EARLY WARNING SYSTEM FOR LANDSLIDES USING WIRELESS SENSOR NETWORKS

DKP Perera[1], MKK Munasinghe[2], WMCM Rajakaruna[3], BGE Priyanwada[4] and PN Karunanayake[5]

[1,2,3,4,5]Department of Electrical, Electronic and Telecommunication Engineering, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka

[2]kusalakalani@gmail.com

Abstract - Landslides occur in many areas in Sri Lanka, and they cause considerable damage to natural habitat, environment, economy and other resources. Monitoring, predicting and controlling are the three major challenges associated with landslides due to the randomness of the event. Yet, developing an accurate prediction mechanism with an effective early warning system has become a need of the hour since the damages and the losses caused by the landslides are intolerable. Although there are expensive and advanced mechanisms deployed in foreign countries to predict the possibility of occurring landslides, such as satellites and radar systems with artificial intelligence capabilities, Sri Lanka finds it difficult to afford them due to the high cost and the advanced technologies used. When compared with the existing high-end systems, a simple wireless sensor network which is capable of identifying the underground movements and soil conditions is a cost effective, practical solution. But, dealing with a large number of variables manually with no proper understanding about their contribution for the occurrence of a landslide is difficult. Machine learning, which is a method used to create complex models and algorithms that lend themselves to predict is a fruitful solution for that issue. This research work is carried out to develop a cost-effective early warning system for land slides using WSNs incorporating machine learning.

Keywords - Wireless sensor network, Machine Learning, Landslide Prediction, Early Warning

I. INTRODUCTION

Landslide can be considered as a main problem which occurs in many areas in Sri Lanka. It causes a considerable damage to the natural habitat, environment, economy and other resources. Landslide monitoring, prediction and managing are the three major challenges associated with landslides due to the randomness of the event. Yet, developing an accurate prediction mechanism with an effective early warning system has become a need of the hour since the damages and the losses caused by landslides are intolerable. So far, many static and dynamic models and prediction mechanisms based on different approaches have been locally tested, validated and improved in various parts of the world. However, a numerous number of problems remain still unsolved hindering accurate prediction of landslide hazards especially when real time forecasting is in concern. In fact, the soil conditions in different parts of the world are much deviated from one another, making it more difficult to adopt one model or a prediction mechanism all over the world. Therefore, more and more studies with the introduction of various factor combinations into modern analytical methodologies are still necessary to come out with more appropriate models.

In the Sri Lankan context, landslide susceptibility maps have been prepared; yet they have no any temporal implications or information about the intensity of triggering events. Hence, their role in managing a landslide event is very much limited. Due to lack of resources, expert knowledge and research interest, the amount of attempts that have been taken so far for the improvement of models and for the development of real time forecasting methodologies in the country are inadequate.

Identifying the necessity for a locally developed accurate landslide prediction mechanism, the challenge to develop an accurate landslide prediction and early warning system has been taken up by this research. It ultimately delivers an intelligent landslide prediction model which incorporates machine learning capabilities to predict the possibility for a landslide by analysing the actual data obtained from a particular landslide prone area, and an effective early warning system that helps to mitigate loss of lives in future landslides. Further, a Wireless Sensor Network is prototyped and the prediction model plus early warning system is tested with the prototype to illustrate that the suggested system can be practically deployed and functioned in the real world.

II. METHODOLOGY

A. Data Filtering

A large data set which contains sensor readings of landslide prone areas in Sri Lanka for nearly three years was obtained from NBRO (National Building Research Organization), the collected data were filtered and re-arranged into a suitable format.

1) Sensor readings

The data set included sensor readings of five sensor types which had been deployed in Kahagalla, Sri Lanka. The number of sensors deployed, their frequency of data collection and the total duration are mentioned in the table below;

Table 1. Summary of the filtered data

Sensor Type	Number of Sensors Deployed	Frequency of data collection	Total Duration
Rain Gauge	1	1 Day	3 Years
Extensometer	4	1 Hour	3 Years
Strain Gauge	2	1 Day	3 Years
Water Level Meter	1	1 Hour	3 Years
Inclinometer	3	1 Month	3 Years

Source: National Building & Research Organization

2) Issued warnings by NBRO

The data set was matched with the dates where a landslide warning was issued to Kahagalla area. During 2015 to 2016, there had been five warning situations where the people who lived in that area were asked to evacuate.

