https://doi.org/10.5194/egusphere-2023-1417
Preprint. Discussion started: 10 August 2023

Addressing Class Imbalance in Soil Movement Predictions

Praveen Kumar[1], Priyanka Priyanka[1], Kala Venkata Uday[2], Varun Dutt[1]

[1]*Applied Cognitive Science Lab, Indian Institute of Technology Mandi, Himachal Pradesh, 175075, India*
[2]*Geotechnical Engineering Lab, Indian Institute of Technology Mandi, Himachal Pradesh, 175075, India*

Corrospondance to: Praveen Kumar (dr.praveenkumar.ml@gmail.com)

Abstract

Landslides threaten human life and infrastructure, resulting in fatalities and economic losses. Monitoring stations provide valuable data for predicting soil movement, which is crucial in mitigating this threat. Accurately predicting soil movement from monitoring data is challenging due to its complexity and inherent class imbalance. This study proposes developing machine learning (ML) models with oversampling techniques to address the class imbalance issue and develop a robust soil movement prediction system. The dataset, comprising two years (2019-2021) of monitoring data from a landslide in Uttarakhand, was split into a 70:30 ratio for training and testing. To tackle the class imbalance problem, various oversampling techniques, including Synthetic Minority Oversampling Technique (SMOTE), K-Means SMOTE, Borderline SMOTE, Support Vector Machine SMOTE, and Adaptive SMOTE (ADASYN), were applied to the dataset. Several ML models, namely Random Forest (RF), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (Light GBM), Adaptive Boosting (AdaBoost), Category Boosting (CatBoost), Long Short-Term Memory (LSTM), Multilayer Perceptron (MLP), and dynamic ensemble models, were trained and compared for soil movement prediction. Among these models, the dynamic ensemble model with K-Means SMOTE performed the best in testing, with an accuracy, precision, and recall rate of 99.68% each and an F1-score of 0.9968. The RF model with K-Means SMOTE stood out as the second-best performer, achieving an impressive accuracy, precision, and recall rate of 99.64% each and an F1-score of 0.9964. These results show that ML models with class imbalance techniques have the potential to significantly improve soil movement predictions in landslide-prone areas.

Keywords: Soil Movement Prediction; Class Imbalance; Oversampling; Machine Learning; Landslide Prone Areas.

1. Introduction

Landslides pose a significant threat to infrastructure, resulting in numerous fatalities and substantial economic losses each year (Parkash, 2011). These destructive events occur globally, particularly in hilly and mountainous regions, driven by gravity and characterized by the movement of large rocks, debris, and soil (Crosta, 1998). Factors such as heavy rainfall, earthquakes, and the impacts of climate change contribute to the occurrence and severity of landslides (Crosta, 1998).

Monitoring, predicting, and warning people about slope movements in landslide-prone areas are crucial for mitigating landslide risks. Advanced technologies like GPS, LIDAR, GIS, and remote sensing have proven effective for assessing and analyzing slope failure hazards (Ray et al., 2020). However, their high cost and the need for specialized expertise limit their accessibility, especially in developing countries where cost-effective IoT technologies are necessary (Pathania et al., 2020).

Machine learning (ML) models have been extensively studied for predicting soil movement in landslide-prone areas (Kumar et al., 2021a; Kumar et al., 2021b). This prediction problem could be divided into classification and regression tasks. The classification task aims to predict the degree of soil movement using various ML models. On the other hand, the regression task involves estimating the acceleration or displacement of soil under observation.

One common challenge in landslide prediction is a class imbalance, where certain classes have significantly more data samples than others. This imbalance can adversely affect the performance of ML models. To address class imbalance issues, techniques such as Synthetic Minority Oversampling Technique (SMOTE), K-Means SMOTE, Support Vector Machine SMOTE (SVM-SMOTE), Borderline SMOTE, and Adaptive Synthetic Minority Oversampling Technique (ADASYN) are employed to balance the dataset (Chawla et al., 2002; Douzas et al., 2018; Tang et al., 2008; Han et al., 2005; He et al., 2008).

Several researchers have dedicated their efforts to addressing class imbalance problems in ML. Notably, Chawla et al. (2002) introduced the SMOTE, Douzas et al. (2018) devised the K-Means SMOTE, Tang et al. (2008) developed the SVM-SMOTE, Han et al. (2005) proposed the Borderline SMOTE, and He et al. (2008) introduced the Adaptive Synthetic Minority Oversampling Technique (ADASYN). These techniques were developed to generate synthetic data and balance imbalanced datasets.

The field of soil movement prediction requires further investigation, particularly considering the complexities associated with a class imbalance in the datasets. Despite extensive research on ML models' predictive abilities for soil movement in landslides, there still needs to be more understanding regarding how class imbalance affects the models' performance and accuracy. This study aims to bridge this knowledge gap by examining different approaches to tackle class imbalance and exploring diverse ML models to improve the prediction of soil movement. Various multivariate classification models, including Random Forest (RF), Adaptive Boosting (AdaBoost), Extreme Gradient Boosting (XGBoost), Light Gradient Boosted Machine (Light GBM), Category Boosting (CatBoost), Long Short-Term Memory (LSTM), Multilayer Perceptron (MLP), and an ensemble of RF, AdaBoost, XGBoost, Light GBM, and CatBoost are developed to predict soil movement when coupled with class imbalance techniques (Kumar et al., 2019; Semwal et al., 2022; Wu et al., 2020; Pathania et al., 2021; Zhang et al., 2022; Sahin, 2022; Kumar et al., 2020).

This study delves into the field of soil movement prediction, making significant advancements by developing specialized ML models and techniques tailored to this domain. A notable aspect that has received limited attention in the existing literature is the challenge of class imbalance in landslide datasets. While previous research has primarily focused on ML models for soil movement prediction, this work addresses the issue of imbalanced data head-on. Multiple variants of the SMOTE and other balancing strategies are introduced and implemented to enhance the efficacy and accuracy of the ML models.

