Applied Mathematics and Computational Intelligence
Volume 15, No. 2, 2026 [92-103]

Key Human Activity Variable for Landslide Prediction in Western Sarawak

Nur Hisyam Ramli[1*], Siti Noor Linda Taib[2], Norazzlina M. Sa'don[3], Raudhah Ahmadi[4], Imtiyaz Akbar Najar[5], Dayangku Salma Awang Ismail[6], Rosmina Ahmad Bustami[7], Tarmiji Masron[8], Nazeri Abdul Rahman[9]

[1,2,3,4,5,6]Department of Civil Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, 94300, Sarawak, MALAYSIA
[7]UNIMAS Water Centre (UWC), Faculty of Engineering, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, MALAYSIA
[1,8]Centre for Spatially Integrated Digital Humanities (CSIDH), Faculty of Social Sciences and Humanities, Universiti Malaysia Sarawak, 94300, Sarawak, MALAYSIA
[9]Department of Chemical Engineering and Energy Sustainability, Faculty of Engineering, Universiti Malaysia Sarawak, 94300, Sarawak, MALAYSIA

[*] Corresponding author: 22020251@siswa.unimas.my

Received: 14 February 2025
Revised: 10 April 2025
Accepted: 17 June 2025

ABSTRACT

Landslide is a common form of natural disaster in a tropical country such as Malaysia, and its presence is mostly concentrated during the wet season. Human activities have been known to influence landslides as development often causes displacement of the original slope. Western Sarawak is a region where development is currently taking place at a steady rate, with the highest population density in Sarawak settlement is expected to increase in the coming years, increasing the impact of human activity on landslide occurrences in the region. However, the most prevalent representation of human activity in the area has not been discussed. It was determined through a Machine Learning based approach that amongst the human activities features of Distance from Road (DFR), Normalised Difference Vegetation Index (NDVI), and Land Use and Land Cover (LULC), DFR was the most influential feature with an importance of 0.130, followed by NDVI with 0.072, and the LULC type of crops and trees, with an importance of 0.36 respectively.

Keywords: Landslide, Machine Learning, Neural Network, Sarawak, Spatial

1 INTRODUCTION

Landslide is a prevalent form of natural disaster in tropical country such as Malaysia where it is widely known to be triggered by rainfall [1]. Due to climate change which causes extreme heat and rainfall, landslide intensity and frequency has been observed to increase throughout the years [2]. Landslide impact can be seen in a multitude of aspect such as economical, societal, and environmental, as major ones can demolish anything that is in its path [3]. Thus, it is an utmost importance to determine the landslide susceptibility of an area of interest to mitigate the ill effects.

Contemporarily, landslide susceptibility is determined through a Machine Learning (ML) based approach that is driven by data to make predictions [4]. ML models particularly through a supervised learning approach has been deployed in both locally and abroad with high degree of success when it comes to predicting landslide susceptibility. Locally, the ML model of an Artificial Neural Network (ANN) has been deployed for the purpose of landslide susceptibility mapping (LSM) where it has been determined to have a high degree of accuracy [5]. An advantage of using ML model to predict landslide is it capability to develop the LSM through Remote Sensing (RS) data with ease.

RS data are spatial information of a certain variable obtained through RS devices such as satellites. It is supported and provided by various agencies and boast a large library of variable types. One of such variables is rainfall, which is the main triggering variable of landslides in Malaysia. Other variables such as the topographical variables of aspect, curvature, elevation, and slope angle are important conditioning factors when it comes to developing an ML model for landslides prediction [6]. Another type of variable that is important to consider when developing a landslide prediction is model is human activity. Human activity often coincides with the degradation of existing slopes to make way for development, which increases the landslide susceptibility. It is commonly represented through variables that are provided by RS organisations such as Distance from Roads (DFR), Land Use and Land Cover (LULC), and Normalised Difference Vegetation Index (NDVI) [7].

The impact that each human activity variables has on landslide susceptibility however has not been evaluated in Western Sarawak. Knowing the impact helps future ML model development when it comes to variable selection for human activity. Western Sarawak is a region in the state which is known to be the most densely populated region in Sarawak, thus increasing the likelihood of exposure to landslide susceptible areas [8]. In this study, the human activity variables of DFR, LULC, and NDVI will be evaluated through the weightage that each variable has on the ANN model prediction to determine the suitable variable to represent human activity in Western Sarawak.

