



Article

Landslide Susceptibility Mapping Based on Deep Learning Algorithms Using Information Value Analysis Optimization

Junjie Ji ^{1,2,3} , Yongzhang Zhou ^{1,2,3,*} , Qiuming Cheng ¹, Shoujun Jiang ⁴ and Shiting Liu ⁵

¹ School of Earth Science and Engineering, Sun Yat-sen University, Zhuhai 519000, China

² Guangdong Provincial Key Laboratory of Geological Processes and Mineral Resource Survey, Guangzhou 510275, China

³ School of Earth Sciences and Engineering, Center for Earth Environment & Resources, Sun Yat-sen University, Guangzhou 510275, China

⁴ Guangdong Geological Survey Institute, Guangzhou 510275, China

⁵ The Sixth Geological Team of Guangdong Geological Bureau, Jiangmen 529000, China

* Correspondence: zhouyz@mail.sysu.edu.cn

Abstract: Selecting samples with non-landslide attributes significantly impacts the deep-learning modeling of landslide susceptibility mapping. This study presents a method of information value analysis in order to optimize the selection of negative samples used for machine learning. Recurrent neural network (RNN) has a memory function, so when using an RNN for landslide susceptibility mapping purposes, the input order of the landslide-influencing factors affects the resulting quality of the model. The information value analysis calculates the landslide-influencing factors, determines the input order of data based on the importance of any specific factor in determining the landslide susceptibility, and improves the prediction potential of recurrent neural networks. The simple recurrent unit (SRU), a newly proposed variant of the recurrent neural network, is characterized by possessing a faster processing speed and currently has less application history in landslide susceptibility mapping. This study used recurrent neural networks optimized by information value analysis for landslide susceptibility mapping in Xinhui District, Jiangmen City, Guangdong Province, China. Four models were constructed: the RNN model with optimized negative sample selection, the SRU model with optimized negative sample selection, the RNN model, and the SRU model. The results show that the RNN model with optimized negative sample selection has the best performance in terms of AUC value (0.9280), followed by the SRU model with optimized negative sample selection (0.9057), the RNN model (0.7277), and the SRU model (0.6355). In addition, several objective measures of accuracy (0.8598), recall (0.8302), F1 score (0.8544), Matthews correlation coefficient (0.7206), and the receiver operating characteristic also show that the RNN model performs the best. Therefore, the information value analysis can be used to optimize negative sample selection in landslide sensitivity mapping in order to improve the model's performance; second, SRU is a weaker method than RNN in terms of model performance.

Keywords: landslide susceptibility mapping; information value analysis; recurrent neural network; simple recurrent unit



Citation: Ji, J.; Zhou, Y.; Cheng, Q.; Jiang, S.; Liu, S. Landslide Susceptibility Mapping Based on Deep Learning Algorithms Using Information Value Analysis Optimization. *Land* **2023**, *12*, 1125. <https://doi.org/10.3390/land12061125>

Academic Editors: Stefano Morelli, Veronica Pazzi and Mirko Francioni

Received: 24 April 2023

Revised: 18 May 2023

Accepted: 22 May 2023

Published: 25 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Faced with current human societal challenges, it is more important than ever for geoscientists to use their understanding of the earth to benefit the society [1]. The most notable development in the field of mathematical geoscience in the last decade has been the introduction of big data and artificial intelligence algorithms. The ability of machine learning (ML) algorithms to handle nonlinear problems has tremendous advantages in dealing with complex geoscience problems [2–4]. As a result, ML is now being fully utilized in geoscience fields. For example, Wang et al. used unsupervised ML algorithms to identify multielement geochemical anomalies [5], and Yu et al. used hierarchical clustering,

singularity mapping, and the Kohonen neural network to identify Ag–Au–Pb–Zn poly-metallic mineralization-associated geochemical anomalies [6]. In general, we are primarily focused on geological events that have a significant impact but occur infrequently, such as earthquakes, typhoons, vein formation, and landslides.

Landslides are natural disasters that pose a serious risk to human lives and property and represent one of the most destructive categories of natural disasters that occur globally [7]. Mountainous areas are especially affected by landslides, whose controlling mechanisms are the complex geological and geographical conditions present in that landscape. Seventy percent of China's area is mountainous, providing favorable conditions for landslide occurrences, resulting in casualties and considerable economic losses [8–11]. As a consequence, landslide susceptibility mapping (LSM), which can analyze possible spatial areas for landslide occurrence, is an effective technique for land managers to mitigate the effects of landslides [12,13].

Machine learning is a subdivision of artificial intelligence (AI) that uses computer technologies to analyze and predict information by learning from the training dataset. A variety of ML methods have been used for LSM, including Bayesian networks, decision trees, support vector machines, random forests, and artificial networks [14–18]. It is to be noted that in recent years, in the implementation and development of natural hazard modelling, researchers have begun to consider the use of hybrid models. Hybrid models combine individual models with metaheuristic algorithms, allowing the hybrid model to eliminate the weak points inherent to the individual models to obtain more accurate results. For example, adaptive neuro-fuzzy system-gradient-based optimization (ANFIS-GBO) is applied to the spatial modelling of flood hazards [19]; cuckoo optimization algorithm-multi-layer perceptron (COA-MLP) and SailFish optimizer- multi-layer perceptron (SFO-MLP) approaches are applied to the landslide susceptibility assessment [20]; and ANFIS integrated three optimization algorithms (ant colony optimization (ACO), genetic algorithm (GA), and particle swarm optimization (PSO)) applied to flood susceptibility maps [21]. A variety of machine learning and deep learning models have been used to improve the accuracy of LSM. In recent years, to obtain better deep learning and machine learning models, researchers have adopted a variety of improved methods, such as the deep-learning optimization algorithm [22], the hybrid ensemble-based deep-learning framework [23], and the class-weighted algorithm combined with ML models [24].

Deep learning models have been increasingly applied in the modeling of environmental variables, such as environmental remote sensing [25], PM_{2.5} prediction [26], and water temperature prediction [27]. Recurrent neural networks (RNNs) are a specific kind of neural network that not only considers the previous moment's input but also gives the network a "memory" function for the previous content. Based on this unique function of the RNN approach, the order of data input will affect the model's effectiveness. Exploring a sequential data representation method can take advantage of the memory function of RNNs, which allows for thorough exploration of the prediction potential of RNNs. RNNs have been applied to LSM. Thi Ngo et al. applied RNN and CNN techniques for an LSM of Iran at the national scale [28]. Liming Xiao et al. used long short-term memory (LSTM) to predict landslide susceptibility along the China–Nepal Highway [29]. The common variants of RNNs are LSTM [30] and gated recurrent units (GRUs) [31]. Recently, a simple recurrent unit (SRU) was proposed as a new RNN variant that has a faster processing speed than the LSTM and GRU approaches. The use of the new RNN variant, using an SRU, has less application in LSM, and its specific performance in LSM should be further studied.

Traditional binary classifiers for machine learning usually require two sets of samples with corresponding labels, including positive and negative samples [32]. There are often imperfect cases in the practical applications, however, most commonly manifesting when only positive and unlabeled samples are used in the training dataset. For non-landslide samples, there still needs to be a specific definition and a reasonable method to obtain them. In general, the study area is divided into landslide and non-landslide areas. Furthermore, samples from non-landslide areas can be drawn randomly from non-landslide

areas. These unlabeled samples cannot be directly considered negative samples, because the areas of these samples are likely to be the only areas where disasters have not yet occurred [33]. At present, the issue of non-landslide sample selection has received some attention. Yang et al. [34] used Bayesian optimization algorithms to optimize the proportion of landslide samples. Chang et al. [35] selected non-landslide samples multiple times and investigated the uncertainty of non-landslide sample selection. Huang et al. [36] selected the non-landslide samples from the non-landslide area with a low landslide susceptibility level based on a semi-supervised multiple-layer perceptron model. Overall, there is no universally accepted method for optimizing non-landslide sample selection due to the differences in study areas and the logic and mechanisms behind different algorithms, which need to be studied thoroughly.

Therefore, the main innovation of this study is to optimize the selection of negative samples using information value analysis. Information value analysis determines the input order of the data by calculating the influence factors and fully explores the prediction potential of RNNs with memory function. In addition, SRU has been less studied on LSM, and both RNN and SRU models are constructed to explore the prediction performance of SRU through a comparative study.

2. Study Area

2.1. Description

Xinhui District, located between latitudes $22^{\circ}5'15''$ and $22^{\circ}35'01''$ N and longitudes $112^{\circ}46'55''$ and $113^{\circ}15'43''$ E, is in the south-central part of Guangdong Province (Figure 1). The land area of the region contains 1354.71 square kilometers. Mountainous areas are distributed in the northwest and southwest of the district, accounting for 35.84% of the total area of the region. Plains are distributed in the southeastern, south-central and west-central parts of the district, accounting for 43.53% of the total area of the district. The region's waters account for 20.63% of the total area of the region. Xinhui has a southern subtropical maritime monsoon climate, abundant rainfall, sufficient sunshine, and mild and humid conditions year round. The average annual temperature is 22.4°C , with the highest and lowest historical temperatures of 38.3°C and 0.1°C , respectively. The annual average precipitation is 1808.3 mm. The precipitation is concentrated from April through September. The average annual sunshine hours are 1734.1 h.