Additionally, this research explores using cost-effective Internet of Things (IoT) technologies in developing regions to improve the investigation and assessment of landslide hazards. The dataset used in this study spans two years, from June 2019 to June 2021, and was collected by an inexpensive IoT monitoring station in Uttarakhand, India. This real-world dataset captures the distinctive characteristics and patterns of soil movements prevalent in the landslide-prone area. By employing a comprehensive methodology, this work advances soil movement prediction and effectively addresses the challenge of class imbalance. It commences with a thorough overview of the collected data, emphasizing the measured weather and soil-related factors. Various SMOTE variants and other balancing techniques are employed to rectify the class imbalance, resulting in the generation of synthetic samples and ensuring a balanced representation of soil movement classes. The intricate correlations and patterns in the soil movement data are captured using a variety of ML models, including RF, AdaBoost, XGBoost, Light GBM, CatBoost, MLP, LSTM, and a dynamic ensembling of RF, AdaBoost, XGBoost, and CatBoost. Overall, this study's findings show potential for accurately reducing landslide risks, increasing the accuracy of landslide prediction, and encouraging the use of cost-effective IoT technologies in landslide-prone locations.

2. Background

Several techniques have been proposed to address the challenge of learning from imbalanced datasets, where the classification categories are not evenly represented. For example, Chawla et al. (2002) proposed the SMOTE, which involves generating synthetic minority class examples to balance the dataset. SMOTE has been shown to improve model performance compared to only undersampling the majority class. Douzas et al. (2018) introduced K-Means SMOTE, a method that combines SMOTE with k-means clustering to effectively overcome imbalances between and within classes without generating unnecessary noise. Tang et al. (2008) modified SVMs by incorporating different rebalancing heuristics, such as cost-sensitive learning, over-sampling, and undersampling. Among the variations of SVM, the granular SVMs-repetitive undersampling model (GSVM-RU) has been found to be the most effective and efficient. Additionally, Han et al., (2005) developed a Borderline SMOTE method that focuses on oversampling only the minority examples near the class boundary. Experimental results indicate that Borderline SMOTE1 and Borderline SMOTE2 outperform SMOTE and random oversampling methods in terms of true positive rate and F-value. Lastly, He et al. (2008) developed the ADASYN, which addresses class imbalance by generating more synthetic data for minority class examples that are harder to learn. ADASYN reduces bias and adaptively shifts the classification decision boundary toward challenging examples. Simulation analyses have demonstrated the effectiveness of ADASYN across various evaluation metrics. These techniques offer valuable approaches to mitigate the impact of imbalanced data in classification tasks. These class imbalance techniques have limited exploration and application for landslide datasets. Existing studies primarily focus on the general imbalanced dataset scenario but need to consider the unique characteristics and challenges associated with landslide datasets. Therefore, research is required for systematic studies that compare the performance and effectiveness of techniques such as SMOTE, K-Means SMOTE, Borderline SMOTE, and ADASYN in the specific context of soil movement prediction across various evaluation metrics. By bridging this literature gap, we can enhance the accuracy and reliability of models for predicting soil movement in landslide-prone areas and contribute to improved landslide risk mitigation strategies.

Several researchers developed the various ML models to predict soil movement and prediction problems in other fields (Kumar et al., 2019; Semwal et al., 2022; Wu et al., 2020; Pathania et al., 2021; Zhang et al., 2022; Sahin, 2022; Kumar et al., 2020). For example, Kumar et al. (2019) developed the ensemble of ML models (RF, Bagging, Stacking, and Voting) for predicting soil movement at the Tangni landslide in Uttarakhand, India. These models were compared with Sequential Minimal Optimization (SMO) and Autoregression (AR). The results indicate that the ensemble models outperformed the SMO and AR models in predicting soil movement. Furthermore, Semwal et al. (2022) developed the SMOreg, Instance-based Learning (IBk), RF, Linear Regression (LR), MLP, as well as ensemble ML models to predict root tensile strength for different vegetation species. The results show that the MLP performed better than the other models, providing more accurate predictions of root tensile strength. Next, Wu et al. (2020) developed the decision tree (DT) with AdaBoost and bagging ensembles for mapping the susceptibility of landslides in Longxian County, Shaanxi Province, China. Researcher developed the technique with ensemble of Alternating Decision Tree (ADTree) with Bagging and AdaBoost to map landslide susceptibility. The results revealed that ensemble of ADTree and AdaBoost model performed better than the individual ADTree model and ensemble of ADTree and Bagging model. Similarly, Pathania et al. (2021) developed a novel ensemble gradient boosting model, called SVM-XGBoost, for soil movements warning at Gharpa landslide, Mandi, India. They compared the performance of SVM-XGBoost with other models such as individual SVMs, DTs, RF, XGBoost, Naïve Bayes (NB), and different variants of XGBoost. The results showed that the SVM-XGBoost model performed better than other models in soil movement prediction. In their research, Kumar et al. (2021b) directed their attention toward predicting soil movement, specifically at the Tangni landslide site in India. To enhance the accuracy of their predictions, they explored various variants of Long Short-Term Memory (LSTM) models. They introduced a novel ensemble approach called BS-LSTM, which combined bidirectional and stacked LSTM models. The findings of their study indicated that the BS-LSTM model outperformed the other LSTM variants in accurately predicting soil movement. Similarly, Zhang et al. (2022) conducted a study to assess the susceptibility of landslides using gradient-boosting ML techniques coupled with class-balancing methods. Their investigation specifically focused on the aftermath of the 2018 Hokkaido earthquake and employed diverse datasets and methodologies to predict the susceptibility of specific areas prone to landslides. Compared to well-established models such as XGBoost and Light GBM, the proposed model showcased superior performance in accurately assessing landslide susceptibility. Furthermore, Sahin (2022) developed multiple ML models, including XGBoost, CatBoost, Gradient Boosting Machine (GBM), and Light GBM, to model the susceptibility of landslides. By leveraging a comprehensive landslide inventory map and relevant conditioning factors stored in a geodatabase, the study employed feature selection techniques and compared the predictive capabilities of ensemble methods with the widely used RF model. The results highlighted that CatBoost exhibited the highest predictive capability, followed by XGBoost, Light GBM, and GBM, while RF demonstrated comparatively lower predictive capability. The study used a geodatabase with a landslide inventory map and conditioning factors. Feature selection techniques were applied, and the performance of XGBoost, CatBoost, GBM, and Light GBM was compared to RF. The results revealed that CatBoost had the highest prediction capability, followed by XGBoost, Light GBM, and GBM. The literature gap in the context of soil movement prediction is the limited exploration and evaluation of ML models in combination with synthetic data generated by SMOTE techniques. While various ML models, such as ensemble models (e.g., RF), neural networks models (MLP and LSTM), and gradient boosting ML models (e.g., AdaBoost, XGBoost, Light GBM, CatBoost), have been developed and applied for soil movement prediction, their utilization in conjunction with synthetic data generated by SMOTE techniques has received less attention in the literature. Incorporating SMOTE-generated synthetic data into the training process of these models can address the issue of class imbalance in landslide datasets and improve their performance in predicting soil movement. Therefore, further research is needed to investigate the effectiveness of these ML models when combined with SMOTE techniques in the context of soil movement prediction, thereby filling the existing literature gap.

