Slope Movement Prediction and Early Warnings via Machine Learning Algorithms and IoT Technologies

Praveen Kumar
(S17007)

Supervisor: Dr. Varun Dutt

School of Computing and Electrical Engineering
Indian Institute of Technology Mandi

This dissertation is submitted for the degree of
Master of Science (by Research)

October 2020

THESIS CERTIFICATE

This is to certified that the work contained in thesis entitled **"Slope Movement Prediction and Early Warnings via Machine Learning Algorithms and IoT Technologies,"** being submitted by **Mr. Praveen Kumar** (Enroll. No. S17007) has been carried out under my supervision. In my opinion, the thesis has reached the standard fulfilling the requirement of regulation of the M.S. degree. The results embodied in this thesis have not been submitted elsewhere for the award of any degree or diploma.

Dr. Varun Dutt
October 2020

Declaration by the Research Scholar

I hereby declare that the entire work embodied in this thesis is the result of investigations carried out by me in the ***School of Computing and Electrical Engineering***, Indian Institute of Technology Mandi, under the supervision of ***Dr. Varun Dutt***, and that it has not been submitted elsewhere for any degree or diploma. In keeping with the general practice, due acknowledgements have been made where the work described is based on finding of others investigators.

Praveen Kumar
(S17007)
October 2020

Declaration by the Research Advisor

I hereby certify that ***Praveen Kumar*** has carried out the entire work in this thesis under my supervision in the ***School of Computing and Electrical Engineering***, Indian Institute of Technology Mandi, and that no part of it has been submitted elsewhere for any degree or diploma.

Dr. Varun Dutt
October 2020

Acknowledgements

I want to thank my thesis advisor, Dr. Varun Dutt of the School of Computing and Electrical Engineering, at the Indian Institute of Technology Mandi. Prof. Dutt's office door was always open for me whenever I ran into a trouble or had a doubt with my research. His consistent guidance and push towards excellence in this research steered me in the right direction.

I also want to thank the experts who were involved in the validation survey for this research project: [Dr. Kala Venkata Uday of the School of Engineering at Indian Institute of Technology Mandi, and Mr. Pratik Chaturvedi from DTRL-DRDO]. The validation survey would not have been conducted successfully without their highly valued input and passionate participation.

Grants supported the project from National Disaster Management Authority (IITM/NDMA/VD/184), Defence Research and Development Organization (IITM/DRDO-DTRL/VD/179), and Deputy Commissioner of Mandi, India (IITM/DCoM/VD/204) to Varun Dutt. We are also grateful to the Indian Institute of Technology Mandi for providing computational resources for this project.

I would also like to take this opportunity to acknowledge Dr. Shyamashree Dasgupta of the School of Humanities and Social Science and Dr. Arnav Bhavsar of the School of Computing and Electrical Engineering at the Indian Institute of Technology Mandi for being the second readers of this thesis. I am gratefully indebted to them for their valuable comment on this thesis.

My appreciations also extend to my laboratory colleague Mr. Ankush Pathania, whose mentoring and encouragement have been extremely valuable.

Finally, I must express my profound gratitude to my parents and my spouse, Priyanka, to provide me with unfailing support and continuous encouragement throughout my academics and research and also while working on this thesis. This accomplishment would not have been possible without any of them.

Thank you all for being an integral part of my successful journey.

Abstract

Landslides are widespread disasters in hilly regions. These disasters cause lots of injuries and deaths every year. Due to these injuries and deaths, it is imperative to monitor landslides and to warn people about impending disasters timely. It is also essential to predict slope movements ahead of time so that people get enough lead time to evacuate from the sliding region. The existing technologies monitor landslides at a very high cost, and these technologies do not warn people and predict slope movements ahead of time. Thus, one objective of this thesis is to detail the development, deployment, and calibration of a new low-cost IoT-based landslide monitoring, warning, and prediction system. The system is deployed on the soil surface, and it can generate real-time warnings via SMSes, blinker, and hooter in case significant surface movements occur. The main advantage of the new system is that it is low-cost, and it may be deployable at many locations to monitor the landslides and to warn people timely.

The second objective of this thesis is to perform predictive analytics on the time-series data of slope movements. For this purpose, time-series data were collected over 78 weeks from July 2012 to July 2014 using inclinometers that were placed in five boreholes at the Tangni landslide in Chamoli, India. These sensors measured tilt in degree units (essentially the angle the inclinometer tilted over time). Different algorithm parameters were calibrated to the training data (first 62-weeks) and then made to predict the test data (the last 16-weeks) across the five time-series (i.e., one series from each sensor in a borehole).

In the first experiment, moving-average algorithms (SARIMA and AR) and support vector regression algorithms (SMOreg) were developed. Each algorithm was calibrated to each time-series independently. In training, the AR and SMOreg algorithms performed the best and second-best with RMSEs of 0.40° and 0.37°, respectively, compared to the SARIMA algorithm (RMSE: 0.71°). However, when these algorithms were applied on the test dataset, results revealed that SARIMA performed best (RMSE: 0.33°) compared to the SMOreg and AR algorithms (SMOreg, RMSE: 0.54°; AR, RMSE: 0.59°). In this experiment, it was found that the moving-average SARIMA algorithm outperformed the support-vector regression algorithm for slope movement predictions. Also, the test accuracy was higher than the training accuracy. Although we may only speculate, one explanation could be that

the parameters and mechanisms in these algorithms may have enabled them to generalize to the unseen test dataset.

In the second experiment, neural-network algorithms (MLPs and LSTMs) were developed and compared with moving-average algorithms (SARIMA), where the latter moving-average algorithm had performed best in the first experiment. One-step-ahead walk-forward validation was used for algorithm comparisons. In the training of algorithms, LSTM and SARIMA algorithms performed the best and second-best with RMSEs of 0.37° and 0.71°, respectively, compared to the MLP algorithm (RMSE: 0.99°). When these algorithms were evaluated on the test dataset, it was found that the SARIMA algorithm (RMSE: 0.33°) performed better compared to the LSTM (RMSE: 0.37°) and MLP (RMSE: 0.38°) algorithms.

In the third experiment, ensemble and non-ensemble machine-learning (ML) algorithms were compared to predict slope movements. Non-ensemble algorithms (Sequential Minimal Optimization Regression (SMOreg), and Autoregression) and ensemble algorithms (Random Forest, Bagging, Stacking, and Voting) involving the non-ensemble algorithms were used. Results revealed that the ensemble algorithms (Bagging, Stacking, and Random Forest) performed better compared to non-ensemble algorithms. These results also showed that the ensemble algorithms seem to follow the general pattern where the training error was lesser compared to the test error.

In the fourth experiment, moving-average algorithms (Seasonal Autoregressive Integrated Moving Average (SARIMA) algorithm and Autoregressive (AR) algorithm), Lazy algorithms (Instance-based-k (IBk) and Locally Weighted Learning (LWL)) and information-gain algorithms (REPTree and M5P) were compared in their ability to predict slope movements. Results revealed that the moving-average algorithms (SARIMA and AR) performed better compared to the lazy and information-gain algorithms during both training and test. Specifically, the SARIMA algorithm possessed the smallest error compared to other algorithms in test data.

From all the experiments, one could conclude that the moving-average algorithms perform better compared to other algorithms. A likely reason for these findings is the presence of seasonal, auto-regressive, and moving-average components in the moving-average algorithms. The implications of our results on slope movement predictions in the real-world are highlighted.

Contents