The list of landslides used in this paper, completed by the Guangdong Geological Survey Institute, consists of 178 landslides and locations of high-risk points (Figure 1), of which the landslide samples occurred from 2017 to 2020. Most of the landslides are classified as sliding landslides. All the landslides in this study can be classified as moderate ($400\text{--}1000\text{ m}^2$) and small ($<400\text{ m}^2$). In addition, there are rock landslides and earth landslides. According to the report, these landslides were triggered by rainfall events that occurred after anthropogenic activity.

2.2. Datasets

Heckmann et al. [37] stated that the increase in the samples accounted for has had a positive impact on the LSM and has increased the model's effectiveness. However, the training samples used for LSM are insufficient in many cases. To solve this problem, we collaborated with geologists to collect historical landslide points and locations with significant potential for landslides throughout the whole region, totaling 178 points. We used these points as samples to improve the effectiveness of the model.

In this study, 15 landslide influencing factors were considered, including elevation, slope, aspect, plan curvature, profile curvature, degree of relief, land use, rock type, topographic wetness index (TWI), terrain ruggedness index (TRI), topographic position index (TPI), normalized difference vegetation index (NDVI) on 15 April 2014, distance to faults, distance to rivers, and distance to roads. Detailed information on the landslide influencing factors is shown in Table 1. The following describes the preparation for each influencing factor.

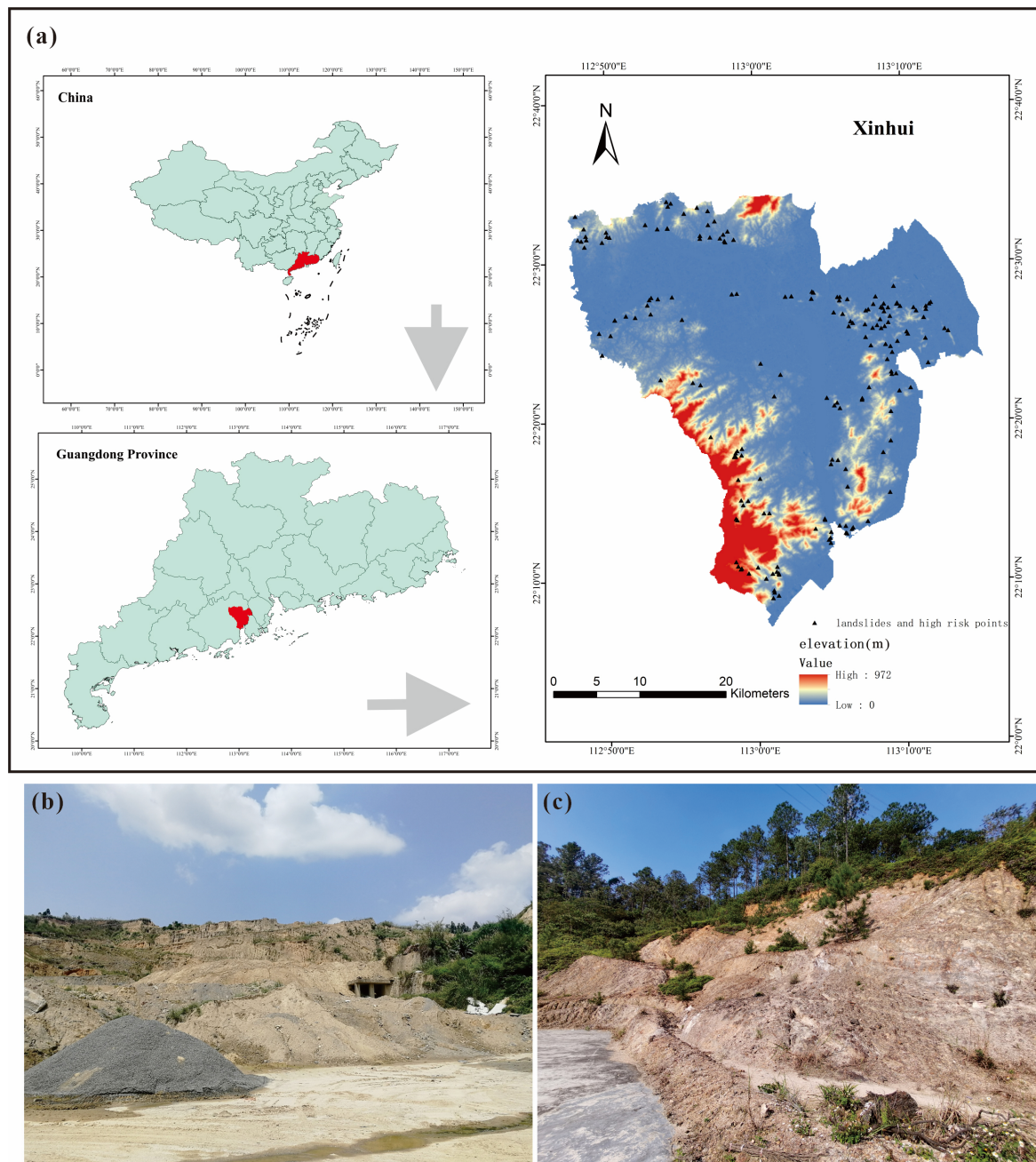


Figure 1. (a) Location of the study area; (b) and (c) are field photos.

The elevation, slope, aspect, plan curvature, profile curvature and degree of relief were extracted from a digital elevation model (DEM) obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM V2) (<http://www.gscloud.cn>, accessed on 11 March 2021). Slope, aspect, plan curvature, profile curvature, and degree of relief were calculated in the MapGIS 10.2 software. The TWI and TPI were generated by the SAGA 6.1 software. The distance to roads and the distance to rivers were produced by ArcGIS based on topographic maps at a scale of 1:50,000. The distance to faults was produced by ArcGIS based on engineering geological maps at a scale of 1:50,000. We obtained NDVI data for the study area from the USGS (<https://earthexplorer.usgs.gov>, accessed on 20 March 2021). Land use data and rock type data were provided by the collaboration with geologists. All factors were converted into a raster form with a spatial resolution of 20 m. The descriptions of these factors are shown in Table 1. Figure 2 shows the spatial distribution of these factors.

Table 1. Description of landslide factors.

Factor Type	Factors	Range
Geologic Factors	Rock type	Granite, Sandstone, Slate, Quaternary sediments and rivers
	Distance to faults (m)	(0, 6046)
	Elevation (m)	(0, 972)
Topographic Factors	Slope	(0, 49.73)
	Aspect	Flat, North, Northeast, East, Southeast, South, Southwest, West, Northwest
	Plan curvature	(0, 65.46)
	Profile curvature	(0, 11.36)
	Degree of relief	(0, 40.73)
	TRI	(0, 83.00)
	TPI	(−6.49, 10.96)
Water-Related Factors	Distance to rivers (m)	(0, 3691)
	TWI	(0, 22.68)
Anthropogenic Factors	Land use	Farmland, Forest and grass, Residential, Bare, Water
	Distance to roads (m)	(0, 2704)
Vegetation Factors	NDVI	(−1, 1)

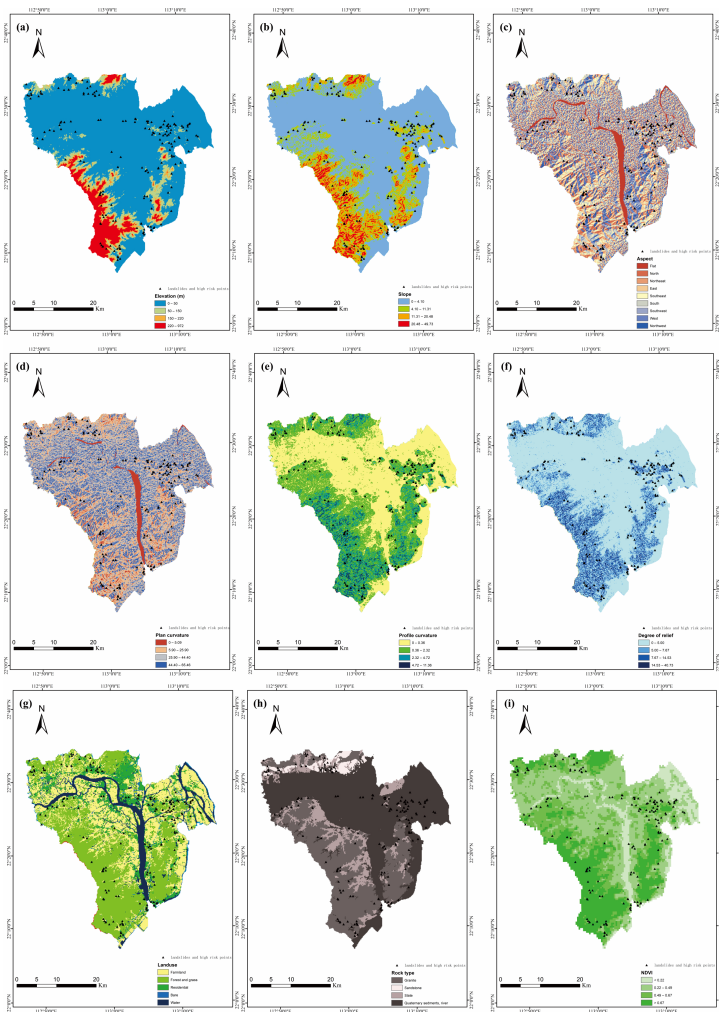


Figure 2. Cont.