2 MACHINE LEARNING FOR LANDSLIDE PREDICTION

An ML based approach is currently gaining traction as it was made accessible with the advancement in both computing hardware and software, as well as the availability of data required for the models development process [9]. The development approach of an ML model is dependent on the use-case, where it can be categorised into a supervised or unsupervised approach, where a supervised models are mainly used to predict trends whereas unsupervised models are mainly used to generate new information [10]. However, in both approaches, the models were not explicitly programmed to develop its predictive or generative ability. Instead, the models learn the intricate relationship between the variables in development dataset based on its hyperparameters. As supervised models are used in prediction-based problems, the models are particularly beneficial in predicting landslide susceptibility.

ANN is one of the many algorithms that are available to develop a supervised ML model for landslide prediction, where it learns the relationship between the landslide features and the outcomes by iteratively adjusting the significance of each features in a series of neurons that is designed to mimic the human brain [11]. The importance of each features in the interconnected neurons within the ANN is determined through weights, where a heavier weights indicates a higher significance towards the outcome [12]. Hence, an ANN model was developed in this study to determine the key human activity variable for landslide prediction in Western Sarawak.

3 STUDY AREA

Western Sarawak is a region located in the Western-most tip of East Malaysia. It experiences heavy rainfall during the Northeast monsoon which commonly starts from November to March. The region is home to the state's capital city of Kuching, higher education institutions, and several industrial zones [13]. Thus, it is a centre of attention for the people of Sarawak as they migrate to the region for job opportunities. Expansion is known to cause an increase in exposure to landslide susceptible areas as well as increasing the susceptibility itself. Figure 1(a) shows the DFR map, where the area furthest from any roads in Western Sarawak is located at 6,355.61 m. NDVI shows the vegetation density in the form of index, with higher values indicating a denser vegetation (Figure 1(b)). The LULC of Western Sarawak can be seen in Figure 1(c) which shows that the major LULC in the area consists of trees.

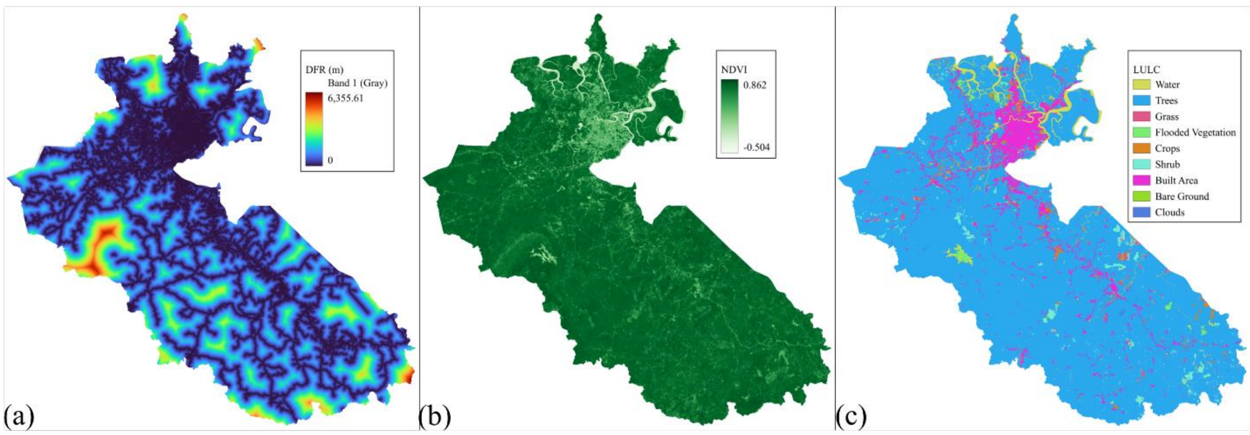


Figure 1: Western Sarawak map of (a) DFR, (b) NDVI, and (c) LULC.

