IoT-Based Landslide Monitoring and Prediction Using Machine Learning

Sahithi K [1]*, *Bhanu Prakash* Saripalli[1], *Karthik* Guntha[2], *Kesava* Datta[1], *Navatej* y [1], *Rajavardhan* reddy[1] *and Vaishnavi* gandikota[1]

[1]School of Technology, Woxsen University, Hyderabad, Telangana
[2]Stanley College of Engineering and Technology for Women, Hyderabad, India

Abstract. Landslides are one of the most devastating natural disasters that result in massive human and infrastructural losses and economic inconveniences. To minimize these effects, it is essential to monitor and make early predictions. In this paper, the author introduces an IoT-based landslide monitoring and forecasting system that uses geotechnical and environmental sensors combined with machine learning algorithms. This system records the real-time data on the main parameters, which are the soil cohesion, intensity of rainfall, the angle of internal friction, the angle of slope, the slope height, and the factor of safety (FOS). These readings are sent through the IoT communication protocols to a cloud storage, pre-processed, and processed by an analytical processing platform. This paper has tested three machine learning algorithms, which include Multilinear Regression, Random Forest, and Decision Tree, to identify and forecast landslide occurrences. It also describes the system architecture, data collection process, feature engineering, and the model performance, giving a comparative analysis of the prediction accuracy of each algorithm. The proposed system integrates the IoT-based sensing with the solutions that are based on data to improve the early warning, enable informed decisions of hazard-management, and safeguard human life, infrastructure, and environment in zones of landslides.

1 Introduction

The most devastating natural hazards include landslides, which result in massive loss of human life, destruction of infrastructure, as well as serious degradation of the environment, especially in hilly areas and mining areas, where slope instability is a common occurrence. The conventional slope stability evaluation methods, despite their common usage, are based on periodic site observations and manual data analysis. Such traditional methods tend to be sluggish, reactive, and unable to respond to quick or sudden slope failures this making them ineffective in high-risk areas, which demand continuous monitoring and timely decision-making. It is possible that the developments in digital technologies recently have made it possible to include real-time sensing and intelligent data processing into landslide hazard monitoring. Computer-based systems made up of Internet of Things (IoT) are a set of

* Corresponding author: ssahithi90@gmail.com

distributed smart sensors and wireless communication devices that can constantly record high-resolution geotechnical and environmental measurements, including rainfall intensity, soil cohesion, internal friction angle, slope geometry, slope height, and factor of safety (FOS). Constant observation of these parameters will give early signs of slope deformation and instability, and it will be a good indicator used to inform prevention strategies, as opposed to emergency responses.

Machine learning (ML) has become an effective tool to predict landslides because it detects complex nonlinear patterns and multivariable relationships in large amounts of data. Multilinear regression, Decision Tree, and Random Forest are used in this study to predict slope instability using real-time sensor measurements. Multilinear regression offers linear interpretability among the input features and FOS, Decision Tree offers rule-based classification that has well-defined decision paths, and random forest enhances the strength of prediction by combining several tree-based estimators to minimize overfitting and erratic or missing data. This study is novel in that it establishes a complete IoT-based landslide monitoring system in tandem with a comparative machine learning predictive model of instant slope failure prediction in mines. In contrast to existing studies, which use a set of static geological data or single-model prediction methods, the proposed study considers continuous multi-sensor data collection. It tests the efficiency of three different ML models to determine the most accurate algorithm for deploying an early warning system. The suggested methodology will result in a shift in landslide risk management, where the traditional, reactive-only method is no longer applicable, but an automated and data-driven early warning system. This study further explores how critical geotechnical parameters influence model behavior and demonstrates the particle significance of IoT-ML integration in enhancing slope safety and disaster resilience.

2 Literature Survey

In recent years, the geotechnical monitoring practice has shown that artificial intelligence and the Internet of Things (IoT) technologies have great potential to be used as tools to evaluate slope stability and predict landslides. Landslides and slope instability are a global issue that has been a major threat to key infrastructure, mining, and transport systems. The Internet of Things (IoT), combined with Machine learning and intelligent sensing led to great advancements in developing sophisticated monitoring and early-warning systems.