The RF, AdaBoost, XGBoost, Light GBM, CatBoost, MLP, LSTM, and an ensemble of RF, AdaBoost, XGBoost, Light GBM, and CatBoost models were chosen to predict soil movement based on their proven effectiveness in previous research. RF is excellent at capturing complex relationships and has outperformed non-ensemble models in predicting debris flow and landslide susceptibility. AdaBoost has successfully predicted soil movement alerts ahead of time. At the same time, XGBoost and Light GBM have demonstrated their ability to achieve balanced and precise predictions, especially in earthquake-induced landslide susceptibility assessments. Among gradient-boosting models, CatBoost stands out for its superior prediction capability, making it a well-suited option for modelling landslide susceptibility. On the other hand, when it comes to predicting root tensile strength, MLP has demonstrated higher accuracy in its predictions. Additionally, LSTM, a robust recurrent neural network architecture, is particularly effective in capturing temporal dependencies and long-term patterns in sequential data. Collectively, these models offer a diverse set of capabilities that prove valuable in the prediction of soil movement.

3. Data Collection and Description

The dataset for predicting soil movement was collected from an actual landslide site in Uttarakhand, India. Spanning a duration of two years, from June 2019 to June 2021, this dataset holds valuable insights into the behaviour of soil in response to various environmental factors. To gather this data, a cost-effective landslide monitoring station (LMS) was carefully deployed at landslide. Equipped with a range of sensors, the LMS diligently recorded critical weather and soil-related parameters. Weather-wise, it diligently captured temperature readings in degrees Celsius, humidity levels as a percentage, rainfall measurements in inches per hour (in/hr), atmospheric pressure in millibars (mb), and even sunlight intensity in lux. These meticulous recordings shed light on the prevailing weather conditions experienced at the precise location of the landslide. In order to monitor the soil conditions with utmost precision, the LMS relied on an accelerometer sensor. An advanced sensor was utilized to measure the acceleration of the soil in three directions: Ax, Ay, and Az (in m/s^2). This provided valuable insights into the soil's movement and stability. Additionally, a gyroscope sensor was employed to capture the angular rotation of the soil along the Wx, Wy, and Wz axes (in degrees per second). This sensor enhanced the understanding of the soil's behaviour by accurately detecting its angular movements. Furthermore, the LMS was equipped with a capacitive soil moisture sensor, enabling it to measure the volumetric moisture content of the soil in percentage. The LMS transmitted all these twelve attributes, including weather parameters, soil g-force, angular rotation, and soil moisture content, to the cloud every ten minutes. The dataset obtained from the LMS consisted of approximately thirty-nine thousand data points, covering a wide range of environmental and soil-related attributes. Table 1 is showcasing the statistics for the recorded soil movement prediction parameters. For each attribute, the table provides the mean value, representing the average measurement, along with the standard deviation (Std. Dev.), indicating the variability of the data. The minimum and maximum values highlight the range of measurements observed, offering insights into the extreme values and overall data distribution.

Table 1. Summary statistics of recorded parameters for soil movement prediction dataset.

Parameter	Mean	Std. Dev.	Min	Max
Temperature (° C)	16.18	10.48	0.00	39.00
Humidity (%)	66.69	35.46	0.00	99.00
Rain (in/hr)	0.00	5.60	0.00	15.00
Pressure (mb)	1040.96	27.96	921.61	1065.41
Light (lux)	5025.35	10154.75	0.00	54612.00
Ax (m/s^2)	0.02	1.23	-28.02	40.25
Ay (m/s^2)	0.00	1.37	-100.08	100.08
Az (m/s^2)	0.00	2.28	-149.61	315.61
Wx (°/s)	0.00	15.86	-1994.51	1997.24
Wy (°/s)	0.00	15.85	-1998.05	1998.73
Wz (°/s)	0.00	6.95	-932.00	932.00
Moisture (%)	80.00	20.30	40.00	100.00

4. Methodology

4.1. Data Pre-processing

The sensors installed at the landslide locations experienced malfunctions, resulting in multiple missing values within the collected data. To address this issue, we employed a method to fill these gaps by replacing the missing values with the average values recorded at the corresponding timestamps during the previous week. By calculating the average values for parameters such as light intensity, humidity, temperature, and pressure from the same time periods in the preceding week, we obtained estimates to replace the skewed or missing data points.

4.2. Class Labeling

The dataset contained values for acceleration and angular rotation in three directions: x, y, and z. The changes in acceleration and angular rotation were calculated by subtracting the current values from the past values, allowing for the assessment of movement. To classify the movement data, four categories were defined: no movement, low movement, moderate movement, and high movement. These categories were determined based on standard deviation thresholds derived from the acceleration and angular rotation values. Specifically, values within ± 1 standard deviation from the mean were categorized as no movement, ± 2 standard deviations as low movement, ± 3 standard deviations as moderate movement, and values exceeding ± 3 standard deviations as high movement. This classification approach considered the variability in acceleration and angular rotation changes to determine the intensity of movement.

During the analysis, each timestamp was assigned to a movement class based on the class associated with the highest standard deviation observed in any acceleration or angular rotation element. If an individual element had the highest standard deviation at a specific timestamp, that timestamp was assigned to the corresponding movement class with the maximum standard deviation.

Table 2 presents the distribution of movement intensity within the dataset, which consisted of 38,900 data points. The table shows the percentage distribution of movement categories: high, moderate, low, and no movement. The majority of the dataset (97.8%) falls under the "No Movement" category, indicating a lack of significant movement. On the other hand, the high movement category represents only a small fraction (1.1%) of the dataset. This distribution highlights the class imbalance issue present in the dataset, which needs to be taken into account when developing a classification model for predicting soil movement.

Table 2. Class distribution of soil movement data points.

Movement Class	Number of Data Points	Percentage
High Movement	423	1.1%
Moderate Movement	146	0.7%
Low Movement	268	0.4%
No Movement	38063	97.8%

4.3. Sliding Window Packets

The sliding window packets technique involves dividing a given dataset into fixed-length subsequences or packets and their corresponding labels. To achieve this, a sequence length parameter is used to determine the length of each subsequence. The sliding window approach is then employed, where a window starts at the beginning of the dataset and moves through the data with a step size of 1. A subsequence of the specified length is extracted from the dataset at each window position. The label for prediction is taken from the next position after the window.