4 METHODOLOGY

4.1 Data Preparation

Data in an integral component for the ANN model where it is used to develop and evaluate the model's performance with the training data and testing data respectively with the data preparation procedures as seen in Figure 2. The targets initially consisted of 50 landslide points supplied by the Department of Minerals and Geosciences, and 1,000 non-landslide points obtained by randomly sampling areas with slope angle lower than 5° [14]. The features on the other hand were provided in a raster file format, which is a continuous image file, and can be categorised into topographical, triggering, and human activities. The topographical features provided by the National Space Agency (NASA) consisted of aspect, curvature, elevation, and slope angle [15]. The sole triggering factor of rainfall intensity for the wet season was supplied by the Climate Hazard Group [16]. The human activity features that were to be analysed in this study consisted of DFR, NDVI, and LULC, which were respectively provided by the OpenStreetMap organization, NASA, and Environmental Systems Research Institute [17]. To prepare the primary dataset the features raster files were stacked and the values of each features at the respective target points were obtained by using the Quantum Geographical Information System (QGIS) raster sampling tool [18]. As the targets were at an imbalance with non-landslide points greatly outweighing the landslide points, the Synthetic Minority Oversampling Technique (SMOTE) was applied to the primary dataset which generates synthetic data for the minority landslide points by computing the k-nearest neighbours [19]. 150 landslide points were generated in this study, as a ratio of 1:5 between landslides and non-landslides were observed to be the optimum ratio in a previous study [20]. The primary dataset then underwent a multicollinearity check with the threshold set at 0.8 where if a pair of features exhibited a correlation coefficient, one of the variables were to be removed [21]. High collinearity are known to degrade an ML model predictive performance through overfitting where the model could not generalise well on never-before-seen data [22]. The multicollinearity assessment was done through the "Data Analysis" plugin in Microsoft Excel with the correlation function.

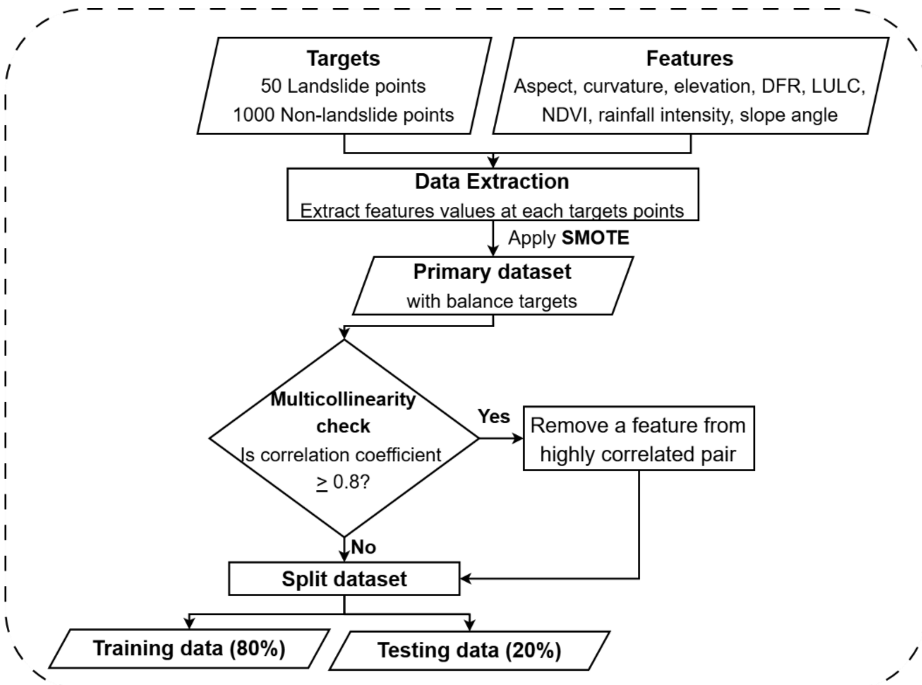


Figure 2: Data preparation procedures.

4.2 ANN Model Development

The ANN model was developed in R through the "neuralnet" package [23]. The ANN model hyperparameter consisted of four neurons in a single hidden layer, a learning rate of 0.01, a backpropagation learning algorithm, and a maximum step of 1E+6 for the weight adjustment. To evaluate the ANN model training success rate and prediction success rate from the training data and testing data respectively, the metrics of precision, recall, F1-Score, Area Under the Receiver Operating Characteristics Curve (AUROC), and accuracy were deployed.

Precision determines the correctness of the positive prediction which was the landslide occurrence in this study (Equation 1) [24].

$$\text{Precision} = \text{True Positive} / (\text{True Positive} + \text{False Positive}) \quad (1)$$

Recall or True Positive Rate determines the model's capability in predicting all positive instances, and it was crucial in this study, as the number of positive instances was the minority target (Equation 2) [19] .

$$\text{Recall} = \text{True Positive}/\ (\text{True Positive} + \text{False Negative}) \tag{2}$$

F1-score is the harmonic mean between precision and recall where in an imbalanced dataset, it provides a better measure of accuracy in compared to the traditional accuracy metric (Equation 3) [25].

$$\text{F1-score} = 2\ [(\text{Precision} * \text{Recall})/\ (\text{Precision} + \text{Recall})] \tag{3}$$

AUROC shows the model's ability to differentiate the targets classes, and it is determined through finding the area under the True Positive Rate vs False Positive Rate curve [26]. Accuracy provides an overall view of the model's ability to predict all instances regardless of positive or negative instances. However, it can be misleading in imbalanced datasets such that of this study [27].