The work discusses historical research that resulted in the development of data-driven slope stability measurement and landslide prediction based on IoT and Machine Learning. According to Onyelowe et al., (2025) explored slope behavior to predict geophysical flow by utilizing the hybrid machine learning combinations and showed that ensemble-based structure greatly improves predictive accuracy of the multi-faceted slope system [1]. Indukala et al. (2024) developed an IoT-based microseismic sensing and monitoring system to detect landslides in real-time and actively assist in monitoring the results and early notifications, involving the use of distributed sensors and a cloud-based data-processing system [2]. Das et al. (2022) implemented an IoT system using machine learning to enhance slope stability in open-cast mines, showing that data-driven models can be effective in the mining industry to assess safety [3]. Time domain reflectometry was used by Lin et al. (2019) to measure subsurface slope deformation through shear failure and demonstrated its appropriateness in identifying early-stage instability [4]. Park et al. (2019) studied slope failure observation with displacement sensors on a model slope and demonstrated that the patterns of progressive deformation are effectively observed with displacement-based measurements [5]. Recently, Babu Thoppil et al. (2023) proposed a highly developed sensor-based landslide and earthquake detection system that uses machine learning and computer vision and improves detection accuracy by multimodal data fusion [6]. Moayedi et al. (2019) examined slope stability control using remote sensing data and fuzzy logic, demonstrating its reliability in scenarios with uncertainty and limited data [7]. Dong et al. (2021) suggested a deep-learning-based deformation prediction model based on real-time monitoring data and showed high efficiency in predicting unstable slope behavior [8]. According to Sharma et al. (2023), sharp landslide detection, monitoring, and early warning in an IoT-cloud setting constitute an ensemble learning framework discussed in the article, which confirms ensemble models as more robust and more accurate than individual ones [9]. Lastly, Pathania et al. (2020) created a low-cost subsurface IoT landslide monitoring, warning, and prediction framework that focuses on affordability and scalability to deploy in resource-constrained areas [10].

2.1 Methodology

The proposed methodology is a scientific approach to designing an IoT-based landslide monitoring and prediction system with the help of machine learning algorithms. It is a combination of software and hardware development, lab testing, data visualization, and predictive modelling to offer reliable and real-time slope stability analysis, as represented in fig.1.

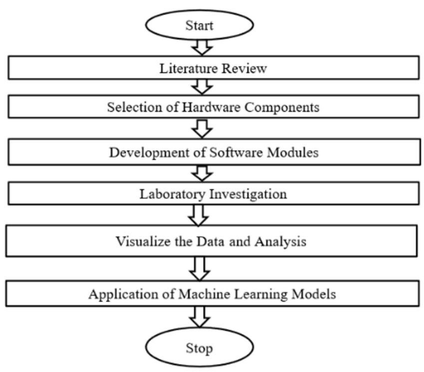

Fig. 1 Methodology for Machine Learning Development in Slope Stability Assessment

The methodology starts with determining the problem of the research and carefully conducting a literature review to learn the available technologies, sensing systems, and methods of analysis applied to assess slope stability and predict landslides. The hardware components are then chosen according to the performance requirements and suitability to the environment, and software modules are developed to acquire real-time data, to communicate wirelessly, and to visualize on the cloud. These modules are connected with IoT sensor nodes in order to provide constant data transmission and mass surveillance. Laboratory experiments that are conducted after hardware and software integration involve a controlled slope model. The major parameters are cohesion of soil, slope angle, rainfall intensity, and internal friction angle, which are always monitored in order to analyse their effects on slope stability and Factor of Safety (FOS). The obtained data are graphically represented and analyzed to find out trends, patterns, and correlations associated with slope failure. The recorded data is then used to train and test machine learning models (Multilinear Regression, Decision Trees, and Random Forest) to predict FOS and determine the susceptibility of landslides to occur under different circumstances. The models can be used to determine important contributing factors, increase predictive accuracy, and enhance their early-warning abilities. A combination of IoT and ML would lead to a more precise, automated, and trustworthy system of slope monitoring and management of disasters.