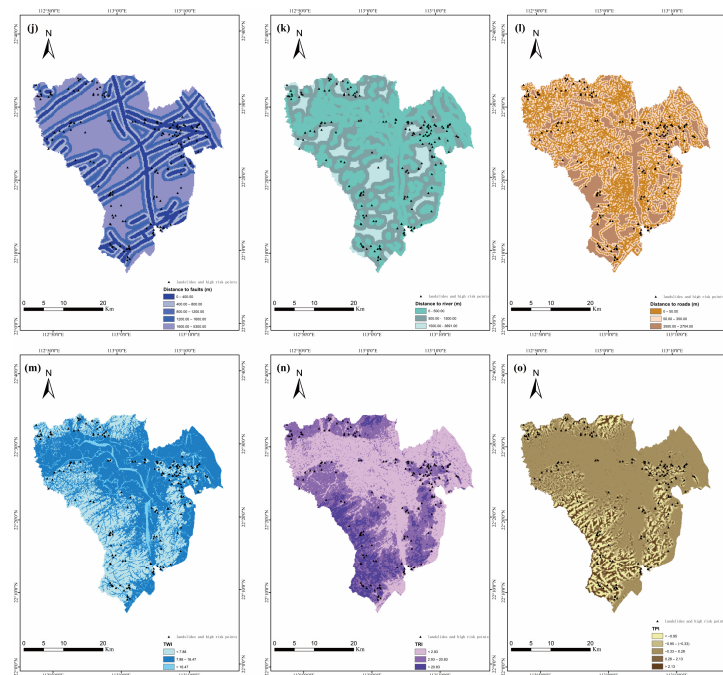


Figure 2. Spatial distribution of landslide influencing factors: (a) elevation, (b) slope, (c) aspect, (d) plan curvature, (e) profile curvature, (f) degree of relief, (g) land use, (h) rock type, (i) NDVI, (j) distance to faults, (k) distance to river, (l) distance to roads, (m) TWI, (n) TRI, and (o) TPI.

3. Materials and Methods

Figure 3 shows the process diagram used in this study. There are six steps in this process: (1) selecting the landslide influencing factors, (2) selecting typical negative samples and representing landslide data in series based on the information values (IVs), (3) preparing both the training and testing datasets by random partitioning, (4) constructing RNN and SRU models, (5) evaluating and comparing the landslide models, and (6) constructing a landslide susceptibility map.

3.1. Information Value Analysis

Information value analysis is a data exploration technique that helps determine which columns in a dataset have predictive power or influence on the value of a specified dependent variable. Information value is a very useful concept for variable selection during model building. The roots of the IVs are in the information theory that was proposed by Claude Shannon [38,39]. The IV analysis is a popular tool in the banking and bond ratings fields [40,41]. The effectiveness of landslide models can be enhanced by introducing IV into the processing of landslide factors for LSM. The correlation coefficient can be calculated as follows:

$$IV(x_i) = (n_{i1}/n_1 - n_{i0}/n_0)WOE(x_i) = (n_{i1}/n_1 - n_{i0}/n_0)\ln \frac{n_{i1}/n_1}{n_{i0}/n_0} \quad (1)$$

$$IV(x) = \sum_{i=1}^N IV(x_i) \quad (2)$$

where n_1 is the total number of landslide rasters, n_0 is the total number of non-landslide rasters, n_{i1} is the number of landslide rasters of class x_i for variable x , and n_{i0} is the number of non-landslide rasters of class x_i for variable x . In practice, the standard rule of using the IVs is shown in the Table 2.

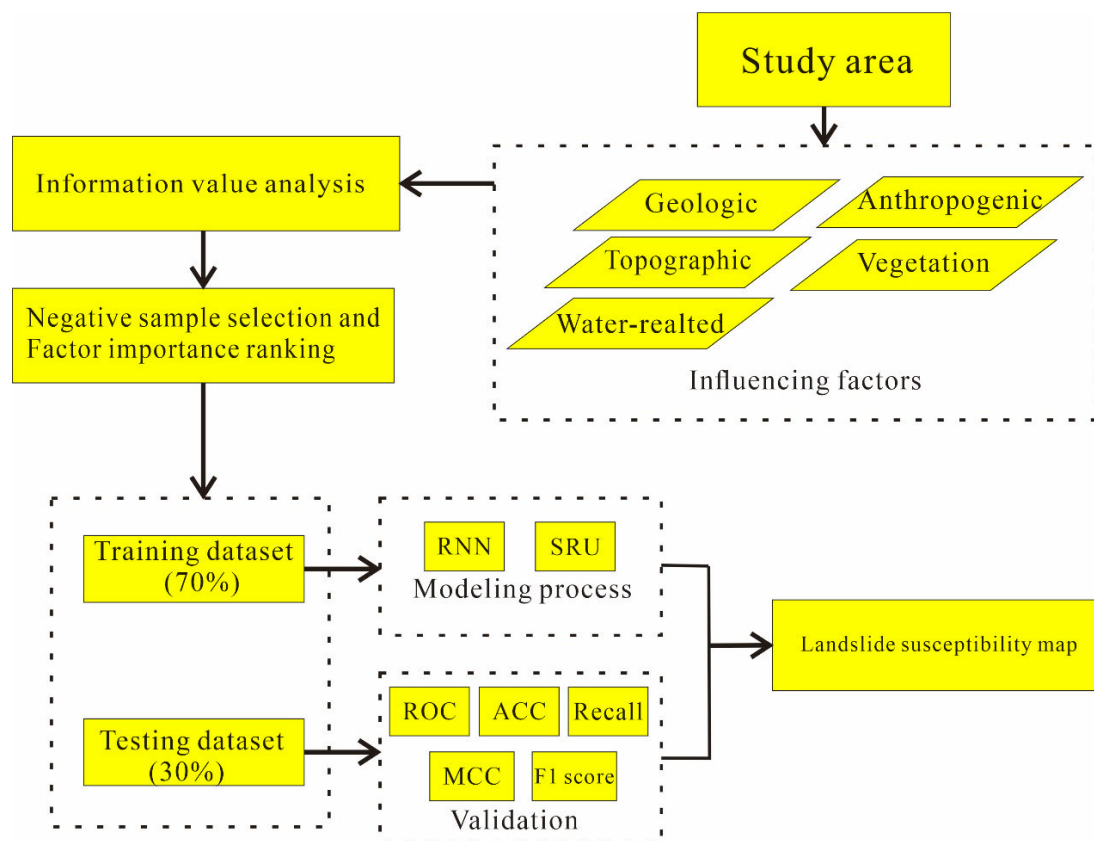


Figure 3. Methodology of the study.

Table 2. Standard rule for using the information value.

Information Value	Predictive Power
<0.02	Useless
0.02–0.1	Weak
0.1–0.3	Medium
0.3–0.5	Strong
>0.5	Suspiciously good

3.2. Recurrent Neural Network and Its Variants

3.2.1. Recurrent Neural Network

In traditional neural network models, the layers are fully connected from the input layer to the hidden layer to the output layer, and the nodes between each layer are unconnected [42,43]. Recurrent neural networks (RNNs) are a class of Artificial Neural Networks (ANNs), and RNNs are intended to be used to process sequential data (Figure 4). Specifically, the network remembers the previous information input and then applies it to the calculation of the current output. The nodes between the hidden layers are no longer connectionless but connected, and the input of the hidden layers includes not only the output of the input layer but also the output of the hidden layer at the previous moment.

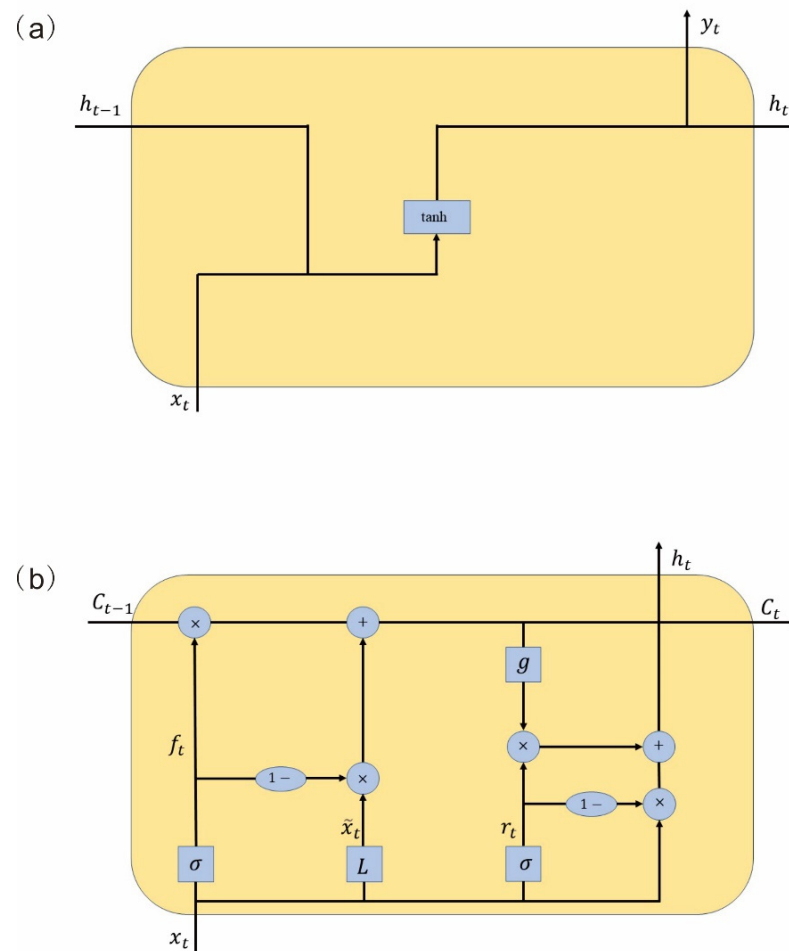


Figure 4. (a) RNN architecture and (b) SRU architecture.

