, one using FRT, the second landslide and non-landslide selected from regions with less than 5° slope, and the third using a random selection of landslide and non-landslide locations. The three samples were used as input for SVM and NB models. The study result showed better performance for SVM and NB models using landslide and non-landslide samples selected using FRT.

Rong et al. [74] explored hybrid techniques, which include the combination of Bayesian optimization (BO) with RF and GBDT. The study also used borderline-SMOTE and random under-sample techniques to handle imbalanced datasets. The samples produced were split into 70% training and 30% validation datasets. The BO method was used for hyperparameter tuning. With BO, the prediction accuracy of the RF and GBDT models was increased by 1% and 7%, respectively.

Wang et al. [55] proposed two hybrid models, which is an integration of MultiBoost with two ML models, radial basis function neural network (RBFN) and Credal decision tree (CDT). The hybrid models and GIS were used to generate LSM for the case study location. The correlation attributes evaluation method was used to determine the influence of LCFs. MultiBoost with CDT had the best AUC value of 0.77. However, the AUC value for all the methods was lower than other studies.

Sahana et al. [120] developed a hybrid model using multi-layer perceptron neural network classifier (MLPC) as the base classifier with bagging, dagging, and DECORATE ensemble methods. Factor selection for the study was implemented by multicollinearity assessment using VIF and TOL. The study result highlighted the improvement made by the hybrid method, with bagging-MLPC achieving the highest AUC value of 0.965.

Xie et al. [79] explored the use of GeoDetector to find spatial autocorrelation and heterogeneity of LCFs. They developed hybrid techniques using GeoDetector with factor detectors and interaction detectors, along with four conventional ML models ANN, BN, LGR, and SVM. The result shows that SVM with GeoDetector had the best performance. However, the large gap between training and validation accuracy indicates that the models were overfitting which can be addressed in future studies.

Alqadhi et al. [121] developed four optimized ML models, namely PSO-ANN, PSO-RF, PSO-M5P, and PSO-SVM, integrated with the LGR model to improve the accuracy of LSM generated. The AUC value of 0.962 for the LGR-based hybrid model outperformed all optimized ML models indicating that hybrid models improved the prediction accuracy of optimized ML models. The authors highlighted the study's limitations, including the use of lower resolution satellite images, limited rainfall gauge, semi-quantitative method due to lack of data, and not enough existing landslide location for training the ML model that could affect the performance of the ML models.

Arabameri et al. [122] explored CDT-based hybrid model for LSM generation. The study compared CDT as a standalone model and CDT-based hybrid models, including CDT-SubSpace, CDT-bagging, and CDT-MultiBoost. The CDT-MultiBoost with an AUC value of 0.993 had the best performance.

Saha et al. [123] used hybrid ensemble methods, with RF as the base classifier, namely, RotFor-RF, RSS-RF, and bagging-RF. The study included the process of factor selection using multicollinearity assessment, and the significance of the LCFs was analyzed using CSAE and IGR techniques. Based on the result of the study, the authors suggested ensemble techniques were better suited for mapping landslide hazards and disasters. The few limitations of the study include lack of data and high-resolution DEM, satellite imagery, and other rasterized map for the study location. The authors also suggest including more LCFs to improve the prediction accuracy of the ML models.

Xing et al. [124] used a hybrid approach where the result of the three ML models, back-propagation (BP), RF, and SVM, were combined using an objective function. Weights were assigned to the output LSM of the individual ML model, and the Grey Wolf optimization algorithm was used to compute the weight coefficients. The combined hybrid method produced the best LSM with an AUC value of 0.790. The AUC value, when compared to other articles, is low.

Hu et al. [125] explored the hybrid implementation of LGR coupled with fuzzy c-means clustering (FCM) and factor analysis (FA) for the generation of LSM. The two methods, FCM and FA, were introduced to compensate limitation of the LGR model. The study result shows that the hybrid method performed better in prediction accuracy and generalizability. The FA-LGR method had the best performance, with an AUC value of 0.827.

Zhou et al. [51] proposed a hybrid method using GeoDetectors and recursive feature elimination method with RF. The study emphasized feature selection to improve the ML model's accuracy for LSM generation. GeoDetectors and RFE were used for LCFs optimization to reduce factor redundancy and collinearity in the data. The result shows the ability of factors optimization techniques to improve the prediction accuracy of LSM generated using ML models. The few limitations of the study include the unavailability of uniform resolution for all LCFs and the non-existence of standards for optimal selection of landslide and non-landslide location ratio. Furthermore, the study suggests using ML-based techniques such as spatial autoregressive and simple linear regression models to reduce spatial heterogeneity and correlation.

Sun et al. [126] implemented a hybrid method using GeoDetector (Geo), RFE, and RF. The hybrid model of Geo-RFE-RF was used to eliminate redundant and noise factors. The factor filtration process reduced twenty-two LCFs to thirteen factors, and the predictive performance of Geo-RFE-RF was better. Geo-RFE-RF with filtered factors had the highest AUC value of 0.982, indicating accurate prediction by the hybrid model.

Lui et al. [61] implemented a hybrid model using GeoDetector for factor optimization to improve the performance of RF in the generation of LSM. The performance of GeoDetector, a tool to detect and utilize spatial heterogeneity, was tested using four datasets prepared using different optimization combinations. The four dataset samples include manual point with 13 features dataset (MPD13), random point with 13 features dataset (RPD13), MPD9, and RPD9. The MPD9 dataset had the best accuracy, with an AUC value of 0.8990. The result shows that the hybrid method using factor optimization improved the performance of ML models. The authors mentioned dynamic combination of quantity and quality of LCFs on the predictive performance of ML models is yet to be explored.

Liang et al. [71] proposed a hybrid method where unsupervised ML models consisting of FA and k-means clustering were used for generating LSM. The generated LSM was used to select the non-landslide locations for modeling the supervised ML model using GBDT. The hybrid method had an AUC value of 0.976, which indicates that the hybrid model performed well in generating LSM.

Dung et al. [127] developed a novel hybrid method consisting of bagging-based rough set (BRS) and AdaBoost-based rough set (ABRS) for the generation of LSM. One-R algorithm was used for feature selection. After filtration of LCFs, hybrid-based models consisting of BRS and ABRS were used to generate LSM. The output was compared with SVM and rough set models. BRS had the highest AUC value of 0.845, demonstrating that the hybrid method accurately predicted landslide susceptibility.

Wei et al. [128] used a hybrid method comprising spatial response (SR) feature with ML classifiers (SR-ML). The technique consists of three steps: extracting spatial features using depthwise separable convolution, extracting features on a different scale using spatial pyramid pooling, and ML classification using the features extracted in the previous steps. The method produced a better result than the ML classifiers considered in the study.

5.2.2. Discussion on Hybrid Techniques

Accurate datasets and feature selection help in improving the prediction accuracy of ML models. The objective of many hybrid LSM generation methods is to filter redundant LCFs and enhance the quality of datasets by integrating optimization techniques with ML models. Some of the hybrid techniques used for the generation of LSM include spatial response feature with ML, GeoDetector with ML, PSO with ML and LGR, CDT with ensemble methods, RF with other ensemble methods, the combination of LSM output using benchmark ML models, FCM and FA with LGR, RF with GeoDetector and RFE, GeoDetector

with RF, the combination of unsupervised learning and supervised ML method, bagging with RS and AdaBoost, GeoSOM with RF and ensemble methods, CNN with conventional ML models, FT with ML methods, BO with RF and GBDT, MultiBoost with RBFN and CDT, MLPC with ensemble methods, SMO with SVM, fractal dimension with IoE and SVM, SVM with geographical weighted regression and PSO, and RS theory with SVM.

The experimentation results of the articles surveyed show hybrid-based LSM generation method can accurately predict landslide susceptibility. There is significant performance improvement in the hybrid method because it can solve the limitation of the conventional ML method. landslide susceptibility mapping is a classification problem. The ML models are used to classify the susceptibility, and most of the ML models used in the generation of LSM do not have filtration or feature selection capability. Integrating feature selection techniques such as GeoDetectors, CNN, and RFE, among other methods, can filter irrelevant LCFs, improving the ML models' performance. Researchers have also used hybrid techniques for the generation of datasets. The selection of non-landslide locations using unsupervised ML learning and FT could generate adequate non-landslide samples for the supervised learning ML method, thereby improving the prediction accuracy of the LSM. Using hybrid techniques in determining geospatial factors also positively impacts the accuracy of ML methods. Use of GeoDetector to find spatial autocorrelation and heterogeneity of LCFs, SR feature extraction for spatial feature and scale of the feature, GWR to obtain LCFs at a local scale, fractal dimension for better spatial distribution, and GeoSOM to cluster regions for solving the heterogeneity problem can enhance the LSM. Combining the output of benchmark ML models using weight factors also improves the quality of the LSM.

5.3. Ensemble Techniques

The conventional ML models used for generating LSM have their limitations. Several studies in landslide susceptibility have used hybrid or ensemble techniques to overcome the limitations and improve prediction accuracy. The ensemble combines the conventional ML models using different averaging or voting systems. With the ensemble approach, the limitation or bias of one conventional ML algorithm will be compensated by another method, thereby producing a more accurate prediction. For this reason, many landslide susceptibility studies have employed ensemble techniques to generate LSM [129]. The ensemble method has been in existence for quite a while and has recently gained popularity in landslide susceptibility mapping studies.

Several ensemble techniques were employed for landslide susceptibility studies. Some of the ensemble techniques used in different landslide studies include lightGBM [73] which is an ensemble algorithm based on DT. It belongs to the boosting ensemble technique family. Microsoft introduced it as a new gradient boosting framework to overcome the limitations of GBDT in massive data. RF [57,72,73,76,123,130,131] is a basic ensemble ML model used in the generation of LSM and is considered a conventional technique because of its popularity in landslide susceptibility mapping. It belongs to the Boostrap aggregation family. RF is a preferred ML model for many ML tasks mainly because it is simple to implement, has low computational overhead, and is robust. Rotation forest [39,47,50,59,76,123,130,132] is a DT based ensemble model. It is a popular ensemble method used in many landslide susceptibility studies. The algorithm works on feature extraction where the training dataset for the base classifier is split into subsets, and PCA is applied to each of the subsets. The technique was proposed by Rodriguez et al. [133] and provides better performance than the individual ML model. Random subspace [39,50,123,130,134,135] is another popular ensemble technique in landslide susceptibility studies. The algorithm was proposed by Ho et al. [136]. It constructs trees by randomly selecting subsets of the training dataset. The key difference between this model and other techniques is the use of different features on an entity space [50]. The main disadvantage of the method is the problem of over-fitting. Canonical correlation forest (CCF) [47] is an ensemble model based on decision forest and canonical correlation statistic. It is considered an advanced method over RF and RotFor.

The main idea behind CCF is to use many DTs to predict unknown samples using the majority voting method. The key difference with CCF is applying bootstrapped samples before using the estimated components for DT construction. Alternative decision tree (ADTree) [137] is a combination of DT with boosting algorithm. The algorithm grows the DT using boosting for the prediction, where two prediction nodes are generated at each boosting iteration step. The final prediction is the weighted sum of all the prediction nodes. Extremely randomized tree [57] is a tree-based ensemble method that is very similar to RF, which is a combination of CART and bagging. The algorithm randomly selects the splitting point in a node. Chi-squared automatic interaction detection (CHAID) [138] is a DT-based ensemble algorithm. As a multivariate, CHAID can analyze many LCFs for automatic classification in LSM. Some properties of the algorithm include modeling categorical and ordinal data, where continuous data are converted to ordinal data during analysis. It builds a non-binary tree and can grow more than two nodes from a single node. The dependent features and causative factors are presented as continuous, nominal, or ordinal data. Data summarization is similar to LR. It considers missing values as a single category. CHAID does not require pre-processing of the relationship between dependant variables and causative factors.