The sliding window packets technique aims to predict future values or events based on preceding subsequences. For instance, if the sequence length is set to five, the sliding window will select five consecutive values from the dataset as a subsequence at each step. The label for prediction will be the value at the sixth position. This process continues until the end of the dataset is reached, resulting in multiple subsequences and their respective labels. Once the packets are created, they are flattened to form a single feature vector. For instance, if the sequence length is five and the dataset has twelve features, each packet will contain sixty elements (5x12). This transformation allows the packets to be treated as individual samples with multiple features suitable for ML models. The primary purpose of creating these packets is to address prediction tasks involving sequences where the input data's order and dependencies are crucial. The model can effectively capture and learn patterns and relationships within the sequential data by utilizing the sliding window packets. The flattened packets generated using the sliding window technique are inputs in oversampling techniques.

4.4. Oversampling

In our analysis, we encountered a significant class imbalance issue in the labelled data. The "No Movement" class, which represents the majority of the data, had a large number of data points. On the other hand, the "High Movement" class, which represents the minority class, had only a minimal representation with just 1% of the total data. This class imbalance posed a challenge for building an effective classification model, as the skewed data distribution made it difficult to classify the minority class accurately.

To overcome the class imbalance challenge, we implemented several oversampling techniques, with a particular focus on SMOTE and its extensions (Chawla et al., 2002; Douzas et al., 2018; Tang et al., 2008; Han et al., 2005; He et al., 2008). SMOTE, which stands for Synthetic Minority Oversampling Technique, addresses the imbalance by generating synthetic data points for the minority class (Chawla et al., 2002). By utilizing the characteristics of existing samples from the minority class, we created new data points, thereby increasing the representation of the "High Movement" class. In addition to the standard SMOTE, we also explored other variations such as K-Means SMOTE (Douzas et al., 2018), SVM-SMOTE (Tang et al., 2008), and Borderline SMOTE (Han et al., 2005) to further enhance the balance of class distribution.

Furthermore, we utilized the ADASYN, an extension of SMOTE that explicitly addresses the classification boundary of the minority class (He et al., 2008). ADASYN assigns higher weights to the minority examples that are more challenging to classify, leading to the generation of additional artificial data points for these instances. By incorporating ADASYN into our oversampling strategy, we enhanced the balance of the class distribution further and improved the classification accuracy for all classes.

Fig. 1. illustrates the application of the K-Mean SMOTE technique for addressing the class imbalance. The figure depicts a scatter plot where the red crosses represent the minority class samples, while the black dots represent the majority class samples. The green crosses indicate the newly generated synthetic samples by the K-Mean SMOTE algorithm. The dashed line represents the decision boundary separating the two classes. K-Mean SMOTE operates by following two simple steps iteratively [8]. Firstly, it assigns each observation to the nearest cluster centroid among the k available. Secondly, it updates the position of the centroids so that they are positioned at the centre between the assigned observations. The information ratio (IR) shown in Fig. 1 helps K-Means SMOTE determine the appropriate amount of oversampling for the minority class, ensuring a balanced representation of the classes in synthetic samples. The value of k was selected 4 in this experiment.

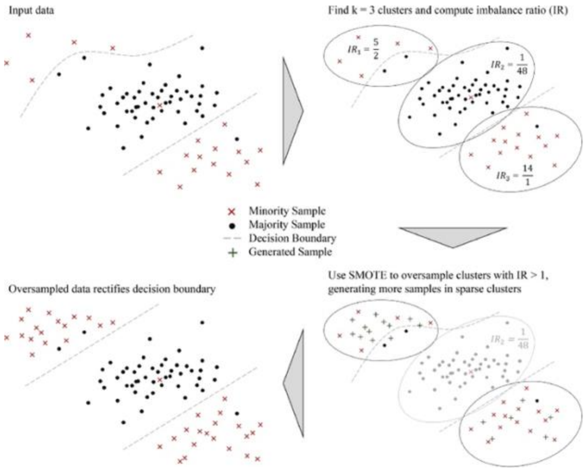
Input data
Find k = 3 clusters and compute imbalance ratio (IR)
$IR_1 = \frac{5}{2}$
$IR_2 = \frac{1}{48}$
$IR_3 = \frac{14}{1}$
× Minority Sample
● Majority Sample
-- Decision Boundary
+ Generated Sample
Oversampled data rectifies decision boundary
Use SMOTE to oversample clusters with IR > 1, generating more samples in sparse clusters
$IR_2 = \frac{1}{48}$

Figure 1. K-Means SMOTE effectively addresses within-class imbalance by oversampling safe areas (Douzas et al., 2018).

4.5. Machine Learning Models

Various models were employed to classify the soil movement. The specific models will be discussed in the following subsection. To evaluate the accuracy of these models, the dataset was divided into two groups: training data (70%) and testing data (30%). Random sampling was used to select 70% of the data points for training the classification models mentioned below, while the remaining 30% of the dataset was reserved for model evaluation.

4.5.1. AdaBoost

AdaBoost, also known as Adaptive Boosting, is a probabilistic classification meta-model designed to enhance the performance of ML models (Wu et al., 2020). It achieves this by combining the results of multiple weak learners, which are learning techniques with slightly better than random guessing capabilities. Through an adaptive process, AdaBoost adjusts subsequent weak models to focus on the cases that were misclassified by earlier models. This adaptive nature helps improve the overall accuracy of the model and reduces the risk of overfitting in certain situations. Although individual weak learners may not perform well on their own, their collective strength allows the final model to converge as a powerful learner capable of making more accurate predictions.

For the AdaBoost model, the number of trees determines the maximum number of weak models to be combined. Increasing the number of trees can improve the performance of the AdaBoost but may also increase the risk of overfitting. The learning rate controls the contribution of each weak model, with a higher learning rate giving more weight to each model. The maximum depth parameter limits the depth of individual weak

https://doi.org/10.5194/egusphere-2023-1417
Preprint. Discussion started: 10 August 2023

models, preventing them from becoming too complex and overfitting the data. Table 3 shows the range of hyperparameters for the AdaBoost model.

4.5.2. XGBoost

XGBoost is an ensemble ML model based on gradient boosting and uses decision trees (Chen and Guestrin, 2016). While deep neural networks excel in predicting unstructured data such as images and text, decision tree-based methods are considered superior for dealing with structured data. Decision trees are particularly effective when the data has a well-defined structure or specific features, making them a preferred choice for certain prediction problems involving structured information.