IoT sensors are also placed at landslide-prone areas to measure the much-needed geotechnical and environmental parameters that include soil cohesion, rainfall intensity, internal angle of friction, slope angle, slope height, and factor of safety (FOS). Such parameters are measured with soil moisture sensors, rain gauges, tilt sensors, vibration sensors, and geotechnical probes located on slopes that are not stable. The gathered information is sent wirelessly via communication modules to a central gateway, and consequently, the information is sent to a cloud-based analytics system. Field conditions pose such problems as sensor noise, loss of readings in the case of rough weather conditions or power off, abrupt changes during rainfall or mining. Combining these conditions of operations and geotechnical parameters, the offered machine-learning architecture is highly consistent with the real implementation of IoT, making it more practical and more correct in predicting. The authors use various approaches: Multilinear Regression, Decision Tree, and Random Forest to predict and model landslide susceptibility. The Multilinear Regression is a model of linear relationships amongst slope angle, cohesion, rainfall, and FOS as a baseline. Decision Tree recognizes nonlinear patterns and hierarchical rules that can be applied in complex slope stability. The ensemble of several decision trees, the Random Forest, is a better prediction tool due to its ability to decrease overfitting and increase the overall accuracy. These algorithms collectively have a complementary set of strengths towards effective landslide prediction.

3 System Architecture of IoT-based Monitoring

The IoT-based monitoring of landslides (Fig. 2) is developed as a smart and integrated structure that will provide the opportunity to monitor the slope stability at all times and provide an early warning. It integrates highly sophisticated sensors, IoT communication units, and cloud-based data processing and machine learning algorithms. The system is based on jointed layers with each layer sensing, transmitting data, predictive analytics and decision support.

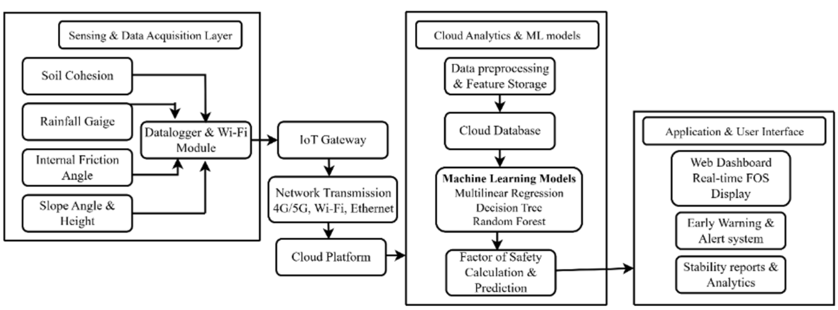

Fig. 2 System Architecture for a Real-Time Slope Stability Monitoring and Early Warning Platform

The system starts with the sensing and data acquisition level, which records real-time geotechnical and environmental parameters. Sensors of soil cohesion, rain gauges, internal friction angle, slope angle, and slope height constantly measure the variables that affect slope stability. These devices are used to collect data, which is aggregated by a Wi-Fi module so that it is properly measured and reliably transferred. The layer of IoT communication and network relays the field data to cloud-based systems. An IoT gateway is an interface that connects sensors and the cloud based on 4G/5G, Wi-Fi, or Ethernet. Such a design will support the flow of wireless data to support remote monitoring continuously, and the cloud-based storage will be resilient and scalable to support analytics and model development. The Cloud Analytics and Machine Learning layer is used to pre-process and extract features to eliminate noise and enhance the quality of its data before data storage in the cloud database. Multilinear Regression, Decision Tree, and Random Forest machine learning algorithms are used to analyse the connection between variables that include rainfall, soil cohesion, and the slope angle. These models are used to estimate the Factor of Safety (FOS) and estimate the likelihood of slope failure, and the best performance of these models is on complex and non-linear datasets. The application and user interface layer provides a visualisation tool, real-time trend of the parameters, FOS values, and early warning messages in SMS, email, or alarm to facilitate effective geotechnical assessment and disaster management.