Natural gradient boosting (NGBoost) [129] is a relatively new ensemble model developed by Duan et al. [139]. The known limitation of existing gradient boosting techniques, such as probabilistic prediction, were compensated using the natural gradient. It was developed to provide a more general approach to the probabilistic prediction method. XGBoost [129,140,141] belongs to the gradient boosting family and is an effective supervised classification model. It is a preferred ensemble technique mainly because of its ability to prevent over-fitting issues by using bagging-bootstrap aggregation. XGBoost can also consider the bias and variance trade-off by implementing feature randomness. Hyperparameters of XGBoost need to be optimized for the preparation of LSM. MultiBoost [39,50] is an ensemble algorithm introduced by Webb [142]. The algorithm is a combination of AdaBoost and the wagging technique. It provides combined advantages such as AdaBoost's high bias and variance reduction and wagging's superior variance reduction. The decision committee generated by MultiBoost produced fewer errors than AdaBoost or the wagging algorithm. MultiScheme ensemble model [50] is a simpler ML technique. The algorithm uses only one classification as the prediction tool and does not depend on the second-level classifications. The main advantages of the algorithm include a more realistic representation of a state and a smaller bias. Real AdaBoost [50] is an improvement over AdaBoost ensemble technique proposed by Schapire et al. [143]. The algorithm employed a real-value classifier and one-half of the log-odds weight predicted by the classifier. The algorithm works by combining weak classifiers repeatedly by using boosting technique. The grouped classifiers provided better performance than the individual weak classifiers. AdaBoost [38,137] known as adaptive boosting, is part of the boosting ensemble family. The algorithm was introduced by Freund et al.[144]. The method employs adaptive resampling of the training dataset to train the base classifiers iteratively. The miss-classified samples are given higher weights in every iteration, and the final classification is the weighted sum of all the base classifiers' predictions. LogitBoost ensemble (LBE) [145] is an ensemble model based on the popular boosting method, AdaBoost. It is an effective algorithm for reducing variance and bias. The main advantage of LBE is the ability to handle noisy samples. Boosted regression trees [131,146] is an ensemble model based on the combination of statistical methods with ML to improve the prediction accuracy of a simple tree. The algorithm is not sensitive to outliers, and the model accuracy can be improved using regression and boosting techniques. Gradient boosting decision tree [57,72] is a boosting based ensemble model. The algorithm use CART as the base classifier. It is a robust classification and regression method. The GBDT employs the negative gradient for the loss value for classification-related problems.

Bagging [38,39,50,54,123,130,132,134] is another popular ensemble technique used in landslide susceptibility mapping studies. Bagging was proposed by Breiman [147] and is frequently used in studies related to natural hazards mainly for its ability to sense changes

in the training variables. It consists of three steps [148]: obtain subsets of the training dataset using random resampling of the original training dataset, use the obtained training subsets to generate multiple classification models, and generate the final classification model by aggregating all the sub-models. Cascade generalization [50] is another ensemble technique. It can improve classification accuracy by removing biases from the training dataset. This is achieved by sequentially using a set of classifiers and introducing new attributes to the original data at each iteration. The new attributes are generated using the base classifier's probability class distribution. Dagging [50,132,135] is an ensemble ML model introduced by Ting et al. [149]. It involves combining the models generated using subsets of the training dataset by the same algorithm or different classification algorithms. It uses the majority voting system to combine the prediction of the sub-models. The disjoint samples are used to generate training subsets for dagging ensemble method. Compared to boosting, it is better at dealing with noise in the training dataset. DECORATE [50,120,132] is an ensemble algorithm first proposed by Melville et al. [150]. The ensemble model, unlike others, creates diversity by generating artificial samples. The artificial samples are generated by taking the standard deviation of the training dataset with Gaussian's rule.

Stacking [54,58,151] is a heterogeneous ensemble method that combines different models. The algorithm has a better non-linear representation and generalization ability by taking advantage of diverse ML models [151]. Blending [151] is a heterogeneous ensemble model. It is a variation of the stacking ensemble model proposed by Toscher [152]. The algorithm employs the hold-out method to split the training dataset into two new subsets, one for training and the other for validation. Use the new training dataset to train the classifiers while the validation dataset is used as the meta-training dataset. The blending ensemble technique can eliminate the information leakage problem. However, it is more likely to overfit. Simple averaging and weighted averaging [151] are simple ensemble techniques. The average of all the base classifiers' predictions is the output prediction for simple averaging. The weighted averaging employs an AUC-based weighted average to integrate the base classifiers. The qualitative matrix ensemble method [56] uses a decision matrix to combine two prediction outcomes to make a comprehensive decision. For example, one base classifier predicts a 'very high' susceptibility for a particular pixel while the other predicts a 'medium' susceptibility. The ensemble susceptibility using the decision matrix for the pixel will be 'high'. The susceptibility level will be the same if both base classifiers predict the same susceptibility level. The semi-quantitative partition [56] is an ensemble model where the region of study is partitioned into several sub-regions based on certain features such as topography in studies related to landslide susceptibility. The best prediction of all the subareas is combined to generate a new LSM. Probability-weighted [56] ensemble method assigns weights to prediction of different base-classifiers. The weights can be determined using certain features, such as the overall success rate. The normalized predictions of all the base classifiers are combined to provide the final prediction. Mean of probability [153] is an ensemble method that works by taking the mean of the probabilities produced by each base classifier. Median of probability [153] is another ensemble method that works by taking the median of the probabilities of each base classifier. It is observed that the median is less sensitive to outliers than the mean probability. Committee averaging [153] is an ensemble method where the probabilities of the selected base classifiers are converted into binary data. The algorithm works on the principle of a simple voting system. The model provides both predictions as well as a measure of uncertainty. Weighted mean of probability [153] is another ensemble technique where the final probability is the mean weight of the probability of the selected base classifiers.

5.3.1. Literature on Ensemble Techniques

A total of 31 articles on ensemble-based landslide susceptibility mapping were identified. A summary of the ensemble methods is presented in Table 4.

From the introduction to the ensemble algorithm, we are aware of many ensemble techniques employed for landslide susceptibility-related studies. We will look at the details of the literature concerning ensemble algorithms.

Althuwaynee et al. [138] explored the use of the DT-based CHAID method to perform the best classification fit for each LCF. The terminal nodes of the tree were then combined with LGR to find the corresponding coefficients of the best-fitting function and assess the optimal terminal nodes. In order to evaluate the proposed model, two LSMs were produced, LSM1 generated using only terminal nodes (CHAID model) and LSM2 generated by integrating terminal nodes in LGR (ensemble model). The best AUC achieved for success rate was CHAID (0.734) and ensemble(0.79), and the prediction rate was CHAID (0.69) and ensemble (0.753). The authors highlighted that the ensemble method, like the one used in their paper, could enhance the accuracy of the models.

Kadavi et al. [38] did a comparative study of various ensemble-based machine learning models such as AdaBoost, LogitBoost, multiclass classifier, and bagging models. The multiclass classifier method had higher prediction accuracy (0.859) compared to bagging (0.854), LogitBoost (0.848), and AdaBoost (0.84) methods. The authors stated ensemble techniques improve the model performance because they can reduce bias, variance, and over-fitting of the base classifiers.

Shirzadi et al. [39] introduced a novel ML algorithm ADTree, which is based on the MultiBoost, bagging, RotFor, and RSS ensemble algorithms under two scenarios of different sample sizes and raster resolutions for spatial prediction of shallow landslides. The IGR method was used for the feature selection of LCFs. The study finding shows that removing factors with low predictive capability improved the performance of the models. RSS and MultiBoost models could further decrease noise and over-fitting problems.

Pham et al. [145] did a comparative study of LogitBoost ensemble, Fisher's linear discriminate analysis (FLDA), LGR, and SVM. The study found that LBE is the best and most promising method compared to the other three ML models considered in the study for the landslide susceptibility mapping. The predictive accuracy AUC value for each method was LBE (0.969), FLDA (0.838), LR (0.840), and SVM (0.957).

Arabameri et al. [146] proposed a novel ensemble approach using a landslide numerical risk factor (LNRF) bivariate model combined with Linear multivariate regression (LMR) and BRT, coupled with radar remote sensing data and GIS for landslide susceptibility mapping. The LNRF and BRT ensemble had the best accuracy. The AUC value for each method was LNRF-BRT (0.912), LNRF-LMR (0.907), and LNRF (0.855). The result shows that LNRF-BRT and LNRF-LMR increase the accuracy of LNRF, and justifying the ensemble approach can improve the predictive accuracy.

Roy et al. [154] proposed a novel ensemble approach by combining the WoE and SVM. The authors explored four SVM kernels: radial basis function, linear kernel, polynomial kernel, and sigmoid kernel. Remote sensing datasets and GIS were used for landslide susceptibility mapping. The AUC for each method was WoE & RBF-SVM (0.87), WoE & Linear-SVM (0.90), WoE & Polynomial-SVM (0.88), and WoE & Sigmoid-SVM (0.85).

Li et al. [46] used three bivariate statistical models, namely WoE, EBF, and IoE, and their ensembles with LGR for the generation of LSM. The AUC value of each methods was EBF-LGR (0.826), IoE-LGR (0.825), WoE-LGR (0.792), EBF (0.791), IoE (0.778), and WoE (0.753). The experimentation results show that the ensemble method has a higher AUC than the base model, and EBF-LGR had the best prediction accuracy.

Hu et al. [49] proposed stacking ensemble models with SVM, ANN, NB, and LGR as the base classifiers. The base classifiers were combined using the stacking technique in various combinations. The ANN-NB-LGR ensemble model had the best prediction accuracy, with an AUC value of 0.940. Based on the result, the authors suggest that stacking ensemble learning models can be an effective tool for modeling landslide susceptibility mapping.

Table 4. Literature on ensemble-based landslide susceptibility mapping.

Author	Year	Ensemble Method
Althuwaynee et al. [138]	2014	CHAID and LGR
Kadavi et al. [38]	2018	AdaBoost, LogitBoost, Multiclass classifier, and bagging models
Shirzadi et al. [39]	2018	ADTree based on the MultiBoost, bagging, RotFor, and RSS ensemble algorithm
Pham et al. [145]	2019	LogitBoost ensemble
Arabameri et al. [146]	2019	Ensemble of landslide numerical risk factor bivariate model with linear multivariate regression and BRT
Roy et al. [154]	2019	Weight-of-evidence and SVM
Li et al. [46]	2019	Ensemble of weight-of-evidence, evidence belief function, and IoE with LGR
Hu et al. [49]	2020	Stacking ensemble of SVM, ANN, NB, and LGR
Di et al. [153]	2020	Ensemble of ANN, generalized boosting model and maximum entropy ML algorithms
Nhu et al [137]	2020	Ensemble model of AdaBoost and alternative decision tree
Nhu et al. [130]	2020	Ensemble of RF with three meta-classifiers bagging, RSS, and RotFor
Sahin et al. [76]	2020	Canonical correlation forest, RF and RotFor
Pham et al. [135]	2020	RBFN ensemble with RSS, attribute selected classifier, cascade generalization, and dagging
Pham et al. [132]	2020	Bagging, dagging, DECORATE, and RotFor
Kalantar et al. [131]	2020	Ensemble of Flexible discriminant analysis, Generalized logistic models, Boosted regression trees, and RF
Song et al. [57]	2020	Stacking ensemble learning method framework to combine CART, RF, extremely randomized tree, GBDT, and XGBoost
Pham et al. [50]	2021	Bagging, Cascade generalization, dagging, DECORATE, MultiBoost, MultiScheme, Real AdaBoost, RotFor, RSS
Saha et al. [123]	2021	Bagging-RF, RotFor-RF and RSS-RF
Kavzoglu et al. [155]	2021	Ensemble of CNN, RNN, and LSTM
Li et al. [151]	2021	Stacking ensemble of CNN and RNN
Fang et al. [58]	2021	Stacking, blending, simple averaging, and weighted averaging
Gong et al. [56]	2021	Qualitative matrix, semi-quantitative partition, and quantitative probability-weighted
Liang et al. [72]	2021	Classification and regression tree, GBDT, AdaBoost-decision tree and RF
Hu et al. [134]	2021	Bagging and RSS-based naive Bayes tree
Kutlug et al. [47]	2021	Canonical correlation forest and RotFor
Hu et al. [54]	2021	Bagging, boosting, and stacking
Fang et al. [59]	2021	Integrating decision trees (DTs) with the RotFor ensemble technique
Zhang et al. [73]	2022	LightGBM and RF
Kavzoglu et al. [129]	2022	Natural gradient boosting compared with RF and XGBoost
Zhou et al. [140]	2022	XGBoost
Zhang et al. [141]	2022	RF and XGBoost

Di et al. [153] implemented four ensemble models: mean, median, committee averaging, and weighted average. The combination of the base classifiers demonstrates the ability of the ensemble models to produce robust and stable outputs compared to the single models. The AUC values of base models are the generalized boosting model (0.84), ANN (0.74), and the maximum entropy model (0.83). The AUC value of ensemble models are mean (0.91), median (0.901), committee averaging (0.899), and weighted mean (0.91). The ensemble improves reliability is testified by the higher AUC values.