4.3 Variable Importance Deduction

To determine the importance of each human activity landslide factors based on the ANN model prediction, the "NeuralNetTools" package was deployed [28]. The Garson's function in the package was used to evaluate the variable importance. The function provides relative importance of input variables in the neural networks as the sum of the weightage from the input to the hidden layer, the hidden layer to the output layer.

5 RESULTS AND DISCUSSION

5.1 Multicollinearity Check

High multicollinearity denoted by a correlation coefficient exceeding 0.8 for a pair of features in an ML model development environment is not a preferable condition as could potentially introduce overfitting during the model's training phase [29]. Overfitting is a problem where the model remembers the solution for each line of data in the training data instead of understanding the intricate relationship between the features and targets, which causes the model to not generalise well on never-before-seen data [30]. In this study, the primary dataset that has underwent SMOTE was tested for multicollinearity through Microsoft's Excel Data Analysis plug-in and the results was represented through a correlation matrix as seen in Table 1. Significant correlation was observed between elevation and slope angle with a correlation coefficient of 0.829, which was above the threshold of 0.8. Amongst the pair of features, slope angle was removed as opposed to elevation as it was determined to have a higher correlation to the target. The removal was due to the aspect of variety, where in combination with the other features, elevation would carry more unique information in comparison to slope angle, as the inclusion of slope angle will lead the model to have an unjustified weightage assignment [31]. Besides the feature pair of slope angle and elevation, no other significant correlation amongst the features were observed, ensuring that the removal of the highly correlated pair were the only variables that might introduce multicollinearity issues if included in the developmental dataset. Post removal of the highly correlated pair, the primary dataset was split into training data and testing data with a ratio of 80:20.

Table 1: Correlation matrix for primary dataset

	Aspect	Curve	Elev	DFR	LULC	NDVI	Rain	Slope	Slide
Aspect	1								
Curve	-0.024	1							
Elev	-0.166	0.283	1						
DFR	-0.082	0.000	0.078	1					
LULC	0.094	-0.020	-0.265	-0.223	1				
NDVI	-0.079	0.074	0.250	0.181	-0.301	1			
Rain	0.036	-0.085	-0.274	0.170	0.145	0.064	1		
Slope	-0.162	0.115	0.829	-0.018	-0.287	0.269	-0.270	1	
Slide	-0.132	0.175	0.670	-0.086	-0.259	0.259	-0.261	0.883	1

Legend for correlation

Negative Positive

Where: Curve is curvature, Elev is elevation, Rain is rainfall intensity, Slope is slope angle, and Slide is landslide