Table 1: Parameters used for landslide prediction

Parameter	Description	Sensor/Computation
Soil Cohesion	Shear strength of soil contributing to slope stability	Geotechnical probes/field data
Rainfall Intensity	The amount of rain over time affects slope saturation	Rain gauge
Internal Angle of Friction	Soil shear resistance property	Geotechnical analysis
Slope Angle	Inclination of the slope influencing stability	IoT tilt sensor
Slope Height	Vertical height of the slope affects mass movement	Field measurement
Factor of Safety (FOS)	Stability index representing slope safety	Computed from geotechnical parameters

The main geotechnical and environmental parameters that have been considered for the current study have been summarized in Table 1. The parameters, including soil cohesion, intensity of rainfall, internal angle of friction, slope angle, slope height, and FOS, are vital when estimating the occurrence of landslides and slope stability. Making them available in the form of a table facilitates explaining their importance and the origin of each of the measurements in the suggested monitoring system.

3.1 Dataset for Landslide Susceptibility Modelling

The dataset used for landslide monitoring and prediction was developed by collecting key geotechnical and environmental parameters that directly influence slope stability. Table 2 shows the parameters considered, their symbols, units, and the range of their values. Such parameters are important inputs to the machine learning models that were created to predict the Factor of Safety (FOS) of slopes and the IoT-enabled monitoring system.

Table 2: Geotechnical Parameters for Deterministic Slope Stability Analysis

Parameter	Symbol	Unit	Minimum	Maximum
Rainfall	*R (or) P*	mm	50	100
Soil Cohesion	c	kPa	0	150
Internal Angle of Friction	ϕ	θ	0	45
Slope Angle	β	θ	16	53
Height	H	m	100	350
Factor of Safety	*FOS*	-	0.6	2

Rainfall in mm is a force that raises the pressure of pore water and is one of the major causes of slope failure. The soil strength and stability are determined by soil cohesion (50-100 kPa) and internal friction angle (0-45°). The slope angle (16-53) and slope height (100 350 m) also determine the probability of failure. Factor of Safety (0.62-2) means a stable or unstable slope. These parameters are geotechnical and environmental parameters that are critical in slope behaviour. They can be tracked in real time with the aid of IoT sensors, and machine learning models apply them to predict the area of landslides more precisely.

4 Machine Learning Models

In the proposed landslide monitoring and prediction system based on the Internet of Things (IoT), the analysis of slope behaviour patterns and predicting the occurrence of landslides are based on the utilization of machine learning (ML) algorithms. The models are trained through field sensor and laboratory data on the intensity of rainfall, cohesion of soil, angle of internal friction, height of slope, and slope angle that directly affect the Factor of Safety (FOS). Combining the data of the IoT and the ML analysis allows predicting the slope stability based on the data. The reason why this study employs multilinear regression, decision trees, and random forests is that they are efficient in handling mixed data. Multilinear regression is a statistical model that identifies linear relationships among multiple independent variables and a continuous outcome through the Ordinary Least Squares (OLS) method. It gives an equation-based prediction that is interpretable and can be used in the identification of linear patterns. Multilinear Regression is used to predict slope stability as the base model by making the relationship between rainfall, cohesion of soil, slope angle, and FOS. It shows the effect of each parameter on stability, but fails in the challenging task of identifying the complex, non-linear interaction that is usually involved in landslide behaviour. A Decision Tree is a non-linear algorithm that divides the data recursively based on a measure, such as Gini or Entropy, into a hierarchical set of interpretable decisions, but can be overfit without adequate pruning. A Decision Tree is a non-linear regression and classification model that divides data into hierarchical branches according to decision rules arising out of input parameters. Conditions (as in rainfall thresholds) are represented as internal nodes, and the predicted stability, or FOS level, is given as leaf nodes. It has a very interpretable structure, can cope with missing data, can capture non-linear relationships that are very complicated, and it is better at prediction than multilinear Regression. Random Forest combines many decision trees through the processes of bootstrap sampling and random features to improve predictive accuracy and reduce variance while providing good performance on noisy or otherwise poorly structured datasets, as well as handling non-linearity and very high-dimensional space. By creating numerous decision trees on random portions of the dataset along with features, Random Forest provides improved accuracy in predictions by averaging the outputs of many decision trees to reduce the chance of overfitting while enhancing the generalised ability of Random Forest. Random Forest works well with IoT data, especially in high dimensions and with noise, and can withstand minor sensor interferences. Feature Importance is an added benefit of using Random Forest because it allows for identifying those geotechnical parameters that most affect landslides.