Traditional recurrent neural networks are often implemented using Elman networks or Jordan networks, both of which are similar and are three-layer networks. The Elman network and the Jordan network are also known as “simple recurrent networks” (SRN) [44,45]. Let x_t , y_t , and h_t be the input vector, the output vector, and the hidden layer vector, then we can obtain

$$h_t = \sigma_h(W_h x_t + U_h h_{t-1} + b_h) \quad (3)$$

$$y_t = \sigma_y(W_y h_t + b_y) \quad (4)$$

where U and W are parameter matrices, b is the bias vector, and σ_h and σ_y are activation functions.

3.2.2. Simple Recurrent Unit

The SRU is a variant of the recently proposed RNN, and the SRU and the related work aim to propose and explore simple, fast, and more explanatory RNNs (Figure 4) [46]. Compared to other RNN variants, such as LSTM and GRU, SRU can achieve faster training speeds due to its designed structure. Figure 5 shows the basic structure of the SRU. The SRU is built on the same “gate” structure as the LSTM and GRU, but the difference is that SRU removes the limitation of parallelization of that LSTM and GRU adhere to, resulting in a much faster processing speed. The SRU has two components: “light recurrence” and “high network”. Let x_t , f_t , c_t , r_t , and h_t be the input vector, the forget gate vector, the current state from light recurrence, the reset gate vector, and the hidden layer vector. The

light recurrence can be summarized as Equations (5)–(7), and the high network can be summarized as Equations (8) and (9).

$$\tilde{x}_t = Wx_t \quad (5)$$

$$f_t = \sigma(W_f x_t + b_f) \quad (6)$$

$$C_t = f_t \odot c_{t-1} + (1 - f_t) \odot (Wx_t) \quad (7)$$

$$r_t = \sigma(W_r x_t + b_r) \quad (8)$$

$$h_t = r_t \odot g(C_t) + (1 - r_t) \odot x_t \quad (9)$$

where W and b are the parameter matrices. The value \odot is the pointwise multiplication operation [47].

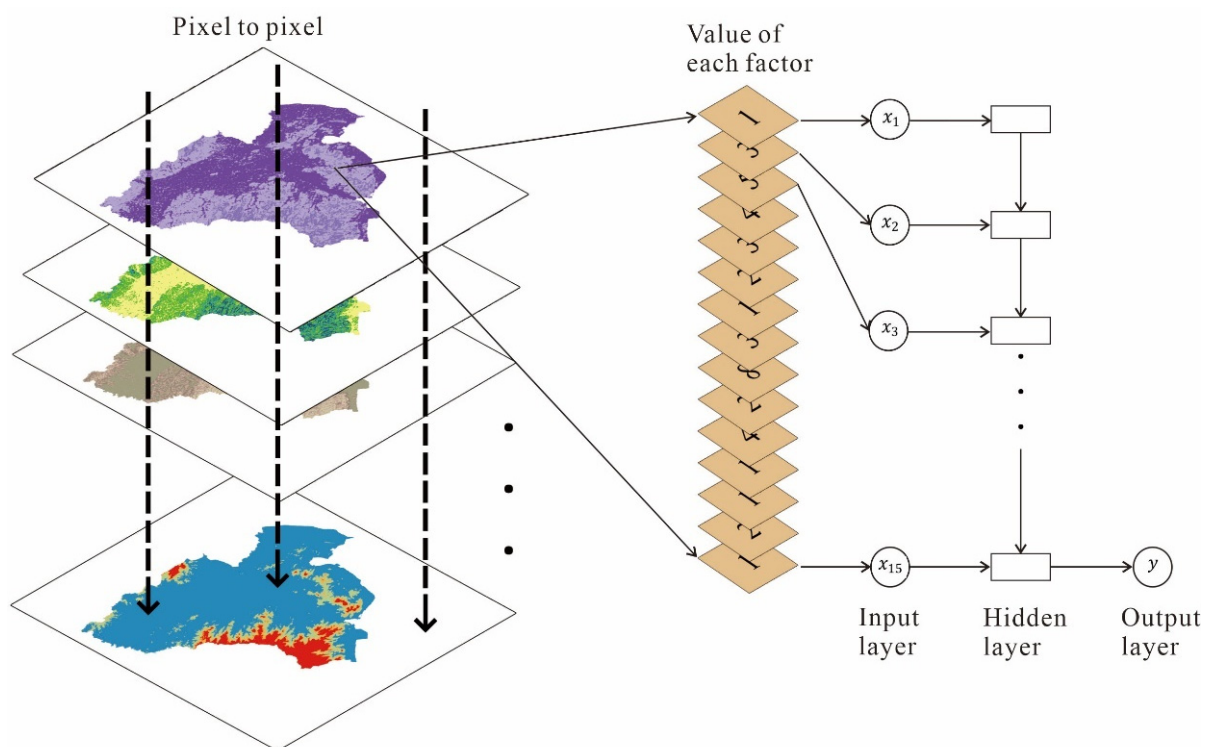


Figure 5. Data representation of models.

3.3. Selection of Landslide Influencing Factor

For LSM models, inputting more data does not necessarily result in a better model, as too much redundancy in the influencing factors considered will reduce the model's predictive capability [48]. Therefore, it is crucial to correctly select the landslide influencing factors [49]. The IV analysis method has been described above, and Table 3 shows the analysis of these influencing factors using Equations (1) and (2).

Table 3. Information value analysis of each landslide influencing factor.

Factor	Class	No. of Pixels	No. of Landslides	WOE	IV of Each Class	IV
Rock type	Granite	10,872	65	0.1029	0.0037	0.0996
	Sandstone	956	15	1.0678	0.0590	
	Slate	3832	11	−0.6308	0.0343	
	Quaternary sediments and rivers	17,340	87	−0.0724	0.0027	
Distance to faults (m)	0–400	7630	68	0.5021	0.0757	0.1031
	400–800	6770	37	0.0131	0.0000	
	800–1200	4824	16	−0.4863	0.0274	
	1200–1600	3634	22	0.1554	0.0016	
	1600–5300	10,142	35	−0.4466	0.0494	
Elevation (m)	0–50	25,588	138	−0.0001	0.0000	0.0492
	50–150	3469	27	0.3667	0.0171	
	150–220	1227	6	−0.0981	0.0003	
	220–972	2716	7	−0.7385	0.0317	
Slope	0–4.10	22,125	79	−0.4125	0.0935	0.3456
	4.10–11.31	5752	75	0.8827	0.2181	
	11.31–20.48	3531	21	0.0977	0.0011	
	20.48–49.73	1592	3	−1.0517	0.0330	
Aspect	Flat	1442	1	−2.0513	0.0781	0.1859
	North	3991	14	−0.4303	0.0182	
	Northeast	3899	17	−0.2128	0.0048	
	East	4313	18	−0.2565	0.0076	
	Southeast	4561	17	−0.3696	0.0158	
	South	3746	30	0.3952	0.0217	
	Southwest	3404	26	0.3479	0.0149	
	West	3670	29	0.3818	0.0197	
Plan curvature	Northwest	3974	26	0.1930	0.0050	0.1889
	0–5.09	2509	4	−1.2189	0.0653	
	5.09–25.90	10,913	92	0.4466	0.0831	
	25.90–44.40	11,148	42	−0.3589	0.0366	
Profile curvature	44.40–65.46	8430	40	−0.1282	0.0039	0.6907
	0–0.36	16,291	24	−1.2978	0.4657	
	0.36–2.32	12,123	125	0.6479	0.2170	
	2.32–4.72	3855	26	0.2234	0.0065	
Degree of relief	4.72–11.36	731	3	−0.2733	0.0014	0.0994
	0–5.00	25,944	115	−0.1963	0.0275	
	5.00–7.67	2721	25	0.5326	0.0309	
	7.67–14.53	3272	30	0.5305	0.0368	
	14.53–40.73	1063	8	0.3331	0.0042	

Table 3. Cont.

Factor	Class	No. of Pixels	No. of Landslides	WOE	IV of Each Class	IV
TRI	<2.93	16,246	19	−1.5287	0.5894	0.8827
	2.93–20.83	13,470	145	0.6910	0.2808	
	>20.83	3284	12	−0.3894	0.0125	
TPI	<−0.95	3392	21	0.1378	0.0021	0.4914
	(−0.95)–0.33	4460	58	0.8800	0.1678	
	0.33–0.28	18,924	46	−0.7971	0.2511	
	0.28–2.13	4648	45	0.5849	0.0655	
	>20.83	1576	6	−0.3484	0.0049	
Distance to rivers (m)	0–500	17,364	116	0.2139	0.0268	0.0985
	500–1500	12,263	55	−0.1845	0.0116	
	1,500–3691	3373	7	−0.9552	0.0601	
TWI	<7.88	11,184	108	0.5824	0.1560	0.3207
	7.88–16.47	20,184	68	−0.4707	0.1081	
	>16.47	1632	2	−1.4819	0.0566	
Land use	Farmland	11,517	35	−0.5738	0.0874	0.2750
	Forest and grass	13,492	95	0.2665	0.0333	
	Residential	3981	39	0.5968	0.0588	
	Bare	646	4	0.1380	0.0004	
	Water	3364	5	−1.2890	0.0952	
Distance to roads (m)	0–50	16,797	100	0.0987	0.0052	0.1372
	50–350	9534	64	0.2187	0.0155	
	350–2704	6669	14	−0.9437	0.1165	
NDVI	<0.22	3356	3	−1.7974	0.1525	0.2058
	0.22–0.49	12,749	61	−0.1199	0.0052	
	0.49–0.67	8974	66	0.3100	0.0306	
	>0.67	7921	32	−0.2891	0.0174	

Table 2 shows the standard rule of using the IV analysis. All IVs are higher than 0.02, indicating that all influencing factors have certain predictive power for the occurrence of landslides. Based on the above results, the TRI has the highest IV of 0.8827, indicating that it may be the dominant factor, and most of the other factors are between 0.1 and 0.4, proving that they also have a positive correlation with the landslide occurrence.