B. Machine learning function

1) Identifying the machine learning approach

Machine learning is used due to the fact that; it can handle a vast amount of data and parameters. Also it does not require a predefined model. To predict whether a landslide will occur or not, the system needs to provide two outcomes. Because of that; binary classification method can be used to build the machine learning function.

If Y is the output, then;

Y =1; Warning Situation

Y =0; Normal Situation

Since adequate amount of previous data are available for the study, and the data set contains several sensor types, the output datasets can be provided for training the system. Hence, the machine learning approach can adopt supervised learning method. To analyse the relationship between the predictors of this study, logistic regression is selected.

2) Feature Selection

Feature selection is one of the paramount tasks in machine learning. Among the available three main methods to choose the most appropriate feature set; K-Fold Cross Validation method was selected which partitions the original sample into several testing sets. The general recommendation for predictive models is K=10. Hence, it was used in this study as well. Feature selection was conducted under two scenarios;

Scenario 01: Each sensor was considered individually and its contribution for the prediction was measured. Based on the obtained results, different feature combinations were selected to train them.

Scenario 02: "Recursive feature elimination method" was also used to choose feature sets. Instead of taking feature sets from each sensor as one whole group, each feature was taken as an individual feature and measured the success rate of each and every one of them and a feature set with highest success rate was selected.

3) Model Evaluation - Confusion Matrix

In order to evaluate the created models, a "Confusion Matrix", which is the technique used for summarizing the performance of a classification algorithm, was used. The confusion matrix of binary classification is a 2×2 table formed by four outcomes;

True positive (TP) : correct positive prediction

False positive (FP) : incorrect positive prediction

True negative (TN) : correct negative prediction

False negative (FN) : incorrect negative prediction

Accuracy of the models were evaluated using the following measures.

Classification Accuracy

Calculated as the number of all correct predictions divided by the total number of the dataset. The best possible accuracy is 1.0, whereas the worst is 0.0.

Classification Error

Error rate (ERR) was calculated as the number of all incorrect predictions divided by the total number of the dataset. The best possible error rate is 0.0, whereas the worst is 1.0.

Sensitivity

Sensitivity was calculated as the number of correct positive predictions divided by the total number of positives. It is also called recall (REC) or true positive rate (TPR). The best possible sensitivity is 1.0, whereas the worst is 0.0.

Specificity (True negative rate)

Specificity was calculated as the number of correct negative predictions divided by the total number of negatives. The best possible is 1.0 and the worst is 0.0.

False Positive Rate (FPR)

FPR was calculated as the number of incorrect positive predictions divided by the total number of negatives. The best possible FPR is 0.0 whereas the worst is 1.0.

Precision

Precision was calculated as the number of correct positive predictions divided by the total number of positive predictions. It is also called positive predictive value (PPV). The best precision possible is 1.0.

C. Prototype Sensor node construction and Web interface development

A prototype of a Wireless Sensor Network, which consists different sensor types to measure the movement of underground soli layers, pressure and the water level, was created to illustrate that the proposed landslide prediction model and the early warning system can be practically deployed and functioned in an actual landslide prone area.

The sensor types, number of sensors, number of slave nodes and their positioning varies from one site to another depending on the condition of that specific site. Hence, designing of sensor nodes for a particular site and deploying them is recommended to be conducted under the supervision geologists. Yet, the basic structure of the sensor network can be created as follows;

1) Sensor Selection

The sensor types which gave high contribution for the prediction model were selected based on the cost and their availability as prototype sensors.

2) Sensor node construction

Two Sensor nodes and a central node was constructed for the prototype sensor network. Plastic boxes were used to place the microcontrollers, transceiver modules and power supplies. The sensors were fixed on pipes in appropriate ways to get sensor readings and the all nodes were powered by the power banks. The modules and the sensors used for each node are as follows;

Central node / Master node

Node MCU was used as the microcontroller and the Wi-Fi module since it has ESP8266 Wi-Fi module embedded with it. NRF24l01 transceiver module was also used to receive sensor data.