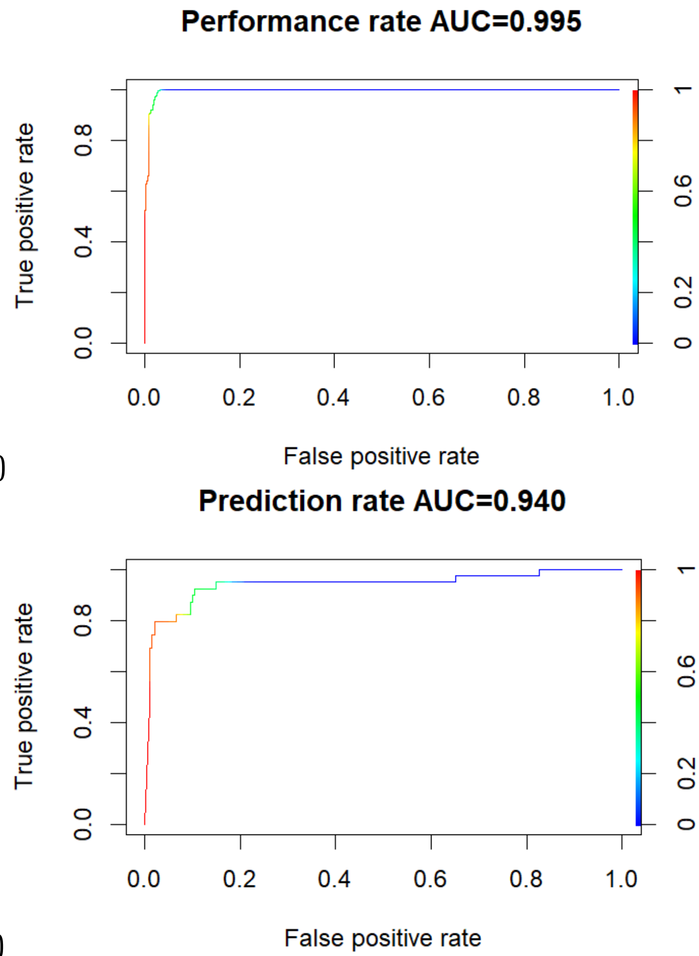
5.2 ANN Model Performance

The results of the ANN model training and testing can be seen in Table 2(a) and Table 2(b) respectively. The ANN model performance was evaluated through precision, recall, F1-score, accuracy, and AUC. Precision was evaluated to the correctness of the positive prediction, where in the training phase, the model was determined to have a precision of 0.943, whereas it's testing phase shows that the model scored a precision of 0.640 [9]. Recall performance of the model was evaluated to be 0.919, and 0.821 for the training phase and testing phase respectively. As compared to the precision metric, recall was more critical to achieve a higher score, as the positive instances of landslide occurrences were more important to be accurately predicted. As for the F1-score and accuracy, the training phase yielded an F1-score of 0.931 and an accuracy of 0.977, and the testing phase yielded an F1-score of 0.718 and an accuracy of 0.896.

Table 2: Performance matrix of ANN Model based on (a) Training data, and (b) Testing data

(a)	**Predicted**		(b)	**Predicted**	
Actual	**FALSE**	**TRUE**	**Actual**	**FALSE**	**TRUE**
0	789	9	0	183	18
1	13	148	1	7	32
Precision	0.943		Precision	0.640	
Recall	0.919		Recall	0.821	
F1-score	0.931		F1-score	0.718	
Accuracy	0.977		Accuracy	0.896	

The final evaluation metric for the ANN model was AUROC, which is a provides an indicator on how well the model classify the targets of landslides and non-landslides [32]. The models performance rate based on the training data can be seen in Figure 3(a) where it shows the model performance rate AUROC was 0.995. As for the model prediction rate based on the testing data, the AUROC for the model was 0.940 as seen in Figure 3(b).



(a)

(b)

Figure 3: AUC for ANN model (a) Training phase, and (b)Testing phase.

The model development was concluded as a success, as the main metrics of recall, and AUROC were determined to have satisfactory results, the analysis on variables importances were then conducted.

5.3 Variable Importance

In Malaysia, human activity has been determined to one of the main contributors of landslide occurrences as it introduces live load unto the slopes and generally involves the modification of the natural slopes [1]. In this study, the human activity features were represented through DFR, LULC, and NDVI. All three of these human activities features are commonly used in numerous landslide predictions study through ML models as the data are provided free for public use [33]. DFR represents human activity in a numerical manner as in Malaysia, live loads from settlements are generally in close proximity to the road, as well as the load from passing vehicles [34]. LULC on the hand is categorical representation of human activity, as it consisted of water, trees, grass, flooded vegetations, crops, shrubs, built areas, bare grounds, and clouds [17]. Thus, when evaluating the importance of LULC as a variable, it must be evaluated in a logic-based evaluation, where each representation of LULC is evaluated based on the type. NDVI is a representation of surface vegetation cover density, a lower NDVI tends be more correlated to a higher human activity as development tends to cause surface vegetation removal. When vegetation is removed, the tendency of landslide is higher due to the lack of support from the surface vegetation roots, and the increase in rainfall seepage [35].

The importance of each features used in developing the ANN model was evaluated through the Garson's function from the "NeuralNetTools" package in R, where it provides the importance of each features in relative terms as the sum of weightage from the start to end layer in the neural networks [28]. It was determined through the Garson's function as seen in Figure 4, that amongst the human activities features, DFR was the most influential with an importance of 0.130, followed NDVI with an importance of 0.072, and finally LULC with crops and trees respectively having an importance of 0.036, built areas with 0.033, grass with 0.028, flooded vegetation with 0.022, bare grounds with 0.020, clouds with 0.013, water with 0.007, and shrubs with 0.003.