3.4. Factor Importance Ranking

From the above introduction of the architecture of RNNs and SRUs, it is clear that RNNs are effective in processing data that have sequential properties due to their special recurrent hidden states. Therefore, constructing models using RNNs should consider the problem of data redundancy and the input sequence of data. In this study, we propose a landslide data representation of RNNs, as shown in Figure 5. According to the results in Section 3.3, first, the IVs of all the influencing factors are arranged in a descending order, and then the influencing factors are ranked via their level of importance. Then, each pixel in the study area is converted into a continuous sample. Thus, the data are the input into the model in the previously ranked order of importance. Due to the special architecture inherent to RNNs, the previous input data are related to the latter input, and the key

information of each influencing factor that induces landslides is passed along the next hidden state.

3.5. Selection of Negative Sample

Landslides are geological events that occur infrequently but are hazardous to our society, and we can further define landslides as being rare events [50]. Identifying classes of rare events and representing them from a large quantity of data are challenging due to the insufficient number of positive samples and the absence of negative samples [51]. The lack of positive samples has been improved by adding the risk points above, and in this section, negative samples are selected by the weight of the evidence (WOE) method.

The WOE is calculated by Equation (1), from which it can be seen that the difference between the ratio of the number of landslides contained in the current class to the number of all landslide occurrences and the ratio of the number of non-landslide samples contained in the current class to the number of all non-landslide samples in this study is the logarithm of the two ratios. The larger the WOE is, the greater the probability of landslide events happening for the pixels belonging to this interval, and the opposite relation results in the probability of landslides being smaller.

To obtain the area for selecting the negative samples, the WOE of the 15 influencing factors for all pixels were summed and averaged in order to obtain a WOE map of the study area, and then the region was divided into two areas: positive WOE and negative WOE (Figure 6). To verify the effectiveness of this method, two groups of negative samples were selected: one group was randomly selected in the area of negative WOE region, and the other group was randomly selected directly in the study area. The number of negative samples in both groups was the same as the number of positive samples.

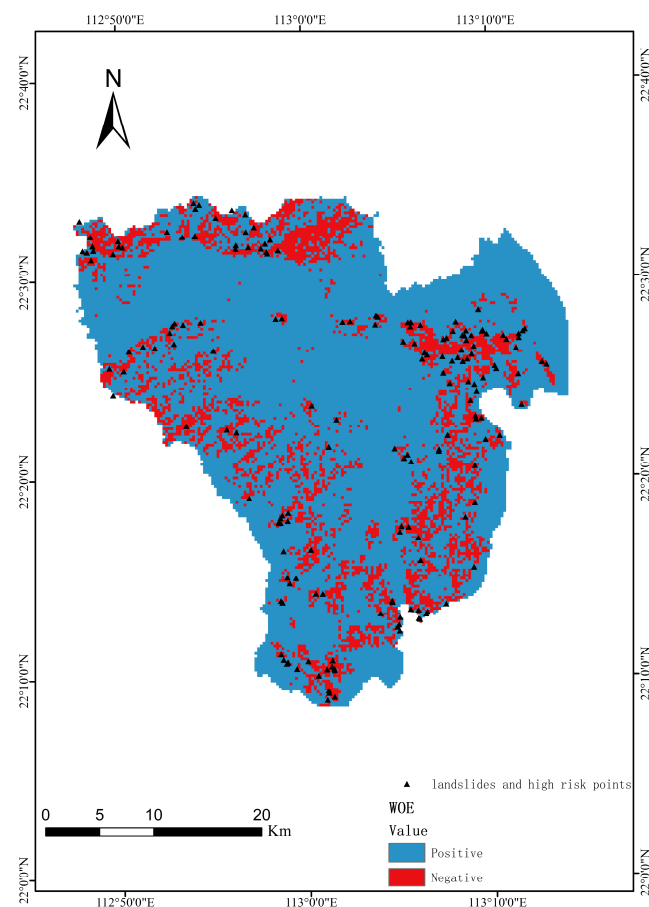


Figure 6. WOE and the selection of negative samples.

3.6. Evaluation and Comparison of Models

The validation of model strength or weakness is a key condition for assessing model performance. The fitting accuracy has been considered a significant feature and is obtained by comparing the model predictions with the true values in the training dataset. The analysis and evaluation of models using the receiver operating characteristic (ROC) curves are common in many related studies. The ROC curve is plotted by including the statistical index values of the false-positive and true-positive ratios. The area under the ROC curve (AUC) represents the model's predicted value. The AUC values range between 0.5 and 1.0, with larger areas indicating a better spatial prediction performance of the model [52]. Statistical indicators such as accuracy (ACC), Matthews correlation coefficient (MCC), F1 score, and recall are added to evaluate the predictive ability of the model, and these are calculated as follows [53–55]:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (10)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \quad (11)$$

$$F1 \text{ score} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (12)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (13)$$

where TP and TN represent true positives and true negatives, and FP and FN denote false positives and false negatives, respectively. The values of ACC, recall, and F1 score range between 0 and 1. MCC ranges between −1 and 1. The higher the ACC, F1, and MCC values, the better the predictive ability of the model.

4. Results

4.1. Performance of the Landslide Models

A dataset with negative samples selected by the IV analysis is input into the RNN and SRU models, named the RNN model and SRU model. A dataset randomly selected directly from the area of negative samples is input to the RNN and SRU models, named the RNN_random model and SRU_random model. The models are implemented in Python under scikit-learn (<https://scikit-learn.org/stable/>, accessed on 21 October 2022) and Keras (<https://keras.io/>, accessed on 21 October 2022). Parameters of the RNN model are as follows: hidden units = 40, learning rate = 0.0001, batch size = 128, epoch = 500. Parameters of the RNN model are as follows: hidden units = 40, learning rate = 0.0001, batch size = 128, epoch = 550, depth = 4, max features = 10,000.

The process of constructing the training and testing datasets is as follows: both of our datasets include 178 positive samples and 178 negative samples in order to construct the training and validation sets for the ML process; 70% of the positive samples (124) and negative samples (124) are used for training, and the remaining 30% (54 and 54) are used for testing. After training and testing the models, four machine learning models were evaluated using five criteria: AUC, ACC, MCC, F1 score, and recall. Table 4 shows the performance of the models. To verify that the method can work across data, we used the five-fold cross-validation, and Table 5 shows the averages of the statistical metrics of the five-fold cross-validation.

Table 4. Performance of the models.

Model Name	ACC	MCC	F1 Score	Recall
RNN	0.8598	0.7206	0.8544	0.8302
SRU	0.7850	0.5949	0.8099	0.9245
RNN_random	0.6887	0.3780	0.6796	0.6604
SRU_random	0.6132	0.2274	0.6306	0.6604

Table 5. The averages of the statistical metrics of 5-fold cross-validation.

Model Name	ACC	MCC	F1 Score	Recall
RNN	0.8220	0.6489	0.8278	0.8549
SRU	0.7591	0.5456	0.7639	0.8228
RNN_random	0.6570	0.3150	0.6601	0.6651
SRU_random	0.5834	0.1915	0.5869	0.5883

The results show that the performance of the RNN model and SRU model are higher than that of the RNN_random model and SRU_random model in all four statistical metrics, indicating that the dataset constructed with negative samples selected by information value analysis model fitting performance is significantly higher than that of the dataset with randomly selected negative samples. Regarding the ACC, the RNN model performs the best and achieves its highest ACC of 0.8598, which is over 0.0748 higher than that of the SRU (0.7850). The RNN model also achieves the highest MCC and F1 score (0.7206, 0.8544), which are 0.1257 and 0.0445 higher than those of the SRU model. In addition, it can be seen that the ML models trained with the IV analysis dataset outperform the ML model trained with the randomly selected negative samples dataset in terms of the RNN and SRU. This is evidenced by the fact that all statistical indicators for the ML models trained with the information value analysis dataset are greater than the ML model trained with the randomly selected negative samples dataset by more than 0.2.

Figure 7 plots the ROC curves of the four models. It can be seen that the AUC values of both the RNN model and the SRU model are above 0.90. In contrast, the AUC values of both the RNN_random model and the SRU_random model are low, indicating that the RNN and SRU techniques combined with the information value analysis show excellent predictive power for LSM. In addition, the RNN model achieves the highest AUC value (0.928), which is superior to the other models.

Figure 8 shows the accuracy and loss curves of four models, which are used to check the robustness of the results. When the model is optimized to the most stable level, the curves are presented as follows: as the epoch increases, the two accuracy curves gradually increase and level off; the two loss curves gradually decrease and level off (the loss curve of the training set decreases and the loss curve of the test set increases, indicating that the model may have an overfitting problem). All four models are optimized to the most robust level without overfitting problems.