Nhu et al. [137] implemented AdaBoost and ADTree ensemble method for the generation of LSM. Their finding supports that AdaBoost-based ensemble methods can reduce the over-fitting and noise problems in the modeling process. The AdaBoost model with an AUC value of 0.96 had better performance than the AdaBoost-ADTree model, which had an AUC value of 0.94.

Nhu et al. [130] implemented ensemble models using RF with three different meta-classifiers bagging, RSS, and RotFor. The AUC value of the RF ensemble with different meta-classifiers is RF-RotFor (0.936), bagging-RF & RSS-RF (0.907), and RF (0.812). Based on AUC values, it is observed that RotFor significantly improved the performance of the RF-based classifier.

Sahin et al. [76] proposed a new ensemble technique named CCF. The CCF was compared with LGR and conventional ensemble learning models, i.e., RF and RotFor, to

test the model's performance, suitability, and robustness. The AUC value of the different methods explored were RF (0.982), CCF (0.970), RotFor (0.966) and LGR (0.826).

Pham et al. [135] proposed four ensemble models for spatially explicit prediction of landslide susceptibility. The study combined RBFN with ensemble techniques such as RSS, attribute selected classifier (ASC), cascade generalization (CG), and dagging for generation of LSM. The training AUC value for each technique were single-RBFN (0.799), ASC-RBFN (0.756), CG-RBFN (0.783), dagging-RBFN (0.773), and RSS-RBFN (0.790). The validation AUC values were single-RBFN (0.79), ASC-RBFN (0.823), CG-RBFN (0.822), dagging-RBFN (0.832), and RSS-RBFN (0.831).

Pham et al. [132] proposed novel approaches to improve the performance of the Credal decision tree by using four ensemble techniques. The techniques include bagging, dagging, DECORATE, and RotFor for generating LSM. The training AUC values were dagging-CDT (0.939), dagging-CDT (838), DECORATE-CDT (0.917), RotFor-CDT (0.946), and CDT (0.88). The validation AUC values were dagging-CDT (0.878), dagging-CDT (861), DECORATE-CDT (0.882), RotFor-CDT (0.886), and CDT (0.842). According to the authors, ensemble methods generally have time-consuming parameter adjustment, which may restrict their development and application in other regions for different purposes.

Kalantar et al. [131] did a comparative study of four conventional ML models and the ensemble consisting of all four ML models. The ML models used were flexible discriminant analysis (FDA), generalized logistic models (GLM), BRT, and RF. The ensemble techniques had the best ROC curve value with 0.904, followed by RF (0.8919), BRT (0.8842), FDA (8641), and GLM (8604).

Song et al. [57] proposed a novel embedded feature selection (EFS) to allow each base classifier to select its own subfeature space. A Stacking ensemble learning method (SELM) framework combines a tree-based classifier (CART, RF, ERT, GBDT, and XGBoost) with a meta-learner LGR to maximize the generalization accuracy. The proposed ensemble method of EFS-SELM had the highest accuracy, with an AUC value of 0.864.

Pham et al. [50] explored twelve ML models, which include decision table, NB, decision table-NB, bagging, RotFor, dagging, CG, DECORATE, multiBoost, RSS, multischeme, and real AdaBoost. The study is a novel combination of ML methods to generate LSM. The RSS-decision-table-NB achieved the best prediction accuracy with an AUC value of 0.839.

Saha et al. [123] implemented bagging, RotFor, and RSS ensemble technique with RF to enhance the accuracy of the RF model. The study also employed feature selection using multicollinearity assessment, and the influence of the LCFs was determined using CSAE and IGR techniques. Based on the results, it was observed that ensemble techniques had better performance for mapping landslide susceptibility. The RSS-RF method produced the highest AUC value with 0.847.

Kavzoglu et al. [155] studied the combination of deep learning and ensemble techniques to overcome the model's low variance and limited generalization capabilities. The authors employed an ensemble technique using CNN, recurrent neural network (RNN), and long short-term memory (LSTM) as the base classifiers. Firstly, the proposed ensemble DL method used the DL models for feature selection. The output of the feature selection process was used by the DL-based prediction classifiers consisting of dropout and dense layer. The prediction output of the individual classifiers was used as input for the final prediction using the ensemble DL block. The ensemble model was able to improve the performance of the DL model up to 7% in overall accuracy.

Li et al. [151] employed a stacking-based ensemble of CNN and RNN DL models to generate LSM. The stacking, a heterogeneous ensemble-learning technique, produced accurate predictions for the landslide susceptibility mapping. In their study, the ensemble model of CNN-RNN had the best performance, with an AUC value of 0.918.

Fang et al. [58] implemented four heterogeneous ensemble models, namely, stacking, blending, simple averaging, and weighted averaging. The result of the study observed that the ensemble technique produced better performance than the conventional ML methods. Furthermore, stacking and blending ensemble models produced better results because they

use meta-classifiers to compensate for the errors of the base classifiers. Simple and weighted averaging also produced good results because it uses AUC values to assign weights.

Gong et al. [56] explored ensemble methods such as qualitative matrix ensemble, semi-quantitative partition ensemble, and quantitative probability-weighted. An extensive comparative analysis was carried out between the ensemble and individual models. The AUC values for the base models were frequency ratio (0.6574), fuzzy assessment (0.7434), BP neural network (0.7098), and SVM (0.7860). The AUC values of the ensemble methods were matrix (0.8378), semi-quantitative partition (0.8009), and quantitative probability-weighted (0.8261). Based on the result, it can be observed that ensemble methods performed better than the base models.

Liang et al. [72] conducted a comparative study of four DT-based ensemble models, namely CART, GBDT, Ada-DT, and RF. The authors suggest exploring the stacking ensemble method with SVM as the base classifier to explore and evaluate the overall potential. The GBDT method had the highest training and validation AUC value of 0.986 and 0.940, respectively.

Hu et al. [134] proposed a novel ensemble method which is a hybridization of Bagging and RSS-based NB-tree named BRSSNBTree for the generation of LSM. The proposed ensemble method was compared with traditional ML models such as SVM and RF and found to have better performance. The BRSSNBTree ensemble model had the highest AUC value of 0.968.

Kutlug et al. [47] tested the performance of the DT-based ensemble models, namely CCF and RotFor, for the generation of LSM. The proposed ensemble models were compared with other models such as RF, bagging, and AdaBoost. The AUC value for all the ensemble models were CCF (0.932), RF (0.931), RotFor (0.925), bagging (0.90), AdaBoost (0.899). Based on the AUC value, the CCF and RotFor could predict landslide susceptibility accurately.

Hu et al. [54] investigated different ensemble learning techniques such as bagging, boosting, and stacking for landslide susceptibility mapping. The stacking model using C4.5 and ANN as base classifiers produced the best modeling robustness. The AUC values for all the ensemble models are boosting-C4.5 (0.945), boosting-ANN (0.903), stacking C4.5-ANN (0.900), bagging-ANN (0.892), bagging-C4.5 (0.878).

Fang et al. [59] presented a new ensemble technique by combining different DTs and RotFor for predicting landslide susceptibility. The selected DTs are alternating decision tree, forest by penalizing attributes (FPA), functional tree, logistic model tree, and Hoeffding tree (VFDT). The LSM was generated using the combination of all the DT classifications using RotFor. The new ensemble technique of DT and RotFor performed better than most popular ensemble models. The AUC values of different models explored for the study are ADTree (0.871), FPA (0.858), FT (0.779), LMT (0.884), and VFDT (0.892) for the single models. The AUC values for the ensemble methods are ADTree-RotFor (0.903), FPA-RotFor (0.907), FT-RotFor (0.900), LMT-RotFor (0.896), and VFDT-RotFor (0.907).

Zhang et al. [73] aim was to explore the implementation of a class-weighted algorithm with LGR and ensemble-based lightGBM and RF algorithm for landslide susceptibility mapping. The class-weighted method was used to handle imbalanced data issues of landslide and non-landslide samples. The imbalanced data issue was converted to cost-sensitive ML by setting unequal weights for both samples, improving the accuracy of LSM generated. The AUC value of weighted-RF (0.913) was the highest indicating that the class-weighted method with RF is effective for landslide susceptibility evaluation.

Kavzoglu et al. [129] compared NGBoost with RF and XGBoost used in landslide susceptibility mapping. Shapley additive explanations based on the game theory approach were used to understand the influence of LCFs on the LSM generated. Multi-collinearity was also employed to analyze the LCFs. The result of the study indicates that NGBoost had the best prediction accuracy with an AUC value of 0.898.

Zhou et al. [140] proposes a novel interpretable model based on the Shapley additive explanation and XGBoost to interpret landslides susceptibility evaluation at global and local levels. The proposed model delivered 0.75 in accuracy and an AUC value of 0.83 for

the test sets. The authors also found that peak rainfall intensity and elevation are the most significant LCFs in the study location.

Zhang et al. [141] used RF and XGBoost ensemble method to predict slope stability for Yunyang County, Chongqing, China considering twelve LCFs. The prediction performance of the ensemble methods was compared with SVM and LGR. The XGBoost, RF, LGR, and SVM test accuracy was 0.905, 0.911, 0.886, and 0.886, respectively. According to the authors, ensemble methods used in the study effectively captures the slope status and can be extended to other landslide-prone locations.

5.3.2. Discussion on Ensemble Techniques

An ensemble algorithm combines conventional ML models as base classifiers with different voting systems to produce better predictions. With the ensemble approach, the limitation or bias of one conventional ML algorithm will be compensated by another method, thereby producing a more accurate prediction. Which why it is a popular approach for generating LSM, among other applications. Some of the popular ensemble approaches used in landslide susceptibility mapping include lightGBM, RF, NGBoost, XGBoost, bagging, cascade generalization, dagging, DECORATE, MultiBoost, AdaBoost, real AdaBoost, RotFor, RSS, stacking, blending, simple averaging, weighted averaging, qualitative matrix, semi-quantitative partition, GBDT, CCF, median of probability, committee averaging, weighted mean probability, ADTree, ERT, LBE, BRT, and CHAID. From the results of various studies, the ensemble techniques consistently outperformed the conventional ML models in terms of prediction accuracy. However, it can be limited by time-consuming parameter tuning. Nevertheless, this can be remedied by using grid search, random search, Bayesian optimization, evolutionary optimization, and gradient-based optimization for parameter tuning to accelerate the model training process using the ensemble approach [132]. A hybrid-ensemble technique using other feature optimization techniques can also be employed for feature tuning and selection to improve the overall process. Due to the observed benefits discussed in the section, future studies can concentrate on hybrid-ensemble-based landslide susceptibility mapping.

