engineering proceedings

Proceeding Paper

# Machine Learning-Based Slope Failure Prediction Model Considering the Uncertainty of Prediction †

**Junhyuk Choi [1], Yongkyu Cho [2], Yongjin Kim [3], Yongseong Kim [4] and Bongjun Ji [4,*]**

[1] Department of Industrial and Management Engineering, Pohang University of Science and Technology, Pohang 37673, Republic of Korea; cjh0102@postech.ac.kr
[2] Department of Industrial Engineering, Kangnam University, Yongin 16979, Republic of Korea; yongkyu.cho@kangnam.ac.kr
[3] Smartgeotech, Chuncheon 24341, Republic of Korea
[4] Department of Regional Infrastructure Engineering, Kangwon National University, Chuncheon 24341, Republic of Korea
* Correspondence: bjji@kangwon.ac.kr

† Presented at the Second International Conference on Maintenance and Rehabilitation of Constructed Infrastructure Facilities, Honolulu, HI, USA, 16–19 August 2023.

**Abstract:** Slope failure is a severe natural disaster that can cause property damage and human costs. In order to develop a warning system for slope failure, various studies have been conducted, including research based on both physics-based models and machine learning-based models. While machine learning-based approaches have shown promise due to their ability to automatically extract hidden patterns in data, conventional machine learning models have their limitations. Specifically, while they can always provide a prediction value, they fail to provide information about the uncertainty of the prediction results. In this study, we developed a machine learning model that can predict the slope failure by training trends in time-series data. Our proposed model addresses the limitations of the conventional machine learning models by incorporating the Monte Carlo dropout to calculate the uncertainty during the prediction stage. The experimental results demonstrated that the proposed model significantly outperforms the conventional machine learning models in terms of both its prediction accuracy and the ability to estimate uncertainty. Furthermore, the model proposed in this study can support decision-makers by providing more accurate information than the conventional models.

**Keywords:** slope failure; machine learning; uncertainty



**Citation:** Choi, J.; Cho, Y.; Kim, Y.; Kim, Y.; Ji, B. Machine Learning-Based Slope Failure Prediction Model Considering the Uncertainty of Prediction. *Eng. Proc.* **2023**, *36*, 6. https://doi.org/10.3390/engproc2023036006

Academic Editor: Hosin (David) Lee

Published: 29 June 2023



## 1. Introduction

Slope failure is a severe natural disaster that can cause significant property damage and even the loss of life [1]. In order to develop a warning system for slope failure, various studies have been conducted in the past. One approach that has gained significant attention in recent years is machine learning [2]. Machine learning-based models have the potential to provide accurate predictions by automatically extracting the hidden patterns in data [3].

In contrast to machine learning, physics-based models are built on the principles of mechanics and physics [4]. These models typically require a detailed understanding of the geology and soil mechanics of the area in question. The major disadvantage of physics-based models is that obtaining an exact mathematical formula is often oversimplified and impractical, especially in complex geological settings [5]. These limitations lead to the formation of models that are not sufficiently precise or efficient for predicting the slope failure [6].

On the other hand, machine learning-based prediction models can be highly effective at predicting the slope failure by training on large datasets and identifying the hidden patterns in the data. However, traditional machine learning models often have their limitations. For example, they may need to consider the uncertainty associated with the prediction result, which can make it difficult for decision-makers to make informed decisions.

To address these limitations, we proposed a machine learning-based slope failure prediction model that considers the prediction result's uncertainty. Our model uses time-series data to train a machine learning algorithm and incorporates the Monte Carlo dropout to calculate the uncertainty of the prediction result [7]. In doing so, we aimed to provide decision-makers with more accurate and reliable information about the potential risks associated with a given slope.

The rest of this paper is organized as follows: Section 2 provides an overview of the materials and methods for slope failure prediction. Section 3 describes the results and discussion of the experiment. Finally, Section 4 concludes the paper and suggests some future works.