Sensor nodes / Slave nodes

One node was made with 4 Strain Gauges and a transceiver module. The other node contained a Rain sensor, Accelerometer, thin film pressure sensor and a transceiver module.

3) Communication link establishment

The link between master node and slave nodes was a point-to-multipoint link. Slave nodes transmitted data using 2.4 GHz NRF24L01 transceivers. Data packets were sent in character array and a specific character was added at the beginning of the array for identification.

After connecting to an access point, master node transmitted received sensor data form slave node to the web server using ESP8266 Wi-Fi module. Transmissions were delayed by necessary time intervals to minimize the data loss. Slave nodes transmitted data for every 5 seconds and Master node transmitted data for every 20 seconds.

4) Web interface and databases

The web interface was created using a public web server and Bootstrap Framework was used to create the interfaces.HTML 5, CSS, JavaScript and php was used for coding. The web interface contains a logging page and a home page. The home page shows all sensor data in corresponding tables created for each and every sensor type used in the prototype. Page is refreshed for every 5 seconds to update data the tables.

A database with 4 tables was created using MySQL to store sensor data. All tables have columns to show date and time. Last 5 values of the tables are shown on the web page. Chart.js which is a JavaScript library, is used to plot graphs. Last 50 values of the tables in the

Classification Accuracy	-	98.1%
Classification Error	-	0.1%
Sensitivity	-	97.6%
Specificity	-	98.8%
False Positive Rate	-	0.1%
Precision	-	93.2%

MySQL database were used to plot the graphs and it is updated for every 20 seconds. The X axis of the graph indicates the time and the Y axis indicates the corresponding sensor values.

Classification Accuracy	-	99.1%
Classification Error	-	0.1%
Sensitivity	-	100%
Specificity	-	98.8%
False Positive Rate	-	0.1%
Precision	-	99.2%

D. Warning Message Generation

Machine learning function was coded in a php script. Real-time sensor data were given as inputs to the machine learning function. After evaluating the function, if the value is greater than 0.5, a warning message was shown on the web page stating a landslide threat with audio

Classification Accuracy	-	99.8%
Classification Error	-	0.001%
Sensitivity	-	99.4%
Specificity	-	99.8%
False Positive Rate	-	0.007%
Precision	-	99.4%

output. To verify the output, the previously mentioned accuracy measures; that is, Classification Accuracy, Classification Error, Sensitivity, Specificity, False Positive Rate, Precision etcetera were used.

III. RESULTS

A. Feature selection

1) Scenario 01 – Feature Selection using individual sensors

11,876 samples and 126 features were used in feature selection process. Based on individual sensors, eight basic

Feature Set	Sensors	10 Fold Cross Validation Accuracy
1	All the sensors	93.3%
2	Rain Gauge	87.0%
3	Extensometers	98.3%
4	Strain Gauge-1	96.1%
5	Strain Gauge-2	91.2%
6	Inclinometer-1	74.9%
7	Inclinometer-2	77.3%
8	Inclinometer-3	77.3%
9	Ground Water	80.5%
10	Rain Gauge+Extensometers+Strain Gauge	92.7%
.....	.....	.....

Figure 1. Results of the Feature Selection
Source: Experimental data

feature collections were selected. Figure 1 indicates how basic sensor types, selected combinations of the sensors and all the sensors as one group contributes to a successful feature selection process with 10-fold cross validation method.

Extensometer and Strain gauge collections shows higher success rate than other basic features. Surprisingly rain sensor success rate is below than extensometer and strain gauges. Among the several types of combinations tested, Extensometer and Strain gauge combination showed more promising results. Therefore Extensometers and Strain Gauges +Extensometers collections were taken as two feature sets to build two models.