5 Results and Discussion

The research presented investigates an Internet of Things (IoT) based system that monitors and predicts landslides through comparisons of three different machine learning algorithms to see how they perform in predicting when landslides will occur based on a variety of geological and environmental factors.

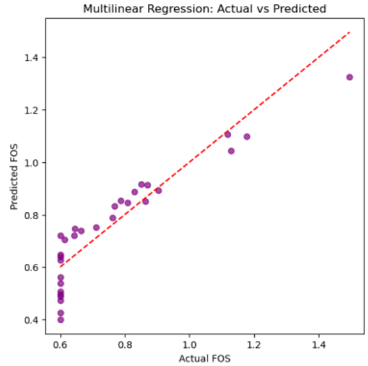

Fig. 3 Multilinear Regression Model Performance

The Multilinear Regression's predicted Factor of Safety (FOS) ranged from 0.4, 0.6, 0.8, and 1.0, whereas the actual FOS ranged from 0.6 to 1.4, as illustrated in Figure 3. While the Multilinear Regression produced relatively good approximations of the actual data set with respect to their general trends, they were significantly deviant from the actual data due in large part to the fact that they estimated an FOS below 0.6, while the observed, established minimum was 0.6. This serves to highlight the limitations of using a linear regression model in contexts that involve non-linearity within the characteristics of the slopes experiencing landslides.

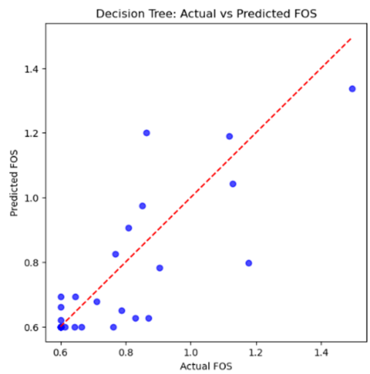

Fig. 4 Decision Tree Regression

The Decision Tree produced predicted values of the FOS that had greater predictive accuracy compared to the previously illustrated Multilinear Regression, producing values matching the actual values recorded (i.e., within the ranges of 0.6 to 1.4), and, as illustrated in Figure 4, demonstrates that the Decision Tree model was able to learn the rules for creating true FOS statistics for this specific data set.

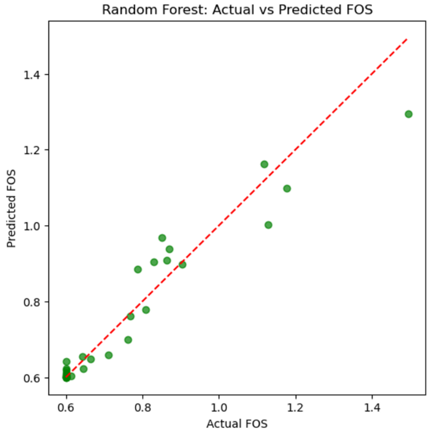

Fig. 5 Random Forest Regression

The Random Forest produced predicted values for the FOS, also matching those of the actual values (FOS), as illustrated in Figure 5. The Random Forest achieved the same accuracy as the Decision Tree on this dataset since it was based on the prediction of multiple decision trees, therefore allowing for a combination, "ensemble," of the multiple decision tree predictions for increased accuracy for this dataset.