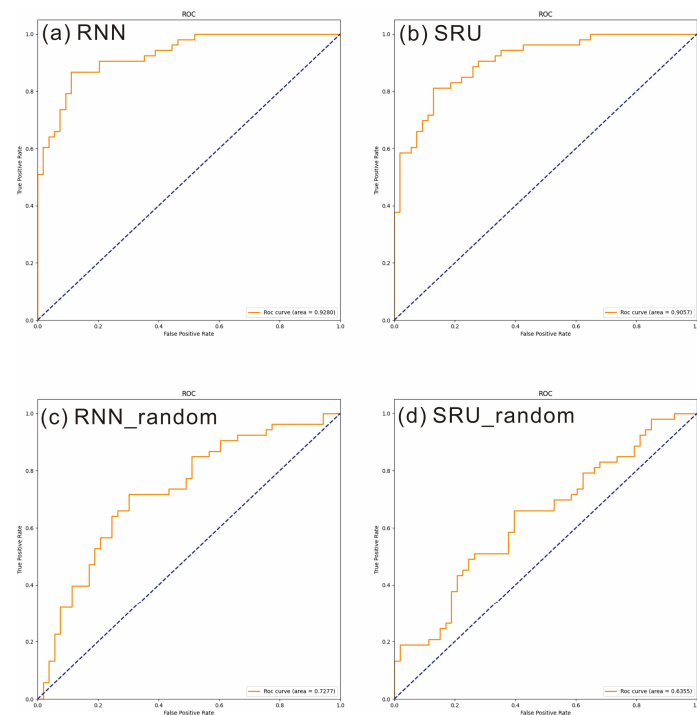


Figure 7. ROC curves of the four models.

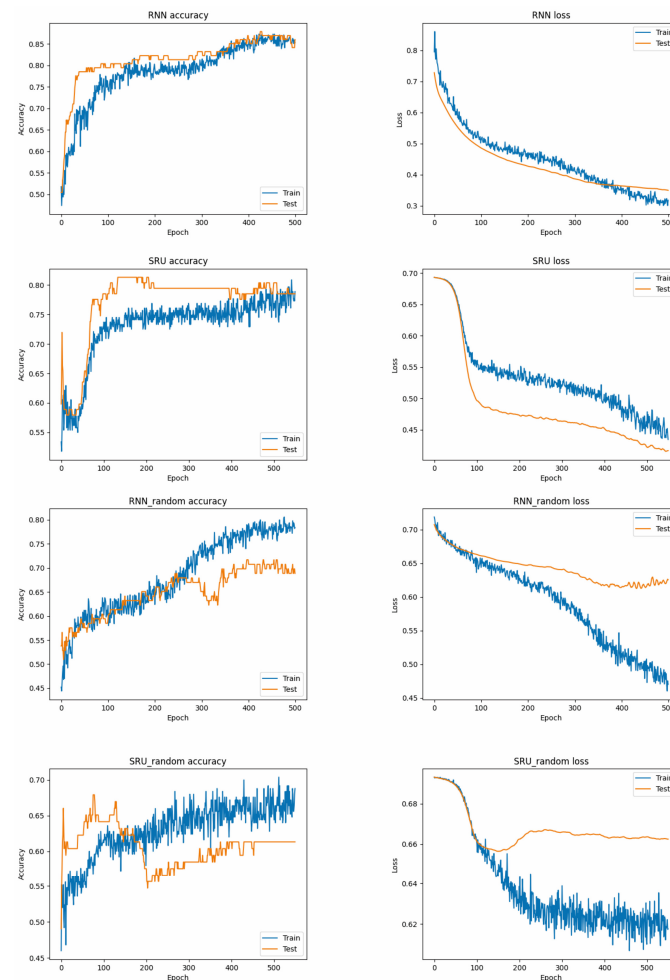


Figure 8. Accuracy and loss curves of the four models.

4.2. Landslide Susceptibility Maps

When LSM is used for comparison, the maps should be classified using quantitative methods [56]. The model output was analyzed and processed using ArcGIS. The maps were divided into five groups: very high, high, medium, low, and very low using the Jenks natural breaks classification method to finally obtain the landslide susceptibility maps (Figure 9). Among the four maps, most of the historical landslide and high-risk sites in Figure 9a–c are in the high landslide susceptibility areas, which are mainly located in the north, southwest, and southeast due to the mountainous terrain in the northwest and southwest of the study area and the strong human engineering activities in the northeast. According to the statistical indicators, the map shown in Figure 9a, which was constructed by the RNN model, is the best, compared to the map shown in Figure 9b, which was constructed by the SRU. Figure 9a does not have too many high susceptibility areas and does not predict low susceptibility areas such as rivers in the study area (Figure 2) as high susceptibility areas. Figure 9c, d also predict that some river areas are moderate and high susceptibility areas, which are not in accordance with the geomorphological conditions of the study area. Therefore, the map shown in Figure 9a is believed to be the best portrayal of the real-world conditions.

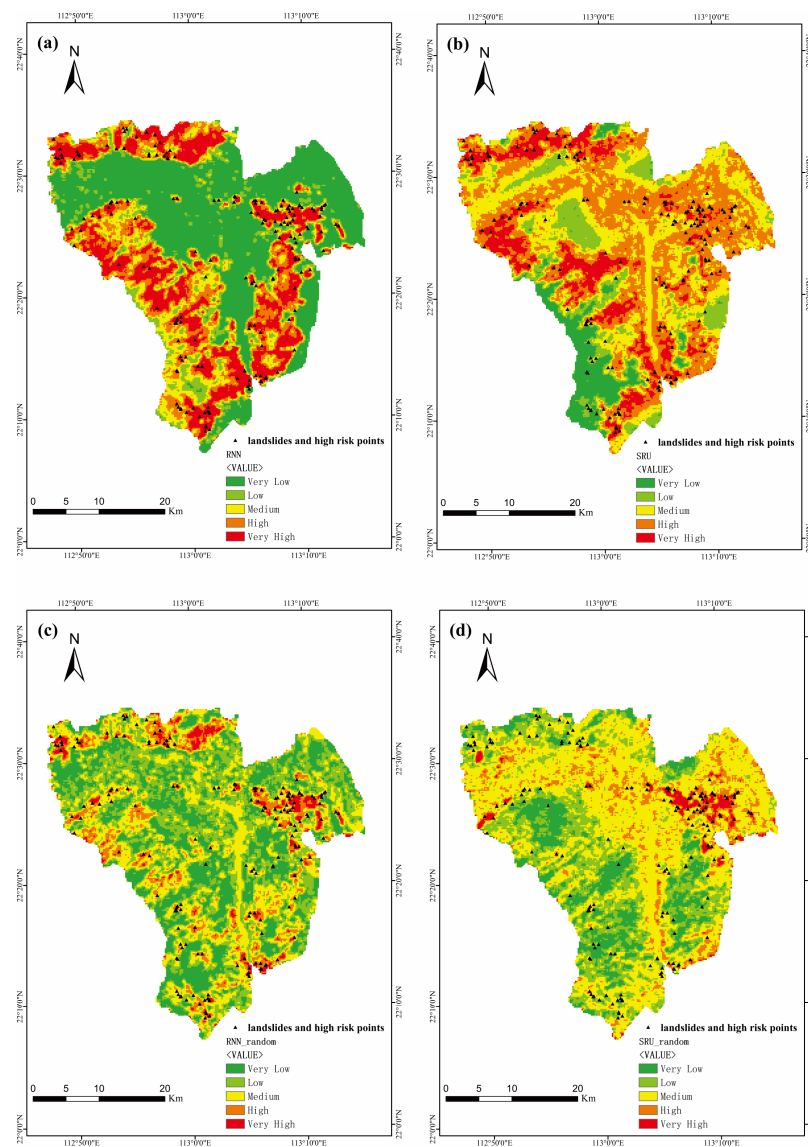


Figure 9. Landslide susceptibility maps by (a) RNN, (b) SRU, (c) RNN_random, and (d) SRU_random.

The visual data analysis initially shows the excellent results of the spatial predictive ability of the RNN model encompassing the LSM of the study area. The model evaluation results can still be described using mathematical-statistical methods (Table 6). LSM produces a model that focuses on high-susceptibility areas and models them simply and efficiently [57]. The evaluation of the practicability of models focuses on two groups, those with a rating of high and very high. First, we introduce the concept of landslide density (LD), which is the frequency ratio, referring to the ratio of the percentage of landslides (IV + V) to the percentage of groups (IV + V) in Table 6. It can be seen that the RNN model is more practical than the SRU model because although the RNN model covers fewer landslide and high-risk points than the SRU model (lower than 3.37%), the high susceptibility regions are much smaller than in the SRU model (lower than 16.31%). The low LD value of the high susceptibility regions of the SRU model also reflects the weak range of real-world applications when compared to that of the RNN model. The RNN_random model and SRU_random model cover too few landslide and high-risk points, indicating that the practical applications of these two models are poor.

Table 6. Practicability of the landslide susceptibility group.