2) Scenario 02 - Recursive Feature Elimination Process

Table 2. Comparison between three models

Measurement	Model 1	Model 2	Model 3
Classification Accuracy	98.1%	99.1%	99.8%
Classification Error	0.1%	0.1%	0.001%
Sensitivity	97.6%	100%	99.4%
Specificity	98.8%	98.8%	99.8%
False Positive Rate	0.1%	0.1%	0.007%
Precision	93.2%	99.2%	99.4%

Source: Experimental data

B. Table updating and plotting graphs

Strain Guages

Date	Time	Strain_guage1	Strain_guage2	Strain_guage3	Strain_guage4
2017-12-05	10:34:[illegible]	92	96	186	20
2017-12-05	10:34:[illegible]	92	96	186	20
2017-12-05	10:34:[illegible]	92	96	186	20
2017-12-05	10:34:[illegible]	92	113	186	20
2017-12-05	10:34:[illegible]	92	96	186	20

Figure 5. Table updating in Web interface
Source: Experimental data

The X axis of the graph indicates the time and the Y axis indicates the corresponding sensor values.

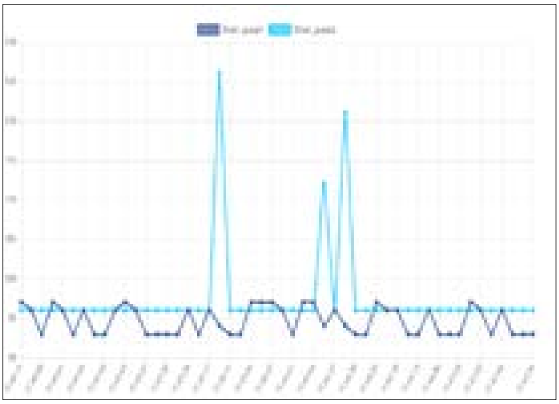

Figure 6. Graph created for two strain gauges
Source: Experimental data

C. Generating warning messages

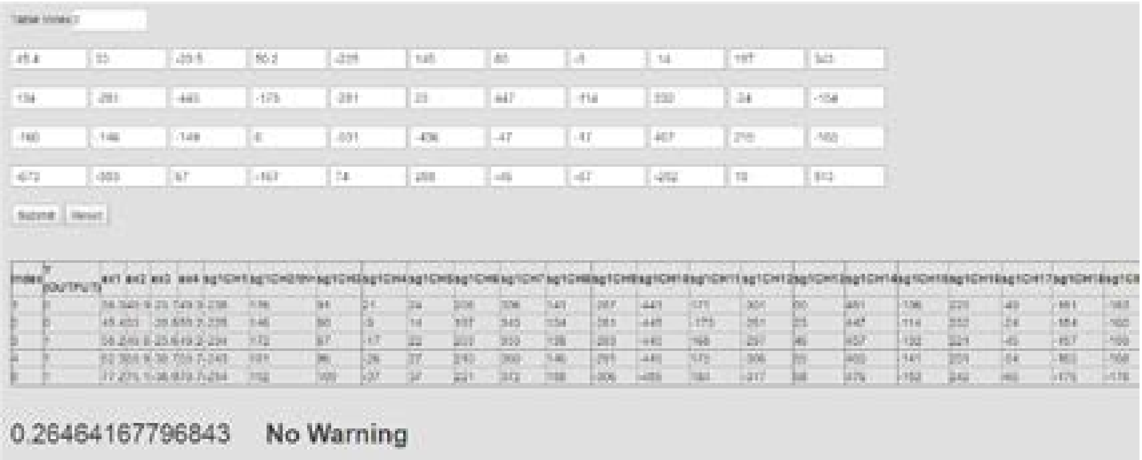

Figure 7. "No Risk" situation
Source: Experimental data

When the system was tested with manually entered data values corresponding to a critical situation the system generates an alert indicating that the system has identified a possibility for a landslide.

IV. CONCLUSION

When considering the past records of natural disasters occurred in Sri Lanka, landslide is a commonphenomenon which has caused intolerable damages to the living beings and properties. Due to the randomness, high uncertainty and the complexity, it is quite difficult to predict the possibility for a landslide and hence, issuing early warnings has become a challenging task. Up to now, the Disaster Management Centre of Sri Lanka adopts a simple method when issuing early warnings for landslides, which is totally based on the amount of rainfall to a particular landslide area. Even though rainfall is considered as one of the two triggering factors of a landslides, relying completely on the rainfall data when issuing landslides is not an acceptable method due to there are more accurate measurable parameters that can be used to predict the possibility for a landslide.