Table 3: Comparative Performance of Machine Learning Models

Model	RMSE	MAE	R^2 Score	Performance Summary
Multilinear Regression	0.184	0.132	0.82	Baseline model: identifies linear relationships but has limited accuracy for nonlinear behavior
Decision Tree	0.145	0.110	0.89	Handles nonlinearity well, but is prone to overfitting and unstable with noise
Random Forest	0.098	0.072	0.94	Best overall performance; robust, stable predictions, reduces variance through an ensemble approach

The method comparison for predicting landslides from live IoT sensors shows that Multilinear Regression produces a predictability result of 0.82 (with a baseline performance). The Multilinear regression model had the highest root mean square error (RMSE) and mean absolute error (MAE); therefore, it did not capture the nonlinear behaviour of the slope predictions based on IoT sensors. The Decision Tree method produced an improved predictability value of 0.89, but also had lower RMSE and MAE values, and because of this, it could model a more accurate non-linear solution. Although due to the structure of how it works, it does have potential for overfitting as well as underfitting, depending on the dataset being used. Random Forest was able to produce the best results (with an R^2 value of 0.94) and has produced the lowest RMSE and MAE values. Random Forest provides both higher reliability and stability for the early detection of landslides as part of IoT-based systems.

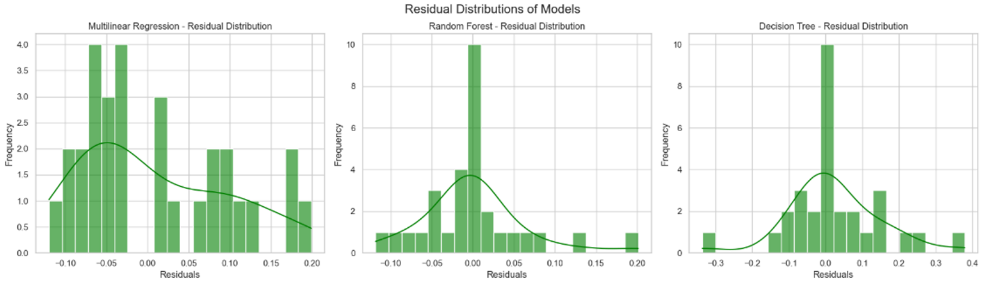

Fig. 6 Comparative Residual Distributions Across All Machine Learning Models

Figure 6 shows the analysis of the residuals of Multilinear Regression (MLR), Random Forest (RF), and Decision Tree (DT) models that are employed in the system of IoT-Based Landslide Monitoring. The values of the MLR residuals are diversified with values of negative and positive values ranging between -0.10 and +0.20, which means that it is less accurate in capturing nonlinear relationships and is likely to deviate from the real trend line, hence moderate prediction errors. On the contrary, the RF model has incredibly low median residual and a near symmetrical distribution around zero. This shows that RF has a high predictive accuracy, low variance, and good consistency. Its ensemble learning algorithm is successful in minimizing bias and enhancing the reliability of models. The DT model was able to fit nonlinear trends of the data; the residual values exhibited moderate skewness and were not distributed as a perfect normal distribution. Although most DT residuals are concentrated around zero, it is also evident that some of the residual values continue to be noticeably skewed towards both positive and negative values, indicating that the model is overfitting certain patterns in the data occasionally, as well as being sensitive to particular data distributions. In general, despite the relatively good performance of DT, RF is obviously providing the strongest and most accurate predictions of the three models. Thus, RF is the best model for real-time landslides monitoring and estimating the Factor of Safety (FOS) of IoT-based geotechnical applications.