5.4. Deep Learning Methods

Deep learning methods outperform conventional ML methods and are advancing in fields that have eluded the ML community for years. DL methods can produce exceptional results in sciences, engineering, and business. It can achieve high performance because of its ability to find intricate structures in high-dimensional data [156]. DL methods are representation-learning techniques with multiple layers of representation. The representation layers of DL include batch normalization, flatten, dense, and dropout layer. The batch normalization layer is used to provide consistent data distribution for training. The flattening layer flattens two- or three-dimensional training data to one-dimensional data. The dense layer predicts the class. The dropout neurons in the dropout layer help improve the model's generalization. The activation function is also an essential component in the DL architecture. The value of the activation function should be differentiable and selected as per the model's application. A commonly used activation function for DL-based landslide susceptibility mapping is the sigmoid activation function mainly because landslide susceptibility mapping is a non-linear [155].

Some popular DL methods employed to generate LSM in landslide studies include CNN, deep neural network (DNN), RNN, and LSTM. ANN was considered a conventional ML method because of its popularity in landslide susceptibility studies. CNN comprises three essential components: convolutional, down-sampling, and fully connected layers. In CNN, the convolutional layer consists of multiple kernels to extract feature information from the previous layer. The shared weight strategy in the convolutional layer allows training with fewer parameters compared to a fully connected network. The activation function follows the convolutional layer and Rectified linear unit function is the most popular and effective activation function. The down-sampling layer or pooling layer is

used to reduce the size of features, and the over-fitting tendency of the model [60]. DNN is a popular method employed in various natural hazard-related studies. It consists of several fully connected layers, dropout layers, hyperparameters for random search, activation function, and optimization algorithm [157]. RNN is another DL model with great success in the temporal data processing. Unlike CNN, it can process sequential data using recurrent hidden states, which can learn useful information from the previous and the current state [151]. However, RNN is not accurate in long data sequences. LSTM uses an advanced, recurrent structure to overcome the limitation of the RNN model. The basic building block of LSTM is the memory cell, and the information stored in the cell is controlled by gates. The input gate regulates the cell status. The forget gate decides how much of the previous memory state is retained or forgotten. The output gate regulates the information passed on to the next layer [155].

5.4.1. Literature on Deep Learning Methods

A total of nine DL-based articles were found in this survey. We shall discuss how the authors of various studies employed DL to generate LSM. As ANN is already discussed in Section 5.1, ANN is excluded in this section. Generally, DL is ANN with a large number of layers. Table 5 presents an overview of the literature on DL-based LSM studies.

Table 5. Literature on deep learning-based landslide susceptibility mapping.

Author	Year	Deep Learning Method
Wang et al. [158]	2019	CNN
Nhu et al. [159]	2020	DNN
Ngo et al. [62]	2021	RNN and CNN
Kavzoglu et al. [155]	2021	CNN, RNN, and LSTM
Azarafza et al. [160]	2021	Deep convolutional neural network (CNN-DNN)
Li et al. [151]	2021	CNN and RNN
Liu et al. [60]	2022	CNN
Wei et al. [128]	2022	DL framework SR-ML
Habumugisha et al. [157]	2022	CNN, DNN, LSTM, and RNN

Wang et al. [158] introduced CNN for landslide susceptibility mapping for the first time. Three different data representation algorithms (1D, 2D, and 3D) were developed to construct three different CNN architectures. Sixteen LCFs were considered, and the historical landslide data of Yanshan County, China, were randomly divided into training and validation in the ratio of 70:30 for the experimentation. The proposed CNNs were evaluated using OA, MCC, ROC, and AUC measures. The authors have also compared the performance with ML methods, namely optimized SVM, DNN, and LeNet-5. The CNN-2D achieved the highest AUC value of 0.813. From the experimentation result, the authors suggest that CNN is more practical for landslide prevention and management than the conventional ML methods.

Ngo et al. [62] compared DL models, namely RNN and CNN, using Iran as the case study location. In order to generate LSM, the entire ROI was converted to a raster format. Trained RNN and CNN were tested by calculating susceptibility indices of each pixel. The indices were then converted to LSM, and the landslide susceptibility was classified into five categories. The AUC value for the RNN model was 0.88 and 0.85 for CNN.

Kavzoglu et al. [155] objective was to address the limitations of DL-based landslide susceptibility mapping, such as insufficient model variance and limited generalizability. The authors proposed a network architecture that consists of two main parts. The first part consist of RNN-CNN-LSTM layer blocks meant for extracting features, and the second part includes the dense and dropout layers for prediction of the class. The trial and error method was employed to find the best batch size, number of layers, epoch, loss function, and optimizer. The proposed model improved the overall accuracy of DL models by up to 7%. The AUC value also shows a 4% improvement in the susceptibility map generated.

The AUC value of the models were CNN-RNN-LSTM (0.93), CNN (0.92), RNN (0.91), and LSTM (0.86).

Azarafza et al. [160] did a comparative study on hybrid CNN-DNN based LSM generation with traditional ML techniques. In their proposed hybrid setup, CNN was used for feature extraction and DNN to sort pixels into susceptibility groups. The CNN-DNN could achieve an AUC value of 0.909. The authors also stated the challenges in the generation of LSM using ML. The limited number of reference landslide in recorded data made modeling challenging. The triggering factors are highly dependent on the spatial resolution, and DEM data quality directly affects the quality of the input data. Furthermore, the models required extensive processing power to manage the inputs during landslide assessment.

Li et al. [151] explored a stacking-based ensemble method using CNN and RNN to generate LSM. The CNN and RNN methods were trained to generate new features and present a detailed description in the training step. Furthermore, the new validation and training dataset was used to build the final landslide susceptibility for the meta-learning step. The AUC values of the models were CNN-RNN (0.918), CNN (0.904), RNN (0.900), and LR (0.877).

Liu et al. [60] did a comparative study of CNN with conventional ML models. The LSM generated by the CNN-based model is sensitive to the high-risk landslide zone and significantly reduces the salt-and-pepper effect, which guarantees the consistency of susceptibility assessment. However, CNN consumed more time to train and predict despite its significant results. Based on the results, CNN-based models exhibit the best predictive capability for LSM on the testing datasets. A few limitations of the study include a smaller scale of the study, and re-sampling of DTM from 5m x 5m to 30m x 30m may have caused a loss of spatial information.

Wei et al. [128] developed a DL framework that integrates spatial response features and ML classifiers (SR-ML). The method includes three steps. In the first step, depthwise separable convolution was used to extract spatial features to prevent confusion of multi-factor features. In the second step, spatial pyramid pooling extracts response features to obtain features under various scales. In the third step, the high-level features are merged into ML classifiers for more effective feature classification. The AUC values of various models explored were SR-LGR (0.903), SR-SVM (0.915), and SR-RF (0.920).

Habumugisha et al. [157] explored four DL methods, namely CNN, DNN, LSTM, and RNN. The study used one-dimensional architecture for the CNN model, which consisted of one convolution layer, one pooling layer, one flattening layer, and two fully connected dense layers. The dropout layer was used to avoid overfitting. Keras tuner library was used to select the number of filters, kernel size, activation function, and the number of neurons in the fully connected layers. The DNN consists of five fully connected layers with dropout layers in the study. A sigmoid activation function was used, and the Adam optimization algorithm trained the model. The study used a bi-directional LSTM. The architecture of LSTM consists of one bi-directional layer, a dropout layer, and two fully coupled layers. The dense layer used sigmoid as the activation function, and the Adam optimization algorithm trained the model. The RNN model was set up using a simple RNN layer using the Keras library in the study. The architecture consists of one RNN layer, the dropout layer, and two fully connected layers. The AUC values for all the DL models were DNN (87.30), LSTM (86.50), CNN (85.60), and RNN (82.90).

5.4.2. Discussion on Deep Learning Methods

DL methods have outperformed traditional ML models in accurately predicting landslide susceptibility. DL-based algorithms are used in hybrid setup for feature selections, as base classifiers in an ensemble setup, and for classification of landslide susceptibility zones in a standalone setup. Popular DL methods include CNN, RNN, DNN, and LSTM. Algorithms such as quick propagation, batch back-propagation, Levenberg–Marquardt, and conjugate gradient descent are used to train the DL models. The limitations of DL

models include extensive training time and processing power requirements. A trade-off between the accuracy of landslide susceptibility and time and resource requirements must be decided to remedy the limitation.

Based on the articles surveyed, DL models can accurately predict landslide susceptibility with less uncertainty. However, DL models have low model variance and limited generalization capabilities [155]; these issues can be overcome by using hybrid and ensemble setups. Hybrid and ensemble of DL models also increase the overall accuracy of the base DL models [155,160].

6. Discussion

This article provides a comprehensive survey of ML-based landslide susceptibility mapping studies using the ROSES technique. We have discussed the different types of ML techniques employed by different studies. This survey also covers the essential elements of ML-based LSM generation, such as LCFs, landslide inventory and datasets, ML models, and evaluation techniques. We have provided the current trends, including the number of articles published per year, the number of ML model names used in the title of the papers, the most preferred journals, most cited articles, and the countries with the most case study locations. Advanced ML techniques such as hybrid, ensemble, and DL methods were also discussed in this article.

The current trend indicates a growing interest in ML-based landslide susceptibility mapping. The number of papers published per year is gradually increasing over the years. Before 2019 the average number of papers per year was less than ten, 2020 had 26 papers, and 2021 had 41 papers. The substantial increase in papers is due to easily accessible landslide inventory, satellite data, advancements in ML techniques, and GIS tools. The choice of the journal for publishing the articles is diverse. According to our survey statistics, *Geocarto International journal* is the most popular choice, with 12 article publications, followed by *Environmental Earth Science* and *Remote Sensing*. To determine the most popular ML models, we have used the name of ML models in the title of the article. Many authors choose to put the specific ML model name used in the study, while some choose to put the blanket term 'Machine Learning' in the title. Based on the survey, 35 articles have 'Machine Learning', and 25 articles have 'Support Vector Machine', which is the highest amongst the specific ML names appearing in the title. According to the statistics, SVM can be considered the most popular ML model. The study on LSM spanned different countries; with 35 articles, China was the most studied location. Turkey, Iran, India, and South Korea were the other popular case study locations.

LCFs are the primary factors responsible for causing landslide. In landslide susceptibility mapping, ML models are used to find the relationship between the occurrence of landslide and LCFs. Once the relationship is determined susceptibility index or LSM of the study location can be generated. Many LCFs are required to find the relationship. On average, 10 to 20 causative factors are commonly used. Popular LCFs include the degree of slope, aspect of slope, plan curvature, lithology, distance from river, TWI, and LULC. The choice of LCFs depends on the case study location. Identification of LCFs that directly influence the landslide is essential. Including irrelevant LCFs will degrade the performance. There is no standard guideline for selecting the LCFs. Many studies employed different optimization, feature selection, and hybrid techniques to find the LCFs with the most correlations.