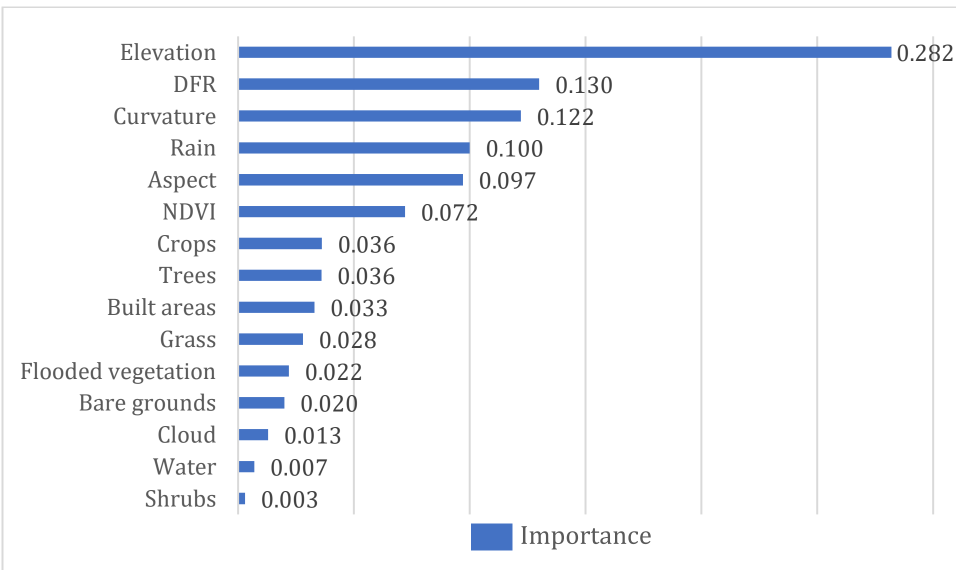


Figure 4: Features importance in the ANN model.

Thus, it was concluded that in terms of human activities features that are most influential on the occurrence of landslides in Western Sarawak, DFR was the most influential features, followed by NDVI, and the LULC of crops, trees, built areas, grass, flooded vegetation, bare grounds, cloud, water, and shrubs. DFR provides the most information to the ANN model

6 CONCLUSION

Human activities are amongst the prevalent features that are related to landslides not only in Western Sarawak, but throughout Malaysia. Commonly, it is represented through publicly available features such as Distance from Roads, Normalised Difference Vegetation Index, and Land Use and

Land Cover. Knowing the importance of each human activities features provides insight into the level of importance for a proper approach for landslide prevention methods. It was determined that through a Machine Learning based approach with Artificial Neural Networks, the most important features of human activities to be considered in the development process is Distance from Road with a relative importance of 0.130, followed by Normalised Difference Vegetation Index with 0.072, and the Land Use and Land Cover of crops and trees with an importance of 0.036 respectively. Future research should focus on developing a new feature with represents human activities more comprehensively.

ACKNOWLEDGEMENT

The authors acknowledge Ministry of Higher Education Malaysia, Fundamental Research Grant Scheme, FRGS/1/2022/TK06/UNIMAS/01/1 and Universiti Malaysia Sarawak for supporting this project.

REFERENCES

[1] N. A. Majid, M. R. Taha, and S. N. Selamat, "Historical landslide events in Malaysia 1993-2019," *Indian J. Sci. Technol.*, vol. 13, no. 33, p. 3387, 2020, doi: 10.17485/IJST/v13i33.884.

[2] M. I. N. Diana, N. Muhamad, M. R. Taha, A. Osman, and M. M. Alam, "Social vulnerability assessment for landslide hazards in Malaysia: A systematic review study," Mar. 01, 2021, *MDPI AG*. doi: 10.3390/land10030315.

[3] E. N. C. Perera, D. T. Jayawardana, P. Jayasinghe, R. M. S. Bandara, and N. Alahakoon, "Direct impacts of landslides on socio-economic systems: a case study from Aranayake, Sri Lanka," *Geoenvironmental Disasters*, vol. 5, no. 1, Dec. 2018, doi: 10.1186/s40677-018-0104-6.

[4] K. Ullah, Y. Wang, Z. Fang, L. Wang, and M. Rahman, "Multi-hazard susceptibility mapping based on Convolutional Neural Networks," *Geosci. Front.*, vol. 13, no. 5, p. 101425, 2022, doi: 10.1016/j.gsf.2022.101425.

[5] S. N. Selamat, N. Abd Majid, and A. Mohd Taib, "A comparative assessment of sampling ratios using artificial neural network (ANN) for landslide predictive model in Langat River Basin, Selangor, Malaysia," *Sustain.*, vol. 15, no. 861, pp. 1–21, Jan. 2023, doi: 10.3390/su15010861.