In the XGBoost model, the number of trees determines the number of boosting rounds or iterations. Increasing the number of trees can improve the performance of XGBoost but also increases computational complexity. The learning rate controls the step size during the boosting process and affects the model's convergence speed and generalization ability. The maximum depth parameter restricts the depth of the decision trees in the ensemble, preventing overfitting and promoting interpretability. Table 3 shows the range of hyperparameters for the XGBoost model.

4.5.3. Light GBM

Light GBM is a gradient-boosting framework that utilizes a decision-branching technique for various ML tasks such as ranking and classification (Ke et al., 2017). Unlike other boosting methods that divide the tree lengthwise or layerwise, Light GBM employs a leaf-wise approach, where the tree is divided leaf by leaf, selecting the best split at each step. The leaf-wise strategy employed in Light GBM offers advantages over traditional boosting techniques by reducing loss and improving accuracy. This strategy focuses on growing the tree leaf-wise rather than level-wise, resulting in a more effective learning process. Light GBM's efficient implementation is reflected in its speedy execution, earning it the moniker "Light" due to its fast performance.

In the Light GBM model, the number of trees determines the number of boosting rounds, and increasing this number has the potential to enhance the model's performance. The learning rate parameter influences the step size taken during the boosting process, striking a balance between convergence speed and model accuracy. To control the complexity of the model and mitigate the risk of overfitting, the maximum depth parameter limits the depth of the decision trees employed by Light GBM.

Table 3 provides an overview of the hyperparameters range for the Light GBM model, allowing for fine-tuning and customization to optimize its performance according to the specific dataset and problem at hand.

4.5.4. CatBoost

CatBoost, which stands for Category Boosting, is an ML model developed by Yandex and recently released as an open-source tool (Prokhorenkova et al., 2018). Its integration capabilities with popular frameworks like TensorFlow and Core ML make it highly adaptable across different platforms. CatBoost excels in handling diverse datasets, making it a valuable tool for addressing various industry problems. This model's exceptional performance and predictive accuracy contribute to its reputation as a leading choice in the field.

The loss function selection is important in the CatBoost model, as it can significantly impact its performance. Different loss functions, such as log, entropy, or hinge, are tailored to handle specific classification problems, and their selection can lead to varying results. To fine-tune the CatBoost model and optimize its performance for a given dataset, Table 3 presents the range of hyperparameters that can be adjusted according to the specific problem and data characteristics.

4.5.5. Random Forest

The RF technique, an ensemble learning method combining predictions from multiple decision trees (Breiman, 2001), was used to construct the regression or classification model. RF offers advantages such as handling relationships and non-linearities without requiring variable independence assumptions or dummy variables. It has shown exceptional performance in various industries and applications, including landslide prediction and site recognition.

The RF model's performance is optimized by adjusting parameters such as the number of trees (DTs), the splitting criteria (Gini or Entropy), and the maximum depth of the trees. These parameters control the model's robustness, accuracy, and complexity. Table 3 outlines the hyperparameter ranges for the RF model.

4.5.6. Multilayer Perceptron

The MLP (Multilayer Perceptron) is a well-known neural network architecture (Rosenblatt, 1961). It comprises interconnected layers, including input, hidden, and output layers. Neurons in the MLP calculate a weighted sum of inputs and pass it through an activation function to capture intricate relationships. The hidden layers extract valuable features from the input data. Dropout layers are employed to prevent overfitting by randomly deactivating neurons during training, enhancing the MLP's ability to generalize and reducing reliance on specific patterns. The MLP is a versatile and powerful approach for tackling classification problems.

In the MLP model, the look-back period determines the number of preceding time steps considered for prediction, influencing the model's capacity to capture temporal dependencies. The number of layers and nodes per layer govern the complexity of the network and its ability to learn intricate representations from the data. Table 3 provides an overview of the hyperparameter ranges for the MLP model.

4.5.7. LSTM

The LSTM is a recurrent neural network architecture (Hochreiter and Schmidhuber, 1997). It addresses the vanishing gradient problem and analyses sequential data by capturing long-term dependencies. The key strength of LSTM lies in its ability to process and predict sequences of varying lengths. LSTM networks selectively retain or forget information over time by utilizing memory cells and gating mechanisms. As a result, they have proven effective in various domains, including natural language processing, speech recognition, and time series forecasting.

In our LSTM model, we conducted experiments to examine the impact of different parameters on its performance. We investigated the influence of various LSTM unit sizes (32, 64, 128, and 256) to understand how the complexity of the model affects its capacity to capture patterns in the data. Additionally, we explored different activation functions (sigmoid, tanh, and ReLU) to assess their effect on the model's ability to learn complex relationships within the dataset. We selected the categorical cross-entropy loss function for multi-class classification tasks due to its suitability. Furthermore, we varied the look-back period from 3 to 10 to evaluate how it influences the model's ability to capture temporal dependencies. The range of hyperparameters for the LSTM model is presented in Table 3.

4.5.8. Dynamic Ensembling

Dynamic ensembling is a highly effective technique in ML that takes advantage of the adaptability and ongoing improvement of predictive models (Ko et al., 2008). It involves creating a versatile and continuously evolving ensemble by harnessing the strengths of multiple models, including RF, CatBoost, XGBoost, Light GBM, and AdaBoost. Traditionally, ensembling methods like bagging and boosting have focused on fixed ensembles. However, dynamic ensembling goes beyond this by introducing the ability to add or remove models based on their performance dynamically. In the case of dynamic ensembling with the models, as mentioned earlier, the monitoring criterion used is accuracy. Accuracy as the monitoring criterion ensures that the dynamic ensemble maintains a high level of accuracy in its predictions. If a model falls below a predefined accuracy threshold, it is considered underperforming and may be replaced to enhance the ensemble's overall performance.

Dynamic ensembling offers numerous advantages, including handling concept drift, where the underlying data distribution changes over time. By incorporating new models that capture updated patterns and relationships in the data, the dynamic ensemble can effectively adapt to concept drift and maintain accurate predictions.

The dynamic ensembling model utilized base models such as RF, CatBoost, XGBoost, Light GBM, and AdaBoost. Each base model was trained individually with the same default parameter settings as their standalone counterparts. The parameter values for each model were mentioned in Table 3.

Table 3. The range of hyperparameters varied in the models.