Landslide datasets and inventory maps are essential in training the ML models. Popular methods to obtain landslide inventory and datasets include historical data, aerial photograph interpretation, Google earth image interpretation, satellite imagery, and field survey is the most popular method. The size of the dataset is essential. If sufficient data or landslide points are not available, oversampling techniques such as SMOTE may be used. We need balanced data of landslide points and non-landslide points in the preparation of the dataset. The landslide points are often collected from historical data or field surveys, and the non-landslide points are generated randomly. Random selection of non-landslide

points has limitations [42] it is suggested to use a systematic approach such as clustering, and ML, among other techniques. The positioning method used to represent the landslide point impacts the accuracy of LSM, as the LCFs are derived by using these points. Researchers have different observations in this regard. Polygon data, data point at the center of the landslide, and landslide scrap are recommended by the researchers. The data splitting pattern is also an essential factor. Researchers have split the data based on events (such as earthquakes), time range, and the ratio of the available dataset. It is recommended to split the dataset in 70:30 for training and testing. However, if the objective is to study the effect of an event such as a natural disaster, the data splitting may be performed before and after the event for training and testing the ML model.

ML methods are advanced algorithms that can be used for finding complex relationships between LCFs and the occurrence of landslide. The ML methods can be classified into conventional, hybrid, ensemble, and DL techniques. The conventional models are popular ML models such as RF, SVM, LGR, ANN, and NB. RF is an ensemble of decision trees, and it can achieve high prediction accuracy by taking average or votes of weak base classifiers' results. Since RF is based on many base classification trees, it does not have over-fitting issues and can automatically handle outliers and missing values. SVM is the most popular ML model used to generate LSM. The basic structure of SVM includes separating hyper-plane, maximum-margin hyper-plan, the soft margin, and the kernel function. Researchers have explored different SVM kernel functions; Gaussian is commonly used. SVM is effective in the generation of LSM and has the least false-positive rate, which is essential in risk assessment application [99]. To further improve the prediction accuracy, hybridization of SVM with optimization techniques was also explored [40]. LGR is a popular statistical method in earth science and is often used in ensemble setup as a meta-classifier. LGR is a multivariate method. In landslide studies, the aim is to find how independent variables, i.e., LCFs, are related to the dependent variable (existence landslide). Many researchers have used LGR in hybrid and ensemble setups. ANN consists of different layers of interconnected neurons. The optimum value of the interconnection weights is determined in the training phase. Researchers have explored different ANN architectures, transfer functions, and training algorithms. Sigmoid and RBF transfer function and back-propagation training algorithms are popular in landslide studies. ANN trained with a good quality dataset can accurately predict LSM. Naive Bayes is based on Bayes' rule. NB has less variance, is computationally efficient, robust, and can handle missing values. NB models have good prediction accuracy and success rate with an average AUC above 80%. NB classifier is effective in generating LSM. All the conventional ML methods provide acceptable prediction accuracy and are often used in hybrid setups as a classifier, ensemble setups as base classifiers, or as a benchmark to compare novel techniques for the generation of LSM.

Hybrid methods combine multiple ML models or different ML models with optimization and feature selection techniques. Hybrid models can identify LCFs with high co-relation to the occurrence of landslide. The model trained with only the selected features has significantly higher predictive capabilities and reduced training complexity. Researchers have also combined oversampling techniques with ML models to produce more landslide data points improving the landslide inventory and dataset artificially, increasing the generalization and predictive capability of the ML models. A total of 23 hybrid-based LSM generation articles were identified in this survey. The hybrid setups used in different landslide studies include spatial response features, Geo-detectors, PSO, ensemble method, GeoSOM, SMOTE, fractal dimension with IoE, GWR, DL, and RS theory with ML models. Based on the results from different studies, there is a noticeable improvement in the prediction accuracy of ML models. This is because hybrid methods can overcome the limitations of conventional ML models.

The ensemble method combines several weak ML models, where a voting or averaging system derives the result of the model. This method can reduce the bias of one ML model by compensating it for the other models. A total of 31 ensemble-based LSM generation

articles were identified in this survey. Some well-known ensemble techniques include lightGBM, RF, NGBoost, XGBoost, bagging, cascade generalization, dagging, DECORATE, MultiBoost, AdaBoost, real AdaBoost, RotFor, RSS, stacking, blending, simple averaging, weighted averaging, qualitative matrix, semi-quantitative partition, GBDT, CCF, median of probability, committee averaging, weighted mean probability, ADTree, ERT, LBE, BRT, and CHAID. From the experimentation result of the articles surveyed, ensemble techniques outperform the conventional ML models in terms of prediction accuracy. Training the ensemble model is time-consuming because it involves training the base models and tuning the parameters/weights involved in combining them. Optimization or searching techniques can be employed to overcome the limitation.

Deep learning is an ML model based on ANN. DL techniques require high computation power and a large training dataset. Due to technical advancements in computational devices, easily accessible landslide inventories, and high-resolution satellite data, recently, DL was applied for the generation of LSM. A total of nine DL-based landslide susceptibility mapping articles are identified in this survey. Popular DL models include CNN, RNN, DNN, and LSTM. Training algorithms such as quick propagation, batch back-propagation, Levenberg–Marquardt, and conjugate gradient descent are used to train the DL models. Based on the articles' results, DL models have higher prediction accuracy than the conventional ML models.

Evaluation methods are used to determine the prediction accuracy of ML models in landslide susceptibility mapping. Popular evaluation methods include AUC, ACC, κ, and RMSE, among other methods. AUC value can represent the summary of overall performance [79], it is the most popular method and is used in almost all the articles. An AUC value greater than 0.70 can be considered to have good accuracy. The summary of the AUC achieved by different categories of ML techniques is shown in Table 6.

Table 6. AUC Performance of machine learning techniques.

ML Techniques	0.70–0.80	0.80–0.90	>0.90
Conventional	12	31	20
Hybrid	1	15	14
Ensemble	4	40	22
Deep Learning	-	6	6

It is observed that conventional techniques are the most used ML model for the generation of LSM because they are also used as the baseline model to prove the superiority of the proposed model in many studies. Deep learning models are the least explored because they are new in landslide studies. Approximately 32% of the conventional, 47% of hybrid, 33.3% of the ensemble, and 50% of deep learning have AUC over 90%. The AUC values in the table indicate that deep learning, hybrid, and ensemble models can generate LSM accurately.

7. Conclusions and Future Scope of Work

Landslide is a natural disaster that disrupts human lives. Scientists and researchers have explored various techniques to minimize the effect of landslide. One of the techniques is generating landslide susceptibility maps, which aid disaster management planners. The methods used to generate LSM include statistical, knowledge-based, and ML techniques. This literature survey provides a comprehensive overview of ML-based landslide susceptibility mapping. We have selected 119 articles for the survey using ROSES protocol to ensure reproducibility and transparency. The trend indicates a growing interest in the field, increasing the number of articles published per year. The statistics signify that SVM is frequently used to generate LSM, the *Geocarto International* journal is a popular choice of authors, and China is the most studied location. This paper also introduces the essential elements in landslide susceptibility studies, including landslide causative factors, datasets and landslide-inventory maps, ML models, and evaluation methods. The commonly used

LCFs are TWI, slope degree, slope aspect, distance to the river, plan curvature, and LULC. Field survey is the most popular method for collecting landslide data. AUC is considered the best evaluation method to assess the performance of ML models in landslide susceptible mapping and is used in almost all the articles. Researchers explored a wide variety of ML models, and they are grouped into conventional, hybrid, ensemble, and deep learning techniques. Popular conventional ML methods include RF, SVM, NB, ANN, and LGR. All the ML models considered in this survey can generate acceptable LSM. Many of them have an AUC value greater than 0.90 and can generate highly reliable LSM. There are studies where the conventional ML models produced highly accurate predictions. Generally, hybrid, ensemble, and DL techniques produced more accurate LSM than conventional techniques.

Based on the survey followings are a few recommendations that researchers may adopt to generate accurate LSM using ML models:

1. There is no standard guideline for selecting LCFs, and their importance differs from one study location to another. Researchers can collect all the available LCFs for the study location, use a suitable feature selection method to determine the important LCFs, and use only the highly co-related LCFs in landslide susceptibility mapping;
2. A large number of high-resolution datasets are required to train the ML models. If sufficient landslide data are not available, researchers can use oversampling techniques to increase the size of the dataset. Landslide and an equal amount of non-landslide data are also required for training and validation. Systematic approaches such as clustering techniques can be used to select the non-landslide locations;
3. Combination of different ML models in a hybrid or ensemble setup can overcome the limitations of standalone ML models and achieve higher accuracy;
4. For evaluating the performance of the landslide susceptibility models, AUC can be used, as the AUC value represents the summary of overall performance [79].

Technology is fast evolving; each passing day, new ML models or improved versions of the existing ML models are proposed by researchers. The following few least explored ML technologies in landslide susceptibility studies that have the potential to generate reliable LSM can be investigated in future work:

1. The Transformer DL model is a sequence-to-sequence model and state-of-the-art in machine translation. It uses self-attention mechanism to decide which part of the input sequence is important in translating to the output sequence. The RNN and LSTM were also designed for sequence-to-sequence translation tasks and have achieved high accuracy in landslide susceptibility mapping;
2. Transfer learning (TL) is a popular ML method. It reuses the previously trained model on a new problem, and the idea is to use the knowledge learned for one task to solve similar ones. The advantages of TL include less training time and high prediction accuracy with a small dataset. The ML models trained on a study location with a rich dataset can be applied in other study locations with a small dataset using transfer learning.

The limitations of this survey are limited keywords and using only the WoS database to search articles on landslide susceptibility studies. Future literature surveys can consider keywords not included in this survey and more academic research databases such as Scopus.

Author Contributions: Conceptualization, M.A. and K.A.; methodology, M.A. and K.A.; validation, A.K.M., E.J., R.G., Z.L. and M.J.; writing—original draft preparation, M.A. and K.A.; writing—review and editing, M.A., K.A., A.K.M., E.J., R.G., Z.L. and M.J.; supervision, K.A.; funding acquisition, A.K.M., E.J., R.G., Z.L. and M.J. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

References

1. Turner, A.K.; Schuster, R.L. *Landslides: Investigation and Mitigation*; Special Report 247; National Academy Press: Washington, DC, USA, 1996.
2. Fleuchaus, P.; Blum, P.; Wilde, M.; Terhorst, B.; Butscher, C. Retrospective evaluation of landslide susceptibility maps and review of validation practice. *Environ. Earth Sci.* **2021**, *80*, 485. [CrossRef]
3. Pardeshi, S.D.; Autade, S.E.; Pardeshi, S.S. Landslide hazard assessment: Recent trends and techniques. *SpringerPlus* **2013**, *2*, 1–11. [CrossRef] [PubMed]
4. Huabin, W.; Gangjun, L.; Weiya, X.; Gonghui, W. GIS-based landslide hazard assessment: An overview. *Prog. Phys. Geogr.* **2005**, *29*, 548–567. [CrossRef]
5. 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]
6. Arabameri, A.; Rezaei, K.; Pourghasemi, H.R.; Lee, S.; Yamani, M. GIS-based gully erosion susceptibility mapping: A comparison among three data-driven models and AHP knowledge-based technique. *Environ. Earth Sci.* **2018**, *77*, 628. [CrossRef]
7. Merghadi, 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]
8. Huang, Y.; Zhao, L. Review on landslide susceptibility mapping using support vector machines. *Catena* **2018**, *165*, 520–529. [CrossRef]
9. Naemitabar, M.; Zanganeh Asadi, M. Landslide zonation and assessment of Farizi watershed in northeastern Iran using data mining techniques. *Nat. Hazards* **2021**, *108*, 2423–2453. [CrossRef]
10. Zhang, T.; Fu, Q.; Wang, H.; Liu, F.; Wang, H.; Han, L. Bagging-based machine learning algorithms for landslide susceptibility modeling. *Nat. Hazards* **2021**, *110*, 823–846. [CrossRef]
11. Clarivate. Web of Science: Discover Multidisciplinary Content. Available online: https://www.webofscience.com/wos/woscc/basic-search (accessed on 2 March 2022).
12. Saha, S.; Saha, A.; Hembram, T.K.; Pradhan, B.; Alamri, A.M. Evaluating the performance of individual and novel ensemble of machine learning and statistical models for landslide susceptibility assessment at Rudraprayag District of Garhwal Himalaya. *Appl. Sci.* **2020**, *10*, 3772. [CrossRef]
13. Fang, Z.; Wang, Y.; Peng, L.; Hong, H. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. *Comput. Geosci.* **2020**, *139*, 104470. [CrossRef]
14. Pham, B.T.; Bui, D.T.; Prakash, I.; Dholakia, M. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *Catena* **2017**, *149*, 52–63. [CrossRef]
15. Haddaway, N.R.; Macura, B.; Whaley, P.; Pullin, A.S. ROSES RepOrting standards for Systematic Evidence Syntheses: pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. *Environ. Evid.* **2018**, *7*, 7. [CrossRef]
16. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* **2009**, *151*, 264–269. [CrossRef] [PubMed]
17. Pradhan, B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput. Geosci.* **2013**, *51*, 350–365. [CrossRef]
18. Yilmaz, I. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat—Turkey). *Comput. Geosci.* **2009**, *35*, 1125–1138. [CrossRef]
19. Yilmaz, I. Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environ. Earth Sci.* **2010**, *61*, 821–836. [CrossRef]
20. 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]
21. Yao, X.; Tham, L.; Dai, F. Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of Hong Kong, China. *Geomorphology* **2008**, *101*, 572–582. [CrossRef]
22. Zare, M.; Pourghasemi, H.R.; Vafakhah, M.; Pradhan, B. Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: A comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arab. J. Geosci.* **2013**, *6*, 2873–2888. [CrossRef]
23. Park, S.; Choi, C.; Kim, B.; Kim, J. Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea. *Environ. Earth Sci.* **2013**, *68*, 1443–1464. [CrossRef]