## 2. Materials and Methods

In this section, we describe the data collection and preprocessing, model architecture, training and validation, and model evaluation procedures that were used to develop a machine learning-based slope failure prediction model considering the uncertainty of prediction.

We used the displacement and remaining failure time data collected from a slope stability monitoring system installed on a slope located in Yeongdeok-gun, Gyeongsangbuk-do [8]. The monitoring system recorded the displacement of the slope at a frequency of day and the remaining failure time of the slope at regular intervals. The dataset consisted of 586 time-series data.

To prepare the displacement and remaining failure time data for use in our proposed model, we performed z-score normalization. This involved calculating the mean and standard deviation values of the data and subtracting the mean from each data point, then dividing the result by the standard deviation. This normalization step was necessary to perform to ensure that the data were on a similar scale, allowing our model to effectively identify patterns and relationships in the time-series data.

We utilized a long short-term memory (LSTM) model, a type of recurrent neural network (RNN) [9], to develop a machine learning-based slope failure prediction model. LSTMs are well-suited for sequence-to-sequence prediction tasks [10], as traditional RNNs often face the vanishing gradient problem where gradients become too small to learn long-term dependencies over many time steps. To address this issue, LSTMs incorporate a memory cell and gating mechanisms that regulate the flow of the information. The memory cell retains information over time, while the gating mechanisms control the amount of information retained, forgotten, or passed to the next time step. These mechanisms comprise three sigmoid functions and one tanh function, which compute the input gate, forget gate, output gate, and candidate memory vector, respectively.

LSTMs are widely used in natural language processing [11], speech recognition [12], and time-series prediction [13]. They are especially effective for tasks requiring long-term memory, such as language modeling, where the network must recall the context of a sentence over many words. In this study, we employed an LSTM network to capture the long-term dependencies present in the time-series data of displacement and remaining failure time.

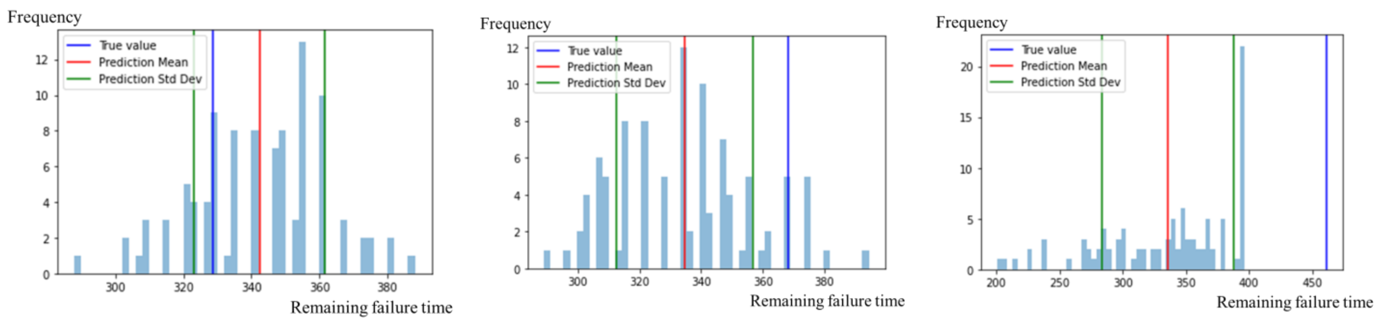
The model architecture is composed of a stack of two LSTM layers, each with 64 units, followed by a dense layer with a single unit. To estimate the uncertainty, we applied the Monte Carlo dropout to the output layer. In further detail, dropout is a regularization technique that is commonly used in deep learning to prevent overfitting. It randomly sets some of the output activations of the layer to zero during training. During inference, dropout is applied by randomly dropping out neurons with a certain probability, and the prediction is then computed by taking the average of the outputs of the network over multiple dropout iterations. This approach allows us to estimate the model's uncertainty by measuring the variance of the predictions across the dropout iterations. Therefore, we applied the MC dropout to the output layer to obtain a more accurate estimate of the model's uncertainty.