Model	Hyperparameter	Range of Hyperparameter
AdaBoost	Number of Trees	[15, 200] in steps of 10
	Learning Rate	[0.1, 2] in steps of 0.1
	Maximum Depth	[3, 33] in steps of 3
XGBoost	Number of Trees	[50, 1100] in steps of 50
	Learning Rate	[0.05, 0.55] in steps of 0.05
	Maximum Depth	[3, 33] in steps of 3
Light GBM	Number of Trees	[20, 400] in steps of 20
	Learning Rate	[0.05, 0.55] in steps of 0.05
	Max Depth	[3, 33] in steps of 3
CatBoost	Loss Function	Log, Entropy, Hinge
	Learning Rate	[0.1, 2] in steps of 0.1
	Max Depth	[3, 33] in steps of 3
RF	Number of Trees	[1, 100] in steps of 5
	Criteria	Gini, Entropy
	Maximum Depth	[1, 100] in steps of 5
MLP	Look-back Period	3 to 10
	Layers	[1, 3]
	Nodes Per Layer	[50, 250] in steps of 50
LSTM	Look-back Period	3 to 10
	LSTM Units	32, 64, 128, 256
	Activation Function	Sigmoid, tanh, ReLU

5. Model Execution, Minimization, and Handling Class Imbalance

A rigorous process was followed to develop an effective model for predicting the intensity of soil movement. The model was trained using grid search techniques, which systematically explored different combinations of hyperparameters to optimize its performance. The training phase utilized the labelled training data, split into a 70:30 ratio for training and testing purposes. One challenge encountered during the training process was the class imbalance issue. The number of samples available for the minority class was insufficient compared to the majority class. To address this, oversampling techniques were employed. By generating synthetic data points for the minority class, we were able to balance the dataset and mitigate the bias toward the majority class. Once the dataset was balanced, multiple ML models were trained using the training data. The primary objective was to optimize the models' parameters for improved performance. Each model underwent a grid search to identify the best configurations, a systematic approach that explores various parameter combinations. The models were then trained using the optimized parameters and evaluated on independent test data to assess their predictive performance. The evaluation primarily focused on accuracy metrics to determine how effectively the models predicted the intensity of soil movement.

6. Results

Table 3 presents the optimized hyperparameter values obtained through the grid search method for each model. These hyperparameters were carefully fine-tuned to ensure the best fit for the given data. In the case of XGBoost, the optimized values included 800 trees, a learning rate of 0.3, and a maximum depth of 9. After a thorough evaluation, these parameter settings were selected to improve the model's generalization capability and ability to capture intricate patterns within the data. The optimized values for the Light GBM model consisted of 220 trees, a hidden layer size of 0.25, and a maximum depth of 12. These settings were determined to enhance the model's performance in terms of both speed and accuracy.

Similarly, the AdaBoost model underwent hyperparameter optimization, selecting 25 trees, a learning rate of 1.7, and a maximum depth of 20. These parameter values were chosen to improve the model's adaptability and accuracy for classification tasks. The CatBoost model also went through optimization, leading to entropy selection as the loss function, a learning rate of 0.9, and a maximum depth of 20. These settings were explicitly chosen to maximize the model's accuracy and robustness. Likewise, the MLP model optimized its hyperparameters with a look-back period of 3, 2 layers, and 200 nodes per layer. These settings were selected to enhance the model's ability to capture complex relationships and improve classification accuracy. Similarly, LSTM has 128 units and tanh activation function. In the RF model, the optimized values were 45 for the number of trees, 25 for the maximum depth, and the evaluation criteria were set to "entropy." These values were chosen to maximize the model's accuracy and predictive power performance. Lastly, the dynamic ensembling model in this study incorporated the optimized RF, CatBoost, XGBoost, Light GBM, and AdaBoost models to improve the accuracy of landslide analysis predictions. By leveraging the strengths of these individually optimized models, as mentioned above, the dynamic ensembling model aimed to improve the accuracy and reliability of landslide analysis predictions.

Table 4. The best value of the hyperparameters calibrated from the training data.

Model	Hyperparameter	Best Value of Hyperparameter
AdaBoost	Number of Trees	25
	Learning Rate	1.7
	Maximum Depth	20
XGBoost	Number of Trees	800
	Learning Rate	0.3
	Maximum Depth	9
Light GBM	Number of Trees	220
	Learning Rate	0.25
	Maximum Depth	12
CatBoost	Loss Function	Entropy
	Learning Rate	0.9
	Maximum Depth	20
RF	Number of Trees	45
	Criteria	Entropy
	Maximum Depth	25
MLP	Look-back Period	3
	Layers	2
	Nodes Per Layer	200
LSTM	Look-back Period	5
	LSTM Units	128
	Activation Function	tanh

Table 5 presents the training results of different classification models combined with various oversampling techniques for landslide prediction. These results provide valuable insights into the performance of each model when trained on the dataset with oversampling. The RF model with K-Mean SMOTE emerges as the best model in training, achieving outstanding accuracy, precision, recall, and F1-score of 100% and 1, respectively. It demonstrates remarkable predictive capability by achieving perfect accuracy in both oversampling and non-oversampling scenarios.

Table 6 presents the test results of various classification models combined with different oversampling techniques for landslide prediction. Among them, the dynamic ensemble model utilizing the K-Mean SMOTE technique demonstrates exceptional performance in accurately predicting landslides on unseen data. It achieves impressive accuracy, precision, and recall rates of 99.68%, along with an F1-score of 0.9968. These outstanding results confirm the effectiveness of the dynamic ensemble approach when combined with K-Mean SMOTE for accurate landslide prediction. The best-performing model is highlighted in bold in Table 6.

Additionally, the RF model incorporating K-Mean SMOTE emerges as the second-best model in the test phase, showcasing high accuracy, precision, and recall rates of 99.64%, along with an F1-score of 0.9964. This result emphasizes the reliability and robustness of the RF model in tackling landslide prediction tasks.

Comparing the dynamic ensemble and RF models with other classification models and oversampling techniques, it becomes evident that the dynamic ensemble model with K-Mean SMOTE and the RF model with K-Mean SMOTE consistently outperform the rest, highlighting their effectiveness in accurately predicting landslides.

These findings underscore the importance of carefully selecting appropriate ML models and employing suitable oversampling techniques to address the class imbalance challenge in landslide prediction. They provide valuable insights into the performance and suitability of these models and techniques for enhancing landslide prediction accuracy, ultimately enabling proactive measures to mitigate landslide risks.

Table 5. The results of the ML models from training dataset.