Model	Group	No. of Pixels	Percentage of Group	Percentage of Group (IV + V)	No. of Landslide	Percentage of Landslide	Percentage of Landslide (IV + V)	LD
RNN	Very low (I)	13,854	40.34%	34.96%	3	1.69%	77.53%	0.042
	Low (II)	3846	11.20%		10	5.62%		0.502
	Medium (III)	4634	13.49%		27	15.17%		1.124
	High (IV)	5611	16.34%		48	26.97%		1.650
	Very high (V)	6396	18.62%		90	50.56%		2.715
SRU	Very low (I)	2537	7.39%	51.27%	4	2.25%	80.90%	0.304
	Low (II)	4054	11.81%		13	7.30%		0.619
	Medium (III)	10,144	29.54%		17	9.55%		0.323
	High (IV)	10,200	29.70%		59	33.15%		1.116
	Very high (V)	7406	21.57%		85	47.75%		2.214
RNN_random	Very low (I)	9165	26.69%	12.99%	3	1.69%	59.55%	0.063
	Low (II)	10,998	32.03%		10	5.62%		0.175
	Medium (III)	9717	28.30%		59	33.15%		1.171
	High (IV)	3374	9.82%		80	44.94%		4.574
	Very high (V)	1087	3.17%		26	14.61%		4.615
SRU_random	Very low (I)	4491	13.08%	12.54%	10	5.62%	32.86%	0.430
	Low (II)	8776	25.56%		41	23.03%		0.901
	Medium (III)	16,766	48.82%		69	38.76%		0.794
	High (IV)	3520	10.25%		45	25.28%		2.466
	Very high (V)	788	2.29%		13	7.30%		3.183

5. Discussion

5.1. Uniqueness of the Study Area

Although Xinhui District is neither an active seismicity area nor an extremely fragile geological environment area, and its climate is not special, its geographic location determines its unique economic location and its research value, as shown in Figure 1. As a new growth pole in the Guangdong Coastal Economic Belt and a destination for industrial transfer from the east to the west of the Guangdong–Hong Kong–Macao Greater Bay Area, the Xinhui District has become an important node district at the strategic intersection of the Guangdong Coastal Economic Belt and the Guangdong–Hong Kong–Macao Greater Bay Area in China, which is both an enormous opportunity and a great challenge. There will be more and more human activities in the Xinhui District, posing a very big challenge to future economic development and land use. Reasonable land planning cannot be separated from reliable geological hazard investigation and evaluation. Therefore, assessing the landslide susceptibility and the potential impacts of landslides on the economic environment can lay the foundation for optimizing the land use patterns and reducing the geological risk in the future.

5.2. Optimization of Non-Landslide Sample Selection

A variety of ML methods have been applied to LSM, with good results in recent years. However, previous studies have mostly focused on applying and comparing various ML methods to improve the performance of the models, but the selection of negative samples used to construct the models has affected the architecture construction of ML models. Randomly selecting non-occurring locations as negative samples will lead to considerable pollution, and conducting unsupervised cluster analysis to select negative samples still results in them being specified artificially, which also leads to a great deal of uncertainty in the resultant performance of the model. Therefore, we use the IV analysis to calculate the influencing factors based on historical landslide points to obtain negative samples that have less pollution to produce the landslide susceptibility maps.

The data in this study are different from the positive and negative sample problem that occurs in supervised learning; however, a positive and unlabeled (PU) problem occurs where there are only definite positive and unlabeled samples. It can only be assumed that the unlabeled samples may be negative samples without a level of certainty. Information value analysis was used to obtain the WOE for the entire study area as a basis for the selection of the negative samples. The final result comparison shows that this method works well, and that the negative data pollution is effectively limited. The groups of influencing factors within each pixel contain important data to consider both the positive and negative influences they have on landslides, and the negative value indicates that the importance is not in accordance with the daily logic. Therefore, we use the WOE with a proportional correction IV as an indicator for the most important factors for determining the order of the data for the input into the RNN model. The results indicate that the two slope-related factors, the TRI and profile curvature, were the most important factors in determining whether there was a chance of the occurrence of a landslide at that pixel location.

The problem of non-landslide sample selection has received attention, and many methods have been proposed recently, such as determining the proportion of non-landslide and landslide samples (because the value of negative samples is weaker than that of landslide samples, more non-landslide samples should be selected to improve the accuracy and avoid the imbalance of positive and negative samples caused by too many non-landslide samples), selecting non-landslide sample sets several times to find the best non-landslide sample set and using semi-supervised learning models. This study obtains negative samples with less pollution through the IV analysis. Overall, various studies on optimizing non-landslide sample selection have achieved satisfactory results. However, due to the differences in study areas and the logic and mechanisms behind different algorithms, there is no universally accepted method for optimizing non-landslide sample selection. A

comparative study using different methods for selecting non-landslide samples under the same conditions should be considered in the future.

5.3. Comprehensive Comparison of the Various Methods

Four datasets were input into the models, and Figure 9 shows that the dataset using less noisy negative data performs significantly better than the dataset with more noisy negative data in regard to their ROC, ACC, MCC, recall, and F1 values. After that, the traditional RNN model was compared to the newly proposed SRU model (which both use datasets that contain less noisy negative data) to produce two landslide susceptibility maps. Both models have excellent accuracy ($AUC > 0.900$), but from Tables 4 and 5, the RNN model generates a more reasonable area of high susceptibility for landslide events and identifies more historical points. Therefore, the map helps regional managers make effective decisions, and this study improves the prediction performance of deep learning techniques represented by RNNs in LSM.

6. Conclusions

This paper focuses on landslide susceptibility mapping (LSM) in the Xinhui District based on the RNN and SRU methods. Using the information value analysis, 15 landslide influencing factors were calculated, and their order of input in the recurrent neural network was determined. Then, the negative data were selected by the information value (IV) analysis. The 178 historical landslide and high-risk points were randomly divided into a training set and a test set for the model calculation, and the final landslide susceptibility maps were produced by the RNN and SRU for comparison purposes. The results led to the following conclusions: (1) the IV analysis method can improve the performance of machine learning methods in LSM by optimizing the selection of negative samples; (2) both the RNN and SRU models obtain excellent results in LSM ($AUC > 0.900$), but the LSM performance of the SRU, a newly proposed variant of RNNs, is weaker than the traditional RNN model in LSM; and (3) the RNN can produce accurate landslide susceptibility maps in areas that have the geography similar to that of the Xinhui District.

However, there are some limitations to be addressed in further studies, such as better consideration of the existing geomechanical properties, which are not well considered. Moreover, in addition to the characteristics of the non-landslide sample itself, whether the surrounding environment of the non-landslide area also influences the performance of the model needs to be better determined. In the future, more focus will be made on selecting more scientific non-landslide samples by increasing the influencing factors and analyzing the mutual influence of the surrounding environment, etc., to ensure the accuracy of the LSM results.

Author Contributions: Conceptualization, J.J. and Y.Z.; data curation, J.J., S.J. and S.L.; software, J.J.; supervision, Y.Z. and Q.C.; writing—original draft, J.J.; writing—review and editing, J.J., Y.Z. and Q.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was jointly supported by the National Natural Science Foundation of China (Grant No. U1911202), Guangdong Provincial Key R&D Project (Grant No.2020B1111370001) and China National Key R&D Project (Grant No. 2022YFF0800101).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Acknowledgments: The authors would like to thank Wei Cao, Hanyu Wang, the editors and anonymous reviewers who provided insightful comments to improve this article.

Conflicts of Interest: The authors declare no competing interest.