6 Conclusion

This research paper highlights an IoT-based solution that monitors and predicts landslides by means of intelligent sensing, real-time data collection, and machine-learning analytics. Important geotechnical parameters that are captured by the system include soil cohesion, internal friction, slope angle, and slope height, as well as meteorological conditions, including rainfall, which affect the slope behaviour and Factor of Safety (FOS). Through the precise collection and processing of these inputs, the system gives a detailed feeling of slope stability. A comparative study of the Multilinear Regression, Decision Trees, and the Random Forest models reveals that data-driven approaches are highly accurate and reliable in the prediction of landslides. Machine learning allows the system to keep learning on the incoming data of the IoT sensors, and refine the relationships between the features and identify early warning signals, say increasing rainfall or increasing soil saturation, and minimize false alarms. Due to the changing nature of environmental and geotechnical conditions with time, continuous monitoring should be carried out to determine the real-time ground stability and address the risks. The results point out that Multilinear Regression is a good tool that is able to describe the dependencies and interactions among geotechnical and environmental variables. All in all, IoT implementation has a great influence on real-time data collection as well as the ability to carry out dynamic risk assessment based on clouds to improve the process of landslide early warning and management.

References

1. Onyelowe, K. C., Ebid, A. M., Hanandeh, S., & Kamchoom, V. Evaluating the slope behavior for geophysical flow prediction with advanced machine learning combinations. Scientific Reports, **15**(1), (2025). https://doi.org/10.1038/s41598-025-90882-8
2. Indukala, P. K., Gosh, U. G., & Vinodini Ramesh, M. IoT-Driven Microseismic Sensing System and Monitoring Platform for Landslide Detection. IEEE Access, **12**, 97787–97805, (2024). https://doi.org/10.1109/ACCESS.2024.3427386

3. Das, S. K., Kumar Pani, S., Samal, A. K., Padhy, S., Dash, S., & Jayanthu, S. (n.d.). Improving Slope Stability in Open Cast Mines via Machine Learning based IoT Framework. In IJACSA) International Journal of Advanced Computer Science and Applications **13**, 10, 970-977 (2022).
4. Lin, Y. S., Chen, I. H., Ho, S. C., Chen, J. Y., & Su, M. bin. Applying time domain reflectometry to quantification of slope deformation by shear failure in a landslide. Environmental Earth Sciences, **78**(5), (2019).
   https://doi.org/10.1007/s12665-019-8124-4
5. Park, S., Lim, H., Tamang, B., Jin, J., Lee, S., Chang, S., & Kim, Y. A Study on the Slope Failure Monitoring of a Model Slope by the Application of a Displacement Sensor. Journal of Sensors, **2019**, 7570517, (2019). https://doi.org/10.1155/2019/7570517
6. Babu Thoppil, B., Premachandran, A., Annapoorna M, Mathew Zachariah, A., & Susan Jacob, B. (n.d.). Advanced Sensor-Based Landslide and Earthquake Detection and Alert System Utilizing Machine Learning and Computer Vision Technologies. International Journal on Emerging Research Areas, **4**, 2, (2024). https://ijera.in/index.php/IJERA/article/view/56
7. Moayedi, H., Bui, D. T., & Foong, L. K. Slope stability monitoring using novel remote sensing based fuzzy logic. Sensors , 19(21), (2019). https://doi.org/10.3390/s19214636
8. Dong, M., Wu, H., Hu, H., Azzam, R., Zhang, L., Zheng, Z., & Gong, X. Deformation prediction of unstable slopes based on real-time monitoring and deepar model. Sensors, 21(1), 1–18, (2021). https://doi.org/10.3390/s21010014
9. Sharma, A., Mohana, R., Kukkar, A., Chodha, V., & Bansal, P. An ensemble learning-based experimental framework for smart landslide detection, monitoring, prediction, and warning in IoT-cloud environment. Environmental Science and Pollution Research, 30(58), 122677–122699, (2023).
   https://doi.org/10.1007/s11356-023-30683-6
10. Pathania, A., Priyanka, P., Kumar, P., Singh Bora, R., Singh, R., Chaturvedi, P., Uday, K. v, & Dutt, V. (n.d.). A Low Cost, Sub-Surface IoT Framework for Landslide Monitoring, Warning, and Prediction, In book: Proceedings of 2020 International conference on advances in computing, communication, embedded and secure systems, (2020).
    https://www.researchgate.net/publication/339850328