[6] Y. Zhou, H. Wang, R. Yang, D. Xie, and J. Liu, "Semantic Segmentation Algorithm of Landslide Based on Remote Sensing Image and DEM," in *2022 5th International Conference on Pattern Recognition and Artificial Intelligence, PRAI 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 595–599. doi: 10.1109/PRAI55851.2022.9904119.

[7] A. Habibi, M. R. Delavar, M. S. Sadeghian, B. Nazari, and S. Pirasteh, "A hybrid of ensemble machine learning models with RFE and Boruta wrapper-based algorithms for flash flood susceptibility assessment," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 122, no. March, p. 103401, 2023, doi: 10.1016/j.jag.2023.103401.

[8] DOSM, “Sarawak Population,” Sarawak Government. Accessed: May 09, 2024. [Online]. Available: https://sarawak.gov.my/web/home/article_view/240/175/

[9] A. Bitzenis, N. Koutsoupias, and M. Nosios, “Artificial intelligence and machine learning in production efficiency enhancement and sustainable development: a comprehensive bibliometric review,” *Front. Sustain.*, vol. 5, 2025, doi: 10.3389/frsus.2024.1508647.

[10] C. Kavitha, G. Tripathi, R. Sridivya, V. Yamuna, N. Supriya, and A. Lakshmanarao, “An Efficient Churn Prediction model using ML supervised and semi supervised Learning Techniques,” in *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, Raghu Engineering College Dakamarri,Department of CSE,Visakhapatnam,A.P,India; Guru Nanak Institutions Technical Campus,Hyderabad,India; Aditya College of Engineering and Technology,Department of CSE,Surampalem,A.P,India; Lendi Institute of Engineering an: IEEE, 2024. doi: 10.1109/i2ct61223.2024.10543994.

[11] N. Menghi *et al.*, “The effects of task similarity during representation learning in brains and neural networks,” *bioRxiv*, 2025, doi: 10.1101/2025.01.20.633896.

[12] H. Borakhade, D. Onkar, and D. Borakhade, “Comparing neural networks in human brains and AI systems with focus on genetics and experiential differences,” in *2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI)*, NKP Salve Institute of Medical Sciences & Research Centre and Lata Mangeshkar Hospital,Nagpur,Maharashtra,India; NKP Salve Institute of Medical Sciences & Research Centre and Lata Mangeshkar Hospital,Department of Anatomy,Nagpur,Maharashtra,India; St. Vin: IEEE, 2024, pp. 1–6. doi: 10.1109/idicaiei61867.2024.10842761.

[13] L. Ai Ling, M. Lucy Gregory, and B. Muniapan, “Marketing education in Sarawak: some issues and challenges from the employers’ perspectives,” 2010. [Online]. Available: http://www.sarawakscore.com.my/

[14] S. Jamaludin, B. B. K Huat, and H. Omar, “Evaluation and development Of cut–slope assessment systems for Peninsular Malaysia in predicting landslides in granitic formation,” *J. Teknol.*, vol. 44, no. B, pp. 31–46, 2006, doi: 10.11113/jt.v44.360.

[15] T. G. Farr *et al.*, “The shuttle radar topography mission,” *Rev. Geophys.*, vol. 45, no. 2, pp. 1–33, 2007, doi: https://doi.org/10.1029/2005RG000183.

[16] C. Funk *et al.*, “The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes,” *Sci. Data*, vol. 2, pp. 1–21, 2015, doi: 10.1038/sdata.2015.66.

[17] K. Karra, C. Kontgis, Z. Statman-Weil, J. C. Mazzariello, M. Mathis, and S. P. Brumby, “Global land use / land cover with Sentinel 2 and deep learning,” in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021, pp. 4704–4707. doi: 10.1109/IGARSS47720.2021.9553499.

[18] QGIS Development Team, “QGIS desktop 3.22 user guide,” 2023. [Online]. Available: https://docs.qgis.org/3.22/pdf/en/QGIS-3.22-DesktopUserGuide-en.pdf

[19] P. Kumar, P. Priyanka, K. V. Uday, and V. Dutt, "Addressing class imbalance in soil movement predictions," *Nat. Hazards Earth Syst. Sci.*, vol. 24, no. 6, pp. 1913–1928, 2024, doi: 10.5194/nhess-24-1913-2024.

[20] M. Li and H. Tian, "Insights from optimized non-landslide sampling and SHAP explainability for landslide susceptibility prediction," *Appl. Sci.*, vol. 15, no. 3, p. 1163, 2025, doi: 10.3390/app15031163.