24. Pourghasemi, H.R.; Jirandeh, A.G.; Pradhan, B.; Xu, C.; Gokceoglu, C. Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran. *J. Earth Syst. Sci.* **2013**, *122*, 349–369. [CrossRef]

25. 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). *Geomat. Nat. Hazards Risk* **2018**, *9*, 49–69. [CrossRef]

26. Koukis, G.; Ziourkas, C. Slope instability phenomena in Greece: A statistical analysis. *Bull. Int. Assoc. Eng. Geol.-Bull. L'Assoc. Int. Géol. L'Ing.* **1991**, *43*, 47–60. [CrossRef]

27. Kavzoglu, T.; Colkesen, I.; Sahin, E.K. Machine learning techniques in landslide susceptibility mapping: A survey and a case study. In *Landslides: Theory, Practice and Modelling*; Springer: Cham, Switzerland, 2019; pp. 283–301.

28. Chen, W.; Chen, Y.; Tsangaratos, P.; Ilia, I.; Wang, X. Combining evolutionary algorithms and machine learning models in landslide susceptibility assessments. *Remote Sens.* **2020**, *12*, 3854. [CrossRef]

29. Lee, S.; Min, K. Statistical analysis of landslide susceptibility at Yongin, Korea. *Environ. Geol.* **2001**, *40*, 1095–1113. [CrossRef]

30. Ercanoglu, M.; Gokceoglu, C. Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey). *Eng. Geol.* **2004**, *75*, 229–250. [CrossRef]

31. Pourghasemi, H.R.; Pradhan, B.; Gokceoglu, C. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Nat. Hazards* **2012**, *63*, 965–996. [CrossRef]

32. Ercanoglu, M.; Gokceoglu, C. Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzzy approach. *Environ. Geol.* **2002**, *41*, 720–730.

33. Survey, P.U.G. Geologic Map of South Asia (geo8ag). Available online: https://catalog.data.gov/dataset/geologic-map-of-south-asia-geo8ag (accessed on 2 March 2022).

34. Gökceoglu, C.; Aksoy, H. Landslide susceptibility mapping of the slopes in the residual soils of the Mengen region (Turkey) by deterministic stability analyses and image processing techniques. *Eng. Geol.* **1996**, *44*, 147–161. [CrossRef]

35. Kirkby, M.; Beven, K. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. J.* **1979**, *24*, 43–69.

36. Moore, I.D.; Grayson, R.; Ladson, A. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrol. Process.* **1991**, *5*, 3–30. [CrossRef]

37. Youssef, A.M.; Pourghasemi, H.R. Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. *Geosci. Front.* **2021**, *12*, 639–655. [CrossRef]

38. Kadavi, P.R.; Lee, C.W.; Lee, S. Application of ensemble-based machine learning models to landslide susceptibility mapping. *Remote Sens.* **2018**, *10*, 1252. [CrossRef]

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

40. Yu, X.; Wang, Y.; Niu, R.; Hu, Y. A combination of geographically weighted regression, particle swarm optimization and support vector machine for landslide susceptibility mapping: a case study at Wanzhou in the Three Gorges Area, China. *Int. J. Environ. Res. Public Health* **2016**, *13*, 487. [CrossRef]

41. Peng, L.; Niu, R.; Huang, B.; Wu, X.; Zhao, Y.; Ye, R. Landslide susceptibility mapping based on rough set theory and support vector machines: A case of the Three Gorges area, China. *Geomorphology* **2014**, *204*, 287–301. [CrossRef]

42. Nhu, V.H.; Shirzadi, A.; Shahabi, H.; Singh, S.K.; Al-Ansari, N.; Clague, J.J.; Jaafari, A.; Chen, W.; Miraki, S.; Dou, J.; 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. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2749. [CrossRef]

43. Abraham, M.T.; Satyam, N.; Lokesh, R.; Pradhan, B.; Alamri, A. Factors Affecting Landslide Susceptibility Mapping: Assessing the Influence of Different Machine Learning Approaches, Sampling Strategies and Data Splitting. *Land* **2021**, *10*, 989. [CrossRef]

44. Wang, Y.; Sun, D.; Wen, H.; Zhang, H.; Zhang, F. Comparison of random forest model and frequency ratio model for landslide susceptibility mapping (LSM) in Yunyang County (Chongqing, China). *Int. J. Environ. Res. Public Health* **2020**, *17*, 4206. [CrossRef]

45. Saha, S.; Roy, J.; Hembram, T.K.; Pradhan, B.; Dikshit, A.; Abdul Maulud, K.N.; Alamri, A.M. Comparison between Deep Learning and Tree-Based Machine Learning Approaches for Landslide Susceptibility Mapping. *Water* **2021**, *13*, 2664. [CrossRef]

46. Li, R.; Wang, N. Landslide susceptibility mapping for the Muchuan county (China): A comparison between bivariate statistical models (woe, ebf, and ioe) and their ensembles with logistic regression. *Symmetry* **2019**, *11*, 762. [CrossRef]

47. Kutlug Sahin, E.; Colkesen, I. Performance analysis of advanced decision tree-based ensemble learning algorithms for landslide susceptibility mapping. *Geocarto Int.* **2021**, *36*, 1253–1275. [CrossRef]

48. Hu, Q.; Zhou, Y.; Wang, S.; Wang, F. Machine learning and fractal theory models for landslide susceptibility mapping: Case study from the Jinsha River Basin. *Geomorphology* **2020**, *351*, 106975. [CrossRef]

49. Hu, X.; Zhang, H.; Mei, H.; Xiao, D.; Li, Y.; Li, M. Landslide susceptibility mapping using the stacking ensemble machine learning method in Lushui, Southwest China. *Appl. Sci.* **2020**, *10*, 4016. [CrossRef]

50. Pham, B.T.; Vu, V.D.; Costache, R.; Phong, T.V.; Ngo, T.Q.; Tran, T.H.; Nguyen, H.D.; Amiri, M.; Tan, M.T.; Trinh, P.T.; et al. Landslide susceptibility mapping using state-of-the-art machine learning ensembles. *Geocarto Int.* **2021**, 1–26. doi: 10.1080/10106049.2021.1914746. [CrossRef]

51. Zhou, X.; Wen, H.; Zhang, Y.; Xu, J.; Zhang, W. Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. *Geosci. Front.* **2021**, *12*, 101211. [CrossRef]

52. Meng, Q.; Miao, F.; Zhen, J.; Wang, X.; Wang, A.; Peng, Y.; Fan, Q. GIS-based landslide susceptibility mapping with logistic regression, analytical hierarchy process, and combined fuzzy and support vector machine methods: A case study from Wolong Giant Panda Natural Reserve, China. *Bull. Eng. Geol. Environ.* **2016**, *75*, 923–944. [CrossRef]

53. 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]

54. Hu, X.; Mei, H.; Zhang, H.; Li, Y.; Li, M. Performance evaluation of ensemble learning techniques for landslide susceptibility mapping at the Jinping county, Southwest China. *Nat. Hazards* **2021**, *105*, 1663–1689. [CrossRef]

55. Wang, G.; Lei, X.; Chen, W.; Shahabi, H.; Shirzadi, A. Hybrid computational intelligence methods for landslide susceptibility mapping. *Symmetry* **2020**, *12*, 325. [CrossRef]

56. Gong, W.; Hu, M.; Zhang, Y.; Tang, H.; Liu, D.; Song, Q. GIS-based landslide susceptibility mapping using ensemble methods for Fengjie County in the Three Gorges Reservoir Region, China. *Int. J. Environ. Sci. Technol.* **2021**, 1–18. [CrossRef]

57. Song, J.; Wang, Y.; Fang, Z.; Peng, L.; Hong, H. Potential of ensemble learning to improve tree-based classifiers for landslide susceptibility mapping. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4642–4662. [CrossRef]

58. Fang, Z.; Wang, Y.; Peng, L.; Hong, H. A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 321–347. [CrossRef]

59. Fang, Z.; Wang, Y.; Duan, G.; Peng, L. Landslide susceptibility mapping using rotation forest ensemble technique with different decision trees in the Three Gorges Reservoir area, China. *Remote Sens.* **2021**, *13*, 238. [CrossRef]

60. Liu, R.; Yang, X.; Xu, C.; Wei, L.; Zeng, X. Comparative Study of Convolutional Neural Network and Conventional Machine Learning Methods for Landslide Susceptibility Mapping. *Remote Sens.* **2022**, *14*, 321. [CrossRef]

61. Liu, Y.; Zhang, W.; Zhang, Z.; Xu, Q.; Li, W. Risk Factor Detection and Landslide Susceptibility Mapping Using Geo-Detector and Random Forest Models: The 2018 Hokkaido Eastern Iburi Earthquake. *Remote Sens.* **2021**, *13*, 1157. [CrossRef]

62. Ngo, P.T.T.; Panahi, M.; Khosravi, K.; Ghorbanzadeh, O.; Kariminejad, N.; Cerda, A.; Lee, S. Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geosci. Front.* **2021**, *12*, 505–519.

63. GSI. Bhukosh | Gateway to All Geoscientific Information of GSI. Available online: https://bhukosh.gsi.gov.in/Bhukosh/MapViewer.aspx (accessed on 2 March 2022).

64. Soldati, M.; Corsini, A.; Pasuto, A. Landslides and climate change in the Italian Dolomites since the Late glacial. *Catena* **2004**, *55*, 141–161. [CrossRef]

65. Wood, J.; Harrison, S.; Reinhardt, L. Landslide inventories for climate impacts research in the European Alps. *Geomorphology* **2015**, *228*, 398–408. [CrossRef]

66. Brardinoni, F.; Slaymaker, O.; Hassan, M.A. Landslide inventory in a rugged forested watershed: A comparison between air-photo and field survey data. *Geomorphology* **2003**, *54*, 179–196. [CrossRef]

67. Fisher, G.B.; Amos, C.B.; Bookhagen, B.; Burbank, D.W.; Godard, V.; Whitmeyer, S. Channel widths, landslides, faults, and beyond: The new world order of high-spatial resolution Google Earth imagery in the study of earth surface processes. *Geol. Soc. Am. Spec. Pap.* **2012**, *492*, 1–22.