Model	Oversampling Technique	Accuracy (in %)	Precision (in %)	Recall (in %)	F1-Score
AdaBoost	SMOTE	71.45	71.41	71.45	0.7142
	K-Means SMOTE	71.13	71.07	71.13	0.7105
	Borderline SMOTE	77.54	77.48	77.54	0.7748
	ADASYN	72.97	72.91	72.97	0.7276
	Without Oversampling	97.87	97.73	97.87	0.9779
CatBoost	SMOTE	99.85	99.85	99.85	0.9985
	K-Means SMOTE	99.86	99.86	99.86	0.9986
	Borderline SMOTE	99.78	99.78	99.78	0.9978
	ADASYN	99.85	99.85	99.85	0.9985
	Without Oversampling	99.72	99.72	99.72	0.9971
XGBoost	SMOTE	99.97	99.97	99.97	0.9997
	K-Means SMOTE	99.98	99.98	99.98	0.9998
	Borderline SMOTE	99.98	99.98	99.98	0.9998
	ADASYN	99.97	99.97	99.97	0.9997
	Without Oversampling	99.99	99.99	99.99	0.9999
Light GBM	SMOTE	99.87	99.87	99.87	0.9987
	K-Means SMOTE	99.91	99.91	99.91	0.9991
	Borderline SMOTE	99.97	99.97	99.97	0.9997
	ADASYN	99.88	99.88	99.88	0.9988
	Without Oversampling	99.87	99.87	99.87	0.9987
RF	SMOTE	100	100	100	1
	K-Means SMOTE	100	100	100	1
	Borderline SMOTE	100	100	100	1

	ADASYN	100	100	100	1
	Without Oversampling	100	100	100	1
MLP	SMOTE	90.32	90.49	90.32	0.9028
	K-Means SMOTE	60.16	73.24	60.16	0.5785
	Borderline SMOTE	95.66	95.72	95.66	0.9566
	ADASYN	88.54	88.81	88.54	0.8854
	Without Oversampling	97.81	96.43	97.81	0.9694
LSTM	SMOTE	68.73	68.81	68.73	0.6799
	K-Means SMOTE	79.51	79.89	79.51	0.7959
	Borderline SMOTE	86.40	86.66	86.40	0.8627
	ADASYN	77.47	77.72	77.47	0.7736
	Without Oversampling	97.88	95.80	97.87	0.9683
Ensemble	SMOTE	99.94	99.94	99.94	0.9994
	K-Means SMOTE	99.96	99.96	99.96	0.9996
	Borderline SMOTE	99.97	99.97	99.97	0.9997
	ADASYN	99.93	99.93	99.93	0.9993
	Without Oversampling	99.99	99.99	99.99	0.9999

Table 6. The results of the ML models from test dataset.

Model	Oversampling Technique	Accuracy (in %)	Precision (in %)	Recall (in %)	F1-Score
AdaBoost	SMOTE	71.57	71.54	71.57	0.7155
	K-Means SMOTE	71.03	70.89	71.03	0.7091
	Borderline SMOTE	77.65	77.57	77.65	0.7757
	ADASYN	72.83	72.77	72.83	0.7263
	Without Oversampling	97.59	97.41	97.59	0.9749
CatBoost	SMOTE	99.56	99.56	99.56	0.9956
	K-Means SMOTE	99.57	99.57	99.57	0.9957
	Borderline SMOTE	99.50	99.50	99.50	0.9950
	ADASYN	99.58	99.58	99.58	0.9957
	Without Oversampling	98.11	97.85	98.11	0.9787
XGBoost	SMOTE	99.63	99.63	99.63	0.9963
	K-Means SMOTE	99.62	99.62	99.62	0.9962
	Borderline SMOTE	99.57	99.57	99.57	0.9957
	ADASYN	99.64	99.64	99.64	0.9964
	Without Oversampling	98.09	97.63	98.09	0.9775
Light GBM	SMOTE	99.55	99.55	99.55	0.9955
	K-Means SMOTE	99.60	99.60	99.60	0.9960
	Borderline SMOTE	99.54	99.54	99.54	0.9954
	ADASYN	99.58	99.58	99.58	0.9958
	Without Oversampling	98.00	97.59	98.00	0.9770
RF	SMOTE	99.52	99.53	99.52	0.9952
	K-Means SMOTE	99.64	99.64	99.64	0.9964
	Borderline SMOTE	99.58	99.58	99.58	0.9958
	ADASYN	99.54	99.54	99.54	0.9953
	Without Oversampling	98.09	97.67	98.09	0.9763
MLP	SMOTE	89.89	90.07	89.89	0.8985
	K-Means SMOTE	59.84	73.31	59.84	0.5754
	Borderline SMOTE	95.44	95.53	95.44	0.9545
	ADASYN	88.19	88.44	88.19	0.8818
	Without Oversampling	97.61	96.00	97.61	0.9670
LSTM	SMOTE	69.02	69.09	69.01	0.6829

	K-Means SMOTE	78.56	78.98	78.55	0.7864
	Borderline SMOTE	86.23	86.52	86.22	0.8609
	ADASYN	76.78	77.06	76.77	0.7664
	Without Oversampling	97.79	95.63	97.78	0.9669
Dynamic Ensemble	SMOTE	99.58	99.58	99.58	0.9958
	K-Means SMOTE	**99.68**	**99.68**	**99.68**	**0.9968**
	Borderline SMOTE	99.55	99.55	99.55	0.9955
	ADASYN	99.58	99.58	99.58	0.9958
	Without Oversampling	98.20	97.83	98.20	0.9783

7. Discussion and Conclusions

In summary, the threat posed by landslides requires the development of effective prediction frameworks, although modelling the chaotic nature of natural data remains challenging. The analysed dataset exhibited a significant class imbalance, with the majority class dominating the samples. This distribution imbalance necessitated careful consideration and appropriate techniques to address the issue.

Various oversampling techniques, including SMOTE and its extensions (K-Means SMOTE, Borderline SMOTE, and SVM-SMOTE), were employed to tackle the class imbalance. ADASYN, which focuses on the minority class boundary, effectively generated synthetic data points, and improved the class distribution balance.

Multiple classification models, such as ADABoost, XGBoost, Light GBM, CatBoost, RF, MLP, LSTM, and a dynamic ensemble, were evaluated for predicting soil movement. The hyperparameters of each model were optimized using a grid search approach. The dynamic ensemble with K-Mean SMOTE and RF with K-Mean SMOTE emerged as the top-performing models, with the dynamic ensemble achieving slightly higher accuracy.