[21] A. R. M. Towfiqul Islam *et al.*, "Flood susceptibility modelling using advanced ensemble machine learning models," *Geosci. Front.*, vol. 12, no. 3, 2021, doi: 10.1016/j.gsf.2020.09.006.

[22] J. M. Lin and L. Billa, "Spatial Prediction of Flood-Prone Areas Using Geographically Weighted Regression," *Environ. Adv.*, vol. 6, no. 100118, pp. 1–19, 2021, doi: 10.1016/j.envadv.2021.100118.

[23] S. Fristch, F. Guenther, M. N. Wright, and S. M. Mueller, "Package 'neuralnet,'" CRAN. Accessed: Jan. 19, 2024. [Online]. Available: https://cran.r-project.org/web/packages/neuralnet/neuralnet.pdf

[24] J. C. Obi, "A comparative study of several classification metrics and their performances on data," *World J. Adv. Eng. Technol. Sci.*, vol. 8, no. 1, pp. 308–314, 2023, doi: 10.30574/wjaets.2023.8.1.0054.

[25] M. Midha, A. K. Jain, V. Sharma, Shubham, G. Kaur, and D. Banerjee, "Enhanced car damage classification: A fusion of deep learning and random forest methods," in *2024 Asia Pacific Conference on Innovation in Technology (APCIT)*, Chitkara University Institute of Engineering and Technology, Chitkara University,Punjab,India; Chandigarh University,Department of Computer Science and Engineering,Punjab,India; Graphic Era Hill University,Computer Science and Engineering: IEEE, 2024, pp. 1–6. doi: 10.1109/apcit62007.2024.10673662.

[26] C. Md.Sharafat, R. Md.Naimur, S. Md.Sujon, S. Md.Abu, K. H. Mahmud, and B. Hafsa, "GIS-based landslide susceptibility mapping using logistic regression, random forest and decision and regression tree models in Chattogram District, Bangladesh," *Heliyon*, vol. 10, no. e23424, pp. 1–19, 2024, doi: 10.1016/j.heliyon.2023.e23424.

[27] A. M. Carrington *et al.*, "Deep ROC analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 329–341, 2023, doi: 10.1109/TPAMI.2022.3145392.

[28] M. Marcus and W. Beck, "Package 'NeuralNetTools.'" Accessed: Mar. 19, 2024. [Online]. Available: https://cran.r-project.org/package=NeuralNetTools/NeuralNetTools.pdf

[29] H. O. Tuffour *et al.*, "Pedotransfer functions for estimating saturated hydraulic conductivity of selected benchmark soils in Ghana," *Asian Soil Res. J.*, no. May, pp. 1–11, 2019, doi: 10.9734/asrj/2019/v2i230046.

[30] H. Shahabi *et al.*, "Flash flood susceptibility mapping using a novel deep learning model based on deep belief network, back propagation and genetic algorithm," *Geosci. Front.*, vol. 12, no. 3,

p. 101100, 2021, doi: 10.1016/j.gsf.2020.10.007.

[31] S. Karami, F. Saberi-Movahed, P. Tiwari, P. Marttinen, and S. Vahdati, "Unsupervised feature selection based on variance-covariance subspace distance," *Neural Netw.*, vol. 166, pp. 188–203, 2023, doi: 10.1016/j.neunet.2023.06.018.

[32] B. Pradhan, S. Lee, A. Dikshit, and H. Kim, "Spatial flood susceptibility mapping using an explainable artificial intelligence (XAI) model," *Geosci. Front.*, vol. 14, no. 6, p. 101625, 2023, doi: 10.1016/j.gsf.2023.101625.

[33] N. F. A. Aziz, N. Ya'acob, A. L. Yusof, and H. Omar, "Statistical Analysis for Forest Fire Factors Using Geography Information System (GIS) and Remote Sensing Imagery," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 45, no. 2, pp. 177–190, 2025, doi: 10.37934/araset.45.2.177190.

[34] S. Moragues, M. G. Lenzano, P. Jeanneret, V. Gil, and E. Lannutti, "Landslide susceptibility mapping in the Northern part of Los Glaciares National Park, Southern Patagonia, Argentina using remote sensing, GIS and frequency ratio model," *Quat. Sci. Adv.*, vol. 13, no. July 2023, pp. 1–18, 2024, doi: 10.1016/j.qsa.2023.100146.

[35] S. Hitouri *et al.*, "Gully erosion mapping susceptibility in a Mediterranean environment: A hybrid decision-making model," *Int. Soil Water Conserv. Res.*, no. xxxx, 2023, doi: 10.1016/j.iswcr.2023.09.008.