68. Sato, H.; Harp, E. Interpretation of earthquake-induced landslides triggered by the 12 May 2008, M7.9 Wenchuan earthquake in the Beichuan area, Sichuan Province, China using satellite imagery and Google Earth. *Landslides* **2009**, *6*, 153–159. [CrossRef]

69. Guerriero, L.; Di Martire, D.; Calcaterra, D.; Francioni, M. Digital Image Correlation of Google Earth Images for Earth's Surface Displacement Estimation. *Remote Sens.* **2020**, *12*, 3518. [CrossRef]

70. Wang, Y.; Wu, X.; Chen, Z.; Ren, F.; Feng, L.; Du, Q. Optimizing the predictive ability of machine learning methods for landslide susceptibility mapping using SMOTE for Lishui City in Zhejiang Province, China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 368. [CrossRef]

71. Liang, Z.; Wang, C.; Duan, Z.; Liu, H.; Liu, X.; Ullah Jan Khan, K. A hybrid model consisting of supervised and unsupervised learning for landslide susceptibility mapping. *Remote Sens.* **2021**, *13*, 1464. [CrossRef]

72. Liang, Z.; Wang, C.; Khan, K.U.J. Application and comparison of different ensemble learning machines combining with a novel sampling strategy for shallow landslide susceptibility mapping. *Stoch. Environ. Res. Risk Assess.* **2021**, *35*, 1243–1256. [CrossRef]

73. Zhang, H.; Song, Y.; Xu, S.; He, Y.; Li, Z.; Yu, X.; Liang, Y.; Wu, W.; Wang, Y. Combining a class-weighted algorithm and machine learning models in landslide susceptibility mapping: A case study of Wanzhou section of the Three Gorges Reservoir, China. *Comput. Geosci.* **2022**, *158*, 104966. [CrossRef]

74. Rong, G.; Alu, S.; Li, K.; Su, Y.; Zhang, J.; Zhang, Y.; Li, T. Rainfall Induced Landslide Susceptibility Mapping Based on Bayesian Optimized Random Forest and Gradient Boosting Decision Tree Models—A Case Study of Shuicheng County, China. *Water* **2020**, *12*, 3066. [CrossRef]

75. Yilmaz, I. The effect of the sampling strategies on the landslide susceptibility mapping by conditional probability and artificial neural networks. *Environ. Earth Sci.* **2010**, *60*, 505–519. [CrossRef]

76. Sahin, E.K.; Colkesen, I.; Kavzoglu, T. A comparative assessment of canonical correlation forest, random forest, rotation forest and logistic regression methods for landslide susceptibility mapping. *Geocarto Int.* **2020**, *35*, 341–363. [CrossRef]

77. Karakas, G.; Kocaman, S.; Gokceoglu, C. Comprehensive performance assessment of landslide susceptibility mapping with MLP and random forest: A case study after Elazig earthquake (24 Jan 2020, Mw 6.8), Turkey. *Environ. Earth Sci.* **2022**, *81*, 1–22. [CrossRef]

78. Narkhede, S. Understanding auc-roc curve. *Towards Data Sci.* **2018**, *26*, 220–227.
79. Xie, W.; Li, X.; Jian, W.; Yang, Y.; Liu, H.; Robledo, L.F.; Nie, W. A novel hybrid method for landslide susceptibility mapping-based geodetector and machine learning cluster: A case of Xiaojin county, China. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 93. [CrossRef]
80. Pham, B.T.; Prakash, I.; Singh, S.K.; Shirzadi, A.; Shahabi, H.; Tran, T.T.T.; Bui, D.T. Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. *Catena* **2019**, *175*, 203–218. [CrossRef]
81. García, S.; Fernández, A.; Luengo, J.; Herrera, F. A study of statistical techniques and performance measures for genetics-based machine learning: Accuracy and interpretability. *Soft Comput.* **2009**, *13*, 959. [CrossRef]
82. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [CrossRef]
83. 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]
84. Akinci, H.; Zeybek, M. Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin), Turkey. *Nat. Hazards* **2021**, *108*, 1515 – 1543. [CrossRef]
85. Ali, M.Z.; Chu, H.J.; Chen, Y.C.; Ullah, S. Machine learning in earthquake-and typhoon-triggered landslide susceptibility mapping and critical factor identification. *Environ. Earth Sci.* **2021**, *80*, 233. [CrossRef]
86. Pham, B.T.; Prakash, I.; Chen, W.; Ly, H.B.; Ho, L.S.; Omidvar, E.; Tran, V.P.; Bui, D.T.; et al. A novel intelligence approach of a sequential minimal optimization-based support vector machine for landslide susceptibility mapping. *Sustainability* **2019**, *11*, 6323. [CrossRef]
87. 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]
88. Liu, Y.; Wang, Y.; Zhang, J. New Machine Learning Algorithm: Random Forest. In Proceedings of the Information Computing and Applications, Chengde, China, 14–16 September 2012; Liu, B., Ma, M., Chang, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 246–252.
89. Liu, L.L.; Yang, C.; Wang, X.M. Landslide susceptibility assessment using feature selection-based machine learning models. *Geomech. Eng.* **2021**, *25*, 1–16.
90. Taalab, K.; Cheng, T.; Zhang, Y. Mapping landslide susceptibility and types using Random Forest. *Big Earth Data* **2018**, *2*, 159–178. Available online: http://xxx.lanl.gov/abs/doi:10.1080/20964471.2018.1472392 (accessed on 11 March 2022). [CrossRef]
91. Pham, Q.B.; Achour, Y.; Ali, S.A.; Parvin, F.; Vojtek, M.; Vojteková, J.; Al-Ansari, N.; Achu, A.L.; Costache, R.; Khedher, K.M.; et al. A comparison among fuzzy multi-criteria decision making, bivariate, multivariate and machine learning models in landslide susceptibility mapping. *Geomat. Nat. Hazards Risk* **2021**, *12*, 1741–1777. Available online: http://xxx.lanl.gov/abs/doi:10.1080/19475705.2021.1944330 (accessed on 11 March 2022). [CrossRef]
92. Zhao, Z.; Liu, Z.Y.; Xu, C. Slope unit-based landslide susceptibility mapping using certainty factor, support vector machine, random forest, CF-SVM and CF-RF models. *Front. Earth Sci.* **2021**, *9*, 589630. [CrossRef]
93. Micheletti, N.; Foresti, L.; Robert, S.; Leuenberger, M.; Pedrazzini, A.; Jaboyedoff, M.; Kanevski, M. Machine learning feature selection methods for landslide susceptibility mapping. *Math. Geosci.* **2014**, *46*, 33–57. [CrossRef]
94. Belousov, A.; Verzakov, S.; Von Frese, J. Applicational aspects of support vector machines. *J. Chemom. J. Chemom. Soc.* **2002**, *16*, 482–489. [CrossRef]
95. Noble, W.S. What is a support vector machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [CrossRef]
96. Chen, W.; Chai, H.; Zhao, Z.; Wang, Q.; Hong, H. Landslide susceptibility mapping based on GIS and support vector machine models for the Qianyang County, China. *Environ. Earth Sci.* **2016**, *75*, 1–13. [CrossRef]
97. 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. [CrossRef]
98. Oh, H.J.; Kadavi, P.R.; Lee, C.W.; Lee, S. Evaluation of landslide susceptibility mapping by evidential belief function, logistic regression and support vector machine models. *Geomat. Nat. Hazards Risk* **2018**, *9*, 1053–1070. [CrossRef]
99. Marjanović, M.; Kovačević, M.; Bajat, B.; Voženílek, V. Landslide susceptibility assessment using SVM machine learning algorithm. *Eng. Geol.* **2011**, *123*, 225–234. [CrossRef]
100. Dreiseitl, S.; Ohno-Machado, L. Logistic regression and artificial neural network classification models: A methodology review. *J. Biomed. Inform.* **2002**, *35*, 352–359. [CrossRef]
101. Subasi, A. Chapter 3—Machine learning techniques. In *Practical Machine Learning for Data Analysis Using Python*; Subasi, A., Ed.; Academic Press: Cambridge, MA, USA, 2020; pp. 91–202. [CrossRef]
102. Belyadi, H.; Haghighat, A. Chapter 5—Supervised learning. In *Machine Learning Guide for Oil and Gas Using Python*; Belyadi, H., Haghighat, A., Eds.; Gulf Professional Publishing: Oxford, UK, 2021; pp. 169–295. [CrossRef]
103. Bartosik, A.; Whittingham, H. Chapter 7—Evaluating safety and toxicity. In *The Era of Artificial Intelligence, Machine Learning, and Data Science in the Pharmaceutical Industry*; Ashenden, S.K., Ed.; Academic Press: Cambridge, MA, USA, 2021; pp. 119–137. [CrossRef]
104. Li, Y.; Chen, W. Landslide Susceptibility Evaluation Using Hybrid Integration of Evidential Belief Function and Machine Learning Techniques. *Water* **2020**, *12*, 113. [CrossRef]

105. Wang, S.C., Artificial Neural Network. In *Interdisciplinary Computing in Java Programming*; Springer: Boston, MA, USA, 2003; pp. 81–100.

106. Choi, J.; Oh, H.J.; Won, J.S.; Lee, S. Validation of an artificial neural network model for landslide susceptibility mapping. *Environ. Earth Sci.* **2010**, *60*, 473–483. [CrossRef]

107. Lucchese, L.V.; de Oliveira, G.G.; Pedrollo, O.C. Mamdani fuzzy inference systems and artificial neural networks for landslide susceptibility mapping. *Nat. Hazards* **2021**, *106*, 2381–2405. [CrossRef]

108. Yilmaz, I. A case study from Koyulhisar (Sivas-Turkey) for landslide susceptibility mapping by artificial neural networks. *Bull. Eng. Geol. Environ.* **2009**, *68*, 297–306. [CrossRef]

109. Quan, H.C.; Lee, B.G. GIS-based landslide susceptibility mapping using analytic hierarchy process and artificial neural network in Jeju (Korea). *KSCE J. Civ. Eng.* **2012**, *16*, 1258–1266. [CrossRef]

110. Wang, Q.; Li, W.; Xing, M.; Wu, Y.; Pei, Y.; Yang, D.; Bai, H. Landslide susceptibility mapping at Gongliu county, China using artificial neural network and weight of evidence models. *Geosci. J.* **2016**, *20*, 705–718. [CrossRef]

111. Can, A.; Dagdelenler, G.; Ercanoglu, M.; Sonmez, H. Landslide susceptibility mapping at Ovacık-Karabük (Turkey) using different artificial neural network models: Comparison of training algorithms. *Bull. Eng. Geol. Environ.* **2019**, *78*, 89–102. [CrossRef]

112. Harmouzi, H.; Nefeslioglu, H.A.; Rouai, M.; Sezer, E.A.; Dekayir, A.; Gokceoglu, C. Landslide susceptibility mapping of the Mediterranean coastal zone of Morocco between Oued Laou and El Jebha using artificial neural networks (ANN). *Arab. J. Geosci.* **2019**, *12*, 1–18. [CrossRef]

113. Webb, G.I.; Keogh, E.; Miikkulainen, R. Naïve Bayes. *Encycl. Mach. Learn.* **2010**, *15*, 713–714.

114. Lee, S.; Lee, M.J.; Jung, H.S.; Lee, S. Landslide susceptibility mapping using naïve bayes and bayesian network models in Umyeonsan, Korea. *Geocarto Int.* **2020**, *35*, 1665–1679. [CrossRef]