Compared with the existing high-end systems, a simple wireless sensor network which is capable of identifying the underground movements and soil conditions is a cost effective, practical solution. But, dealing with a large number of variables manually with no proper understanding about their contribution for occurrence of a landslide is difficult for human brains. Machine learning, which is a method used to create complex models and algorithms that lend themselves to predict; is a fruitful solution for that issue. Incorporating the cost effective Wireless Sensor Networks and machine learning approaches, developing an accurate landslide prediction and early warning mechanism to minimize anticipated damage of future landslides in Sri Lanka was the ultimate goal of this project.

Achieving the expected outcome and the main objectives, an intelligent landslide prediction model which incorporates machine learning capabilities to predict the possibility for a landslide was developed and tested with the prototype to illustrate the practicability of the suggested system. Since the developed model incorporates the parameters which contribute more than rainfall for the occurrence of a landslide, the accuracy and the reliability of the proposed system is higher and the warnings generated by the system are trustworthy and true when compared to the existing warning mechanism. Depending on all the results of the system, it can be concluded that by deploying a Wireless Sensor Networks in the landslide prone areas with the required sensor types and by using a fine-tuned version (to match with the soil condition of the selected site) of the proposed model to predict the possibility of landslides, the anticipated damage of future landslides in Sri Lanka can be reduced up to a great level.

V. ACKNOWLEDGEMENT

The authors are greatly thankful to Mrs. PN Karunanayeke of General Sir John Kotelawala Defence University and Dr. RT Uthyashanker for all the generous encouragement and support that helped and motivated the authors to successfully complete this research. Further, the authors offer their sincere gratitude to Mr. Mr. R.M.S. Bandara of National Building and Research Organization and Dr. Gamini Jayathissa for the valuable advices and support given when making this concept a reality.

VI. REFERENCES

Cui X, Zhao X, et al (2010) Research on landslide prediction based on support vector model. In: 2010 *International Conference onComputer Design and Applications (ICCDA).* 540–545.

Fosalau C, Zet C, et al (2016) Implementation of a landslide monitoring system as a wireless sensor network, In: *Electronics & Mobile Communication Conference (UEMCON).* 66-72.

Jayathissa M, (2009) Socio-Technical approaches in landslide risk reduction, In: *Second International Summer School and Conference for Applied Geo-informatics for Society and Environment (AGSE 2009),* Stuttgart University of Applied Sciences. 12-18.

Kunnath A.T and Ramesh M.V, (2010) Integrating Geophone Network to Real-Time Wireless Sensor Network System for Landslide Detection, In: *First International Conference onSensor Device Technologies and Applications.* 167–172

Lian C, Zeng Z, et al (2014) Performance of combined artificial neural networks for forecasting landslide displacement, In: *2014 International Joint Conference on Neural Networks (IJCNN).* 418–424.

Mahrooghy M, Aanstoos A, et al (2015) A machine learning framework for detecting landslides on Earthen Levees using Spaceborne SAR imagery,In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 8, no. 8, 2015, 3791–3801

Mahrooghy M, et al (2011), Levee soil moisture assessment based on backpropagation neural network using synthetic aperture radar, In: *34th International Symposium -Remote Sensors. Environ,* Sydney, Australia. 123-128.

Pachauri G and Sharma S, (2015). Anomaly Detection in Medical Wireless Sensor Networks using Machine Learning Algorithms. In: 4th International conference on *Eco-Friendly Computing and Communication Systems.* Elsevier, 325-333.

Ramesh MV, (2009) Real-time Wireless Sensor Network for Landslide Detection, In: *The Third International Conference on Sensor Technologies and Applications,* SENSORCOMM 2009, IEEE, Greece, June 18-23, 2009. 405-409.

Ramesh MV, Kumar S, et al (2009), Wireless Sensor Network for Landslide Detection. In: *Third International Conference onSensor Technologies and Applications, 2009. SENSORCOMM '09.* 18 June 2009. IEEE. 78–85.

Warriach E and Tei K, (2013). Fault detection in Wireless Sensor Networks: A Machine Learning Approach, In: *2013 IEEE 16th International Conference on Computational Science and Engineering.* 758-765.