References

1. Press, F. Earth science and society. *Nature* **2008**, *451*, 301–303. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Zhang, Q.; Zhang, Z.Y. Big data helps geology develop rapidly. *Acta Petrol. Sin.* **2018**, *34*, 3167–3172.
3. Zhou, Y.; Zhang, Q.; Shen, W.; Xiao, F.; Zhang, Y.; Zhou, S.; Huang, Y.; Ji, J.; Tang, L.; Ouyang, C. Construction and Applications of Knowledge Graph of Porphyry Copper Deposits. *Earth Sci. Subsoil Use* **2021**, *44*, 204–218. [\[CrossRef\]](#)
4. Li, X.; Ting, G.; Shen, W.; Zhang, J.; Zhou, Y. Quantifying the influencing factors and multi-factor interactions affecting cadmium accumulation in limestone-derived agricultural soil using random forest (RF) approach. *Ecotoxicol. Environ. Saf.* **2021**, *209*, 111773. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Wang, J.; Zhou, Y.; Xiao, F. Identification of multi-element geochemical anomalies using unsupervised machine learning algorithms: A case study from Ag–Pb–Zn deposits in north-western Zhejiang, China. *Appl. Geochem.* **2020**, *120*, 104679. [\[CrossRef\]](#)
6. Yu, X.; Fan, X.; Zhou, Y.; Wang, Y.; Wang, K. Application of hierarchical clustering, singularity mapping, and Kohonen neural network to identify Ag–Au–Pb–Zn polymetallic mineralization associated geochemical anomaly in Pangxidong district. *J. Geochem. Explor.* **2019**, *203*, 87–95. [\[CrossRef\]](#)
7. Petley, D. Global patterns of loss of life from landslides. *Geology* **2012**, *40*, 927–930. [\[CrossRef\]](#)
8. Nohani, E.; Moharrami, M.; Sharafi, S.; Khosravi, K.; Pradhan, B.; Pham, B.T.; Lee, S.; Melesse, A.M. Landslide Susceptibility Mapping Using Different GIS-Based Bivariate Models. *Water* **2019**, *11*, 1402. [\[CrossRef\]](#)
9. Pham, B.T.; Prakash, I.; Bui, D.T. Spatial prediction of landslides using a hybrid machine learning approach based on Random Subspace and Classification and Regression Trees. *Geomorphology* **2018**, *303*, 256–270. [\[CrossRef\]](#)
10. Tsangaratos, P.; Ilia, I. Landslide susceptibility mapping using a modified decision tree classifier in the Xanthi Perfection, Greece. *Landslides* **2016**, *13*, 305–320. [\[CrossRef\]](#)
11. 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. [\[CrossRef\]](#)
12. Akgun, A. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: A case study at İzmir, Turkey. *Landslides* **2012**, *9*, 93–106. [\[CrossRef\]](#)
13. Hong, H.; Pradhan, B.; Sameen, M.I.; Kalantar, B.; Zhu, A.; Chen, W. Improving the accuracy of landslide susceptibility model using a novel region-partitioning approach. *Landslides* **2018**, *15*, 753–772. [\[CrossRef\]](#)
14. Bui, D.T.; Tran, A.T.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378.
15. Chen, W.; Wang, J.; Xie, X.; Hong, H.; Trung, N.V.; Bui, D.T.; Wang, G.; Li, X. Spatial prediction of landslide susceptibility using integrated frequency ratio with entropy and support vector machines by different kernel functions. *Environ. Earth Sci.* **2016**, *75*, 1344. [\[CrossRef\]](#)
16. Felicísimo, Á.M.; Cuartero, A.; Remondo, J.; Quiros, E. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study. *Landslides* **2013**, *10*, 175–189. [\[CrossRef\]](#)
17. Liang, W.-J.; Zhuang, D.-F.; Jiang, D.; Pan, J.-J.; Ren, H.-Y. Assessment of debris flow hazards using a Bayesian Network. *Geomorphology* **2012**, *171*–172, 94–100. [\[CrossRef\]](#)
18. 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. [\[CrossRef\]](#)
19. Nguyen, H.D. Spatial modeling of flood hazard using machine learning and GIS in Ha Tinh province, Vietnam. *J. Water Clim. Change* **2022**, *14*, 200–222. [\[CrossRef\]](#)
20. Ikram, R.M.A.; Dehrashid, A.A.; Zhang, B.; Chen, Z.; Le, B.N.; Moayed, H. A novel swarm intelligence: Cuckoo optimization algorithm (COA) and SailFish optimizer (SFO) in landslide susceptibility assessment. *Stoch. Environ. Res. Risk Assess.* **2023**, *37*, 1717–1743. [\[CrossRef\]](#)
21. Termeh, S.V.R.; Kornejady, A.; Pourghasemi, H.R.; Keesstra, S. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Sci. Total Environ.* **2018**, *615*, 438–451. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Hakim, W.L.; Rezaie, F.; Nur, A.S.; Panahi, M.; Khosravi, K.; Lee, C.-W.; Lee, S. Convolutional neural network (CNN) with metaheuristic optimization algorithms for landslide susceptibility mapping in Icheon, South Korea. *J. Environ. Manag.* **2022**, *305*, 114367. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Lv, L.; Chen, T.; Dou, J.; Plaza, A. A hybrid ensemble-based deep-learning framework for landslide susceptibility mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *108*, 102713. [\[CrossRef\]](#)
24. 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. [\[CrossRef\]](#)
25. Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sens. Environ.* **2020**, *241*, 111716. [\[CrossRef\]](#)

26. Pak, U.; Ma, J.; Ryu, U.; Ryom, K.; Juhyok, U.; Pak, K.; Pak, C. Deep learning-based PM2.5 prediction considering the spatiotemporal correlations: A case study of Beijing, China. *Sci. Total Environ.* **2020**, *699*, 133561. [\[CrossRef\]](#)
27. Ikram, R.M.A.; Mostafa, R.R.; Chen, Z.; Parmar, K.S.; Kisi, O.; Zounemat-Kermani, M. Water temperature prediction using improved deep learning methods through reptile search algorithm and weighted mean of vectors optimizer. *J. Mar. Sci. Eng.* **2023**, *11*, 259. [\[CrossRef\]](#)
28. 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.
29. Xiao, L.; Zhang, Y.; Peng, G. Landslide Susceptibility Assessment Using Integrated Deep Learning Algorithm along the China-Nepal Highway. *Sensors* **2018**, *18*, 4436. [\[CrossRef\]](#)
30. Graves, A. *Long Short-Term Memory*; Springer: Berlin/Heidelberg, Germany, 2012.
31. Dey, R.; Salemt, F.M. Gate-variants of Gated Recurrent Unit (GRU) neural networks. In Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA, 6–9 August 2017; pp. 1597–1600.
32. Denis, F.; Gilleron, R.; Letouzey, F. Learning from positive and unlabeled examples. *Theor. Comput. Sci.* **2005**, *348*, 70–83. [\[CrossRef\]](#)
33. Wu, B.; Qiu, W.; Jia, J.; Liu, N. Landslide Susceptibility Modeling Using Bagging-Based Positive-Unlabeled Learning. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 766–770. [\[CrossRef\]](#)
34. Yang, C.; Liu, L.-L.; Huang, F.; Huang, L.; Wang, X.-M. Machine learning-based landslide susceptibility assessment with optimized ratio of landslide to non-landslide samples. *Gondwana Res.* **2022**, *in press*. [\[CrossRef\]](#)
35. Chang, Z.; Huang, J.; Huang, F.; Bhuyan, K.; Meena, S.R.; Catani, F. Uncertainty analysis of non-landslide sample selection in landslide susceptibility prediction using slope unit-based machine learning models. *Gondwana Res.* **2023**, *117*, 307–320. [\[CrossRef\]](#)
36. Huang, F.; Cao, Z.; Jiang, S.-H.; Zhou, C.; Huang, J.; Guo, Z. Landslide susceptibility prediction based on a semi-supervised multiple-layer perceptron model. *Landslides* **2020**, *17*, 2919–2930. [\[CrossRef\]](#)
37. Heckmann, T.; Gegg, K.; Gegg, A.; Becht, M. Sample size matters: Investigating the effect of sample size on a logistic regression susceptibility model for debris flows. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 259–278. [\[CrossRef\]](#)
38. Shannon, C. The lattice theory of information. *Trans. IRE Prof. Group Inf. Theory* **1953**, *1*, 105–107. [\[CrossRef\]](#)
39. Howard, R.A. Information Value Theory. *IEEE Trans. Syst. Sci. Cybern.* **1966**, *2*, 22–26. [\[CrossRef\]](#)
40. Freedman, S.; Jin, G.Z. The information value of online social networks: Lessons from peer-to-peer lending. *Int. J. Ind. Organ.* **2017**, *51*, 185–222. [\[CrossRef\]](#)
41. Kliger, D.; Sarig, O. The Information Value of Bond Ratings. *J. Financ.* **2000**, *55*, 2879–2902. [\[CrossRef\]](#)
42. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* **2018**, *4*, e00938. [\[CrossRef\]](#)
43. Tealab, A. Time series forecasting using artificial neural networks methodologies: A systematic review. *Future Comput. Inform. J.* **2018**, *3*, 334–340. [\[CrossRef\]](#)
44. Elman, J.L. Finding structure in time. *Cogn. Sci.* **1990**, *14*, 179–211. [\[CrossRef\]](#)
45. Jordan, M.I. Serial order: A parallel distributed processing approach. In *Advances in Psychology*; Elsevier: Amsterdam, The Netherlands, 1997.
46. Lei, T.; Zhang, Y.; Wang, S.I.; Dai, H.; Artzi, Y. Simple recurrent units for highly parallelizable recurrence. *arXiv* **2017**, arXiv:1709.02755.
47. Jiang, C.; Chen, S.; Chen, Y.; Bo, Y.; Han, L.; Guo, J.; Feng, Z.; Zhou, H. Performance Analysis of a Deep Simple Recurrent Unit Recurrent Neural Network (SRU-RNN) in MEMS Gyroscope De-Noising. *Sensors* **2018**, *18*, 4471. [\[CrossRef\]](#)
48. Chen, W.; Xie, X.; Wang, J.; Pradhan, B.; Hong, H.; Bui, D.T.; Duan, Z.; Ma, J. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena* **2017**, *151*, 147–160. [\[CrossRef\]](#)
49. Zhao, X.; Chen, W. Optimization of Computational Intelligence Models for Landslide Susceptibility Evaluation. *Remote Sens.* **2020**, *12*, 2180. [\[CrossRef\]](#)
50. Karpatne, A.; Ebert-Uphoff, I.; Ravela, S.; Babaie, H.A.; Kumar, V. Machine learning for the geosciences: Challenges and opportunities. *IEEE Trans. Knowl. Data Eng.* **2018**, *31*, 1544–1554. [\[CrossRef\]](#)
51. Guo, Q.; Li, W.; Liu, Y.; Tong, D. Predicting potential distributions of geographic events using one-class data: Concepts and methods. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 1697–1715. [\[CrossRef\]](#)
52. Wang, Y.; Fang, Z.; Wang, M.; Peng, L.; Hong, H. Comparative study of landslide susceptibility mapping with different recurrent neural networks. *Comput. Geosci.* **2020**, *138*, 104445. [\[CrossRef\]](#)
53. Huang, J.; Ling, C. Using AUC and accuracy in evaluating learning algorithms. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 299–310. [\[CrossRef\]](#)
54. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Powers, D.M.W. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv* **2020**, arXiv:2010.16061.

56. Goetz, J.N.; Brenning, A.; Petschko, H.; Leopold, P. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Comput. Geosci.* **2015**, *81*, 1–11. [[CrossRef](#)]
57. Bui, D.T.; Tsangaratos, P.; Nguyen, V.-T.; Van Liem, N.; Trinh, P.T. Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. *Catena* **2020**, *188*, 104426. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.