115. Imtiaz, I.; Umar, M.; Latif, M.; Ahmed, R.; Azam, M. Landslide susceptibility mapping: Improvements in variable weights estimation through machine learning algorithms—A case study of upper Indus River Basin, Pakistan. *Environ. Earth Sci.* **2022**, *81*, 1–14. [CrossRef]

116. Chen, W.; Zhao, X.; Shahabi, H.; Shirzadi, A.; Khosravi, K.; Chai, H.; Zhang, S.; Zhang, L.; Ma, J.; Chen, Y.; et al. Spatial prediction of landslide susceptibility by combining evidential belief function, logistic regression and logistic model tree. *Geocarto Int.* **2019**, *34*, 1177–1201. [CrossRef]

117. Zhang, T.y.; Han, L.; Zhang, H.; Zhao, Y.h.; Li, X.a.; Zhao, L. GIS-based landslide susceptibility mapping using hybrid integration approaches of fractal dimension with index of entropy and support vector machine. *J. Mt. Sci.* **2019**, *16*, 1275–1288. [CrossRef]

118. Adnan, M.S.G.; Rahman, M.S.; Ahmed, N.; Ahmed, B.; Rabbi, M.; Rahman, R.M.; et al. Improving spatial agreement in machine learning-based landslide susceptibility mapping. *Remote Sens.* **2020**, *12*, 3347. [CrossRef]

119. Wang, Y.; Feng, L.; Li, S.; Ren, F.; Du, Q. A hybrid model considering spatial heterogeneity for landslide susceptibility mapping in Zhejiang Province, China. *Catena* **2020**, *188*, 104425. [CrossRef]

120. Sahana, M.; Pham, B.T.; Shukla, M.; Costache, R.; Thu, D.X.; Chakrabortty, R.; Satyam, N.; Nguyen, H.D.; Phong, T.V.; Le, H.V.; et al. Rainfall induced landslide susceptibility mapping using novel hybrid soft computing methods based on multi-layer perceptron neural network classifier. *Geocarto Int.* **2020**, 1–25. doi: 10.1080/10106049.2020.1837262. [CrossRef]

121. Alqadhi, S.; Mallick, J.; Talukdar, S.; Bindajam, A.A.; Saha, T.K.; Ahmed, M.; Khan, R.A. Combining Logistic Regression-based hybrid optimized machine learning algorithms with sensitivity analysis to achieve robust landslide susceptibility mapping. *Geocarto Int.* **2021**, 1–25. [CrossRef]

122. Arabameri, A.; Chandra Pal, S.; Rezaie, F.; Chakrabortty, R.; Saha, A.; Blaschke, T.; Di Napoli, M.; Ghorbanzadeh, O.; Thi Ngo, P.T. Decision tree based ensemble machine learning approaches for landslide susceptibility mapping. *Geocarto Int.* **2021**, 1–35. [CrossRef]

123. Saha, S.; Roy, J.; Pradhan, B.; Hembram, T.K. Hybrid ensemble machine learning approaches for landslide susceptibility mapping using different sampling ratios at East Sikkim Himalayan, India. *Adv. Space Res.* **2021**, *68*, 2819–2840. [CrossRef]

124. Xing, Y.; Yue, J.; Guo, Z.; Chen, Y.; Hu, J.; Travé, A. Large-scale landslide susceptibility mapping using an integrated machine learning model: A case study in the Lvliang mountains of China. *Front. Earth Sci.* **2021**, *9*, 622. [CrossRef]

125. Hu, H.; Wang, C.; Liang, Z.; Gao, R.; Li, B. Exploring Complementary Models Consisting of Machine Learning Algorithms for Landslide Susceptibility Mapping. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 639. [CrossRef]

126. Sun, D.; Shi, S.; Wen, H.; Xu, J.; Zhou, X.; Wu, J. A hybrid optimization method of factor screening predicated on GeoDetector and Random Forest for Landslide Susceptibility Mapping. *Geomorphology* **2021**, *379*, 107623. [CrossRef]

127. Dung, N.V.; Hieu, N.; Phong, T.V.; Amiri, M.; Costache, R.; Al-Ansari, N.; Prakash, I.; Le, H.V.; Nguyen, H.B.T.; Pham, B.T. Exploring novel hybrid soft computing models for landslide susceptibility mapping in Son La hydropower reservoir basin. *Geomat. Nat. Hazards Risk* **2021**, *12*, 1688–1714. [CrossRef]

128. Wei, R.; Ye, C.; Sui, T.; Ge, Y.; Li, Y.; Li, J. Combining spatial response features and machine learning classifiers for landslide susceptibility mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *107*, 102681. [CrossRef]

129. Kavzoglu, T.; Teke, A. Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). *Arab. J. Sci. Eng.* **2022**, *47*, 7367–7385. [CrossRef]

130. Nhu, V.H.; Shirzadi, A.; Shahabi, H.; Chen, W.; Clague, J.J.; Geertsema, M.; Jaafari, A.; Avand, M.; Miraki, S.; Talebpour Asl, D.; et al. Shallow landslide susceptibility mapping by random forest base classifier and its ensembles in a semi-arid region of Iran. *Forests* **2020**, *11*, 421. [CrossRef]

131. Kalantar, B.; Ueda, N.; Saeidi, V.; Ahmadi, K.; Halin, A.A.; Shabani, F. Landslide susceptibility mapping: Machine and ensemble learning based on remote sensing big data. *Remote Sens.* **2020**, *12*, 1737. [CrossRef]

132. Pham, B.T.; Van Phong, T.; Nguyen-Thoi, T.; Trinh, P.T.; Tran, Q.C.; Ho, L.S.; Singh, S.K.; Duyen, T.T.T.; Nguyen, L.T.; Le, H.Q.; et al. GIS-based ensemble soft computing models for landslide susceptibility mapping. *Adv. Space Res.* **2020**, *66*, 1303–1320. [CrossRef]

133. Rodriguez, J.J.; Kuncheva, L.I.; Alonso, C.J. Rotation forest: A new classifier ensemble method. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1619–1630. [CrossRef] [PubMed]

134. Hu, X.; Huang, C.; Mei, H.; Zhang, H. Landslide susceptibility mapping using an ensemble model of Bagging scheme and random subspace–based naïve Bayes tree in Zigui County of the Three Gorges Reservoir Area, China. *Bull. Eng. Geol. Environ.* **2021**, *80*, 5315–5329. [CrossRef]

135. Pham, B.T.; Nguyen-Thoi, T.; Qi, C.; Van Phong, T.; Dou, J.; Ho, L.S.; Van Le, H.; Prakash, I. Coupling RBF neural network with ensemble learning techniques for landslide susceptibility mapping. *Catena* **2020**, *195*, 104805. [CrossRef]

136. Ho, T.K. The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 832–844.

137. Nhu, V.H.; Mohammadi, A.; Shahabi, H.; Ahmad, B.B.; Al-Ansari, N.; Shirzadi, A.; Clague, J.J.; Jaafari, A.; Chen, W.; Nguyen, H. Landslide susceptibility mapping using machine learning algorithms and remote sensing data in a tropical environment. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4933. [CrossRef] [PubMed]

138. Althuwaynee, O.F.; Pradhan, B.; Park, H.J.; Lee, J.H. A novel ensemble decision tree-based CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic regression models in landslide susceptibility mapping. *Landslides* **2014**, *11*, 1063–1078. [CrossRef]

139. Duan, T.; Anand, A.; Ding, D.Y.; Thai, K.K.; Basu, S.; Ng, A.; Schuler, A. Ngboost: Natural gradient boosting for probabilistic prediction. In Proceedings of the International Conference on Machine Learning, Virtual, 13–18 July 2020; pp. 2690–2700.

140. Zhou, X.; Wen, H.; Li, Z.; Zhang, H.; Zhang, W. An interpretable model for the susceptibility of rainfall-induced shallow landslides based on SHAP and XGBoost. *Geocarto Int.* **2022**, 1–27. [CrossRef]

141. Zhang, W.; Li, H.; Han, L.; Chen, L.; Wang, L. Slope stability prediction using ensemble learning techniques: A case study in Yunyang County, Chongqing, China. *J. Rock Mech. Geotech. Eng.* 2022, *in press*. [CrossRef]

142. Webb, G.I. Multiboosting: A technique for combining boosting and wagging. *Mach. Learn.* **2000**, *40*, 159–196. [CrossRef]

143. Schapire, R.E.; Singer, Y. Improved boosting algorithms using confidence-rated predictions. *Mach. Learn.* **1999**, *37*, 297–336. [CrossRef]

144. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]

145. Pham, B.T.; Prakash, I. Evaluation and comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, logistic regression and support vector machines methods for landslide susceptibility mapping. *Geocarto Int.* **2019**, *34*, 316–333. [CrossRef]

146. Arabameri, A.; Pradhan, B.; Rezaei, K.; Sohrabi, M.; Kalantari, Z. GIS-based landslide susceptibility mapping using numerical risk factor bivariate model and its ensemble with linear multivariate regression and boosted regression tree algorithms. *J. Mt. Sci.* **2019**, *16*, 595–618. [CrossRef]

147. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]

148. Tien Bui, D.; Ho, T.C.; Pradhan, B.; Pham, B.T.; Nhu, V.H.; Revhaug, I. GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with AdaBoost, Bagging, and MultiBoost ensemble frameworks. *Environ. Earth Sci.* **2016**, *75*, 1–22. [CrossRef]

149. Ting, K.M.; Witten, I.H. *Stacking Bagged and Dagged Models*; University of Waikato: Hamilton, New Zealand, 1997.

150. Melville, P.; Mooney, R.J. Creating diversity in ensembles using artificial data. *Inf. Fusion* **2005**, *6*, 99–111. [CrossRef]

151. Li, W.; Fang, Z.; Wang, Y. Stacking ensemble of deep learning methods for landslide susceptibility mapping in the Three Gorges Reservoir area, China. *Stoch. Environ. Res. Risk Assess.* **2021**, 1–22. [CrossRef]

152. Töscher, A.; Jahrer, M.; Bell, R.M. The bigchaos solution to the netflix grand prize. *Netflix Prize. Doc.* **2009**, 1–52.

153. Di Napoli, M.; Carotenuto, F.; Cevasco, A.; Confuorto, P.; Di Martire, D.; Firpo, M.; Pepe, G.; Raso, E.; Calcaterra, D. Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability. *Landslides* **2020**, *17*, 1897–1914. [CrossRef]

154. 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.* **2019**, *11*, 2866. [CrossRef]

155. Kavzoglu, T.; Teke, A.; Yilmaz, E.O. Shared blocks-based ensemble deep learning for shallow landslide susceptibility mapping. *Remote Sens.* **2021**, *13*, 4776. [CrossRef]

156. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]

157. Habumugisha, J.M.; Chen, N.; Rahman, M.; Islam, M.M.; Ahmad, H.; Elbeltagi, A.; Sharma, G.; Liza, S.N.; Dewan, A. Landslide susceptibility mapping with deep learning algorithms. *Sustainability* **2022**, *14*, 1734. [CrossRef]

158. Wang, Y.; Fang, Z.; Hong, H. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. *Sci. Total Environ.* **2019**, *666*, 975–993. [CrossRef]

159. Nhu, V.H.; Hoang, N.D.; Nguyen, H.; Ngo, P.T.T.; Bui, T.T.; Hoa, P.V.; Samui, P.; Bui, D.T. Effectiveness assessment of Keras based deep learning with different robust optimization algorithms for shallow landslide susceptibility mapping at tropical area. *Catena* **2020**, *188*, 104458. [CrossRef]
160. Azarafza, M.; Azarafza, M.; Akgün, H.; Atkinson, P.M.; Derakhshani, R. Deep learning-based landslide susceptibility mapping. *Sci. Rep.* **2021**, *11*, 24112. [CrossRef]