Combining K-Means SMOTE oversampling with the dynamic ensemble model yielded exceptional results, with high accuracy, precision, recall, and F1-score predicting soil movement. These outcomes demonstrate the effectiveness of oversampling techniques and the dynamic ensemble model in addressing class imbalance and improving landslide prediction accuracy.

This study emphasizes the importance of pre-processing, class labelling, addressing the class imbalance, and selecting suitable classification models for accurate soil movement prediction. The findings contribute to a better understanding of landslide risks and support the development of effective preventive measures.

However, there are several limitations to consider. The generalizability of the findings to other regions or geological conditions may be limited due to the specific dataset used. The synthetic data points generated through oversampling may only partially capture the complexity of real-world landslide occurrences. The choice of classification models and hyperparameter settings could introduce bias, and alternative configurations may yield different results. The study relied on historical data, potentially limiting its ability to account for future changes. Factors such as rainfall intensity, seismic activity, and human influences were not fully accounted for, suggesting the need for further research to enhance landslide prediction accuracy.

In future work, we plan to evaluate the performance of encoder-decoder models or transformer models on the class-imbalanced movement dataset. These models have demonstrated success in sequence-to-sequence tasks and could potentially improve classification accuracy and address class imbalance challenges. This experimentation will provide valuable insights into their suitability for analyzing and modeling imbalanced movement data.

Acknowledgements

We would like to acknowledge and express our sincere gratitude to the DST, India, and the DDMA Kangra (IITM/DDMA-Kan/KVU/357), Kinnaur (IITM/DDMA-Kinn/VD/345), and Mandi (IITM/DDMA-M/VD/325 and IITM/DDMA-M/VD/358) for their invaluable financial support towards this research project. We are also immensely grateful to the IIT Mandi for generously facilitating us with the necessary infrastructure, including

research facilities and computational resources, that have been instrumental in the successful execution of this study.

References

Breiman, L.: Random forests. *Machine learning*, 45, 5-32, 2001.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P.: SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357, 2002.

Chen, T., & Guestrin, C.: Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, pp. 785-794, 2016.

Crosta, G.: Regionalization of rainfall thresholds: an aid to landslide hazard evaluation. *Environmental Geology*, 35(2), 131-145, 1998.

Douzas, G., Bacao, F., & Last, F.: Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. *Information Sciences*, 465, 1-20.

Han, H., Wang, W. Y., & Mao, B. H.: Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In Advances in Intelligent Computing: International Conference on Intelligent Computing, ICIC 2005, Hefei, China, August 23-26, 2005, Proceedings, Part I 1, pp. 878-887, 2005.

He, H., Bai, Y., Garcia, E. A., & Li, S.: ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*, pp. 1322-1328, 2008.

Hochreiter, S., & Schmidhuber, J.: Long short-term memory. *Neural computation*, 9(8), 1735-1780, 1997.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y.: Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.

Ko, A. H., Sabourin, R., & Britto Jr, A. S.: From dynamic classifier selection to dynamic ensemble selection. *Pattern recognition*, 41(5), 1718-1731, 2008.

Kumar, P., Sihag, P., Sharma, A., Pathania, A., Singh, R., Chaturvedi, P., & Dutt, V.: Prediction of Real-World Slope Movements via Recurrent and Non-recurrent Neural Network Algorithms: A Case Study of the Tangni Landslide. *Indian Geotechnical Journal*, 51(4), 788-810 (2021a).

Kumar, P., Sihag, P., Chaturvedi, P., Uday, K. V., & Dutt, V.: BS-LSTM: an ensemble recurrent approach to forecasting soil movements in the real world. *Frontiers in Earth Science*, 9, 696-792 (2021b).

Kumar, P., Sihag, P., Pathania, A., Agarwal, S., Mali, N. C. P., Singh, R., ... & Dutt, V.: Landslide debris-flow prediction using ensemble and non-ensemble machine-learning methods. In *International Conference on Time Series and Forecasting* (Vol. 1), 2019.

Kumar, P., Priyanka, Pathania, A., Agarwal, S., Mali, N., Singh, R., ... & Dutt, V.: Predictions of weekly slope movements using moving-average and neural network methods: a case study in Chamoli, India. In *Soft Computing for Problem Solving 2019: Proceedings of SocProS 2019, Volume 2*, pp. 67-81, 2020.

Parkash, S.: Historical records of socio-economically significant landslides in India. *Journal of South Asia Disaster Studies*, 4(2), 177-204, 2011.

Pathania, A., Kumar, P., Priyanka, P., Maurya, A., Uday, K. V., & Dutt, V.: Development of an Ensemble Gradient Boosting Algorithm for Generating Alerts About Impending Soil Movements. In *Machine Learning, Deep Learning and Computational Intelligence for Wireless Communication: Proceedings of MDCWC 2020*, pp. 365-379, 2021.

Pathania, A., Kumar, P., Sihag, P., Chaturvedi, P., Singh, R., Uday, K. V., & Dutt, V.: A low-cost, sub-surface IoT framework for landslide monitoring, warning, and prediction. In *Proceedings of 2020 International conference on advances in computing, communication, embedded and secure systems*, 2020.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A.: CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 31, 2018.

Ray, R. L., Lazzari, M., & Olutimehin, T.: Remote sensing approaches and related techniques to map and study landslides. *Landslides Investig. Monit.*, 2020.

Rosenblatt, F.: Principles of neurodynamics. perceptrons and the theory of brain mechanisms. *Cornell Aeronautical Lab Inc Buffalo NY*, 1961.

Sahin, E. K.: Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. *Geocarto International*, 37(9), 2441-2465, 2022.

Semwal, T., Priyanka, P., Kumar, P., Dutt, V., & Uday, K. V.: Predictions of Root Tensile Strength for Different Vegetation Species Using Individual and Ensemble Machine Learning Models. In *Trends on Construction in the Digital Era: Proceedings of ISIC 2022*, pp. 87-100, 2022.

Tang, Y., Zhang, Y. Q., Chawla, N. V., & Krasser, S.: SVMs modeling for highly imbalanced classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1), 281-288, 2008.

Wu, Y., Ke, Y., Chen, Z., Liang, S., Zhao, H., & Hong, H.: Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. *Catena*, 187, 104396, 2020.

Zhang, S., Wang, Y., & Wu, G.: Earthquake-Induced Landslide Susceptibility Assessment Using a Novel Model Based on Gradient Boosting Machine Learning and Class Balancing Methods. *Remote Sensing*, 14(23), 5945, 2022.