We randomly split the collected data into a training data and a validation data set with a ratio of 70% and 30%, respectively. We used the Adam optimizer for training the model with a learning rate of 0.01. We used the mean squared error (MSE) loss function with an additional term that calculates the standard deviation of the MC dropout output as a measure of uncertainty. The model was trained for 100 epochs, and batch sizes of 32 were used. The validation loss was used to monitor the model's performance during training.

We assessed the performance of the prediction model using two popular metrics; the root mean squared error (RMSE) and the mean absolute error (MAE). To assess the model's ability to predict the uncertainty of the prediction, we generated 100 MC dropout samples for each test data point and estimated the uncertainty of the prediction. All computations were performed using the TensorFlow deep learning framework.

## 3. Results and Discussion

The results of our study indicate that the machine learning-based slope failure prediction model using a Bayesian LSTM with MC dropout can accurately predict the remaining failure time of a slope, while also estimating the uncertainty of the prediction. The Bayseian LSTM model showed a performance improvement of 3.5 in terms of MAE and 6.9 in terms of RMSE over the original model, respectively. The estimated uncertainty of the prediction allows decision-makers to have a more accurate and reliable assessment of the risks associated with slope instability. Furthermore, the examples of the predicted outputs presented in Figure 1 highlight the significance of estimating uncertainty. The large gap observed between the true and predicted values when the uncertainty of the output is significant emphasizes the importance of considering the uncertainty estimates in making decisions related to the slope stability.




(a) (b) (c)

**Figure 1.** Examples of predicted output distribution which has actual values of; (**a**) 328, (**b**) 368, and (**c**) 460, respectively.

The use of machine learning for slope failure prediction has the potential to provide a more accurate and reliable warning system for slope failure. Compared to physics-based models, machine learning models can better capture the complex relationships between the various factors that contribute to slope failure, such as soil properties, rainfall, and slope geometry. However, the use of machine learning models can also introduce new sources of uncertainty which need to be considered.

In our study, we used a Bayesian LSTM with MC dropout to estimate the uncertainty of the prediction. By estimating the uncertainty, we can provide decision-makers with a more accurate assessment of the risks associated with the slope instability.

## 4. Conclusions

In conclusion, our study developed a machine learning-based slope failure prediction model that can accurately predict the remaining failure time of a slope, while also estimating the uncertainty of the prediction. This model has the potential to provide decision-makers with a more accurate and reliable warning system for slope failure. Future work should focus on validating the model on data from other monitoring systems and slopes, as well as exploring more efficient methods for estimating uncertainty.

**Author Contributions:** Conceptualization, J.C. and B.J.; methodology, B.J.; software, J.C.; validation, J.C., Y.C. and B.J.; formal analysis, J.C.; investigation, J.C. and Y.K. (Yongjin Kim); resources, Y.K. (Yongseong Kim); data curation, Y.C.; writing—original draft preparation, J.C.; writing—review and editing, Y.K. (Yongseong Kim) and B.J.; visualization, J.C.; supervision, B.J.; project administration, B.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** First, this research was supported by "Ministry of the Interior and Safety" R&D program (RS-2022-00155667). Additionally this study was supported by 2022 Research Grant from Kangwon National University.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data have been synthesized for the purpose of this study and are not publicly available.

**Acknowledgments:** Authors would like to acknowledge the significant support that has made this research possible.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Orense, R.P.; Shimoma, S.; Maeda, K.; Towhata, I. Instrumented model slope failure due to water seepage. *J. Nat. Disaster Sci.* **2004**, *26*, 15–26. [CrossRef]
2. Zhang, J.; Wang, Z.; Hu, J.; Xiao, S.; Shang, W. Bayesian machine learning-based method for prediction of slope failure time. *J. Rock Mech. Geotech. Eng.* **2022**, *14*, 1188–1199. [CrossRef]
3. Vasavi, S. Extracting hidden patterns within road accident data using machine learning techniques. In Proceedings of the Information and Communication Technology, Singapore, 12–13 December 2016.
4. White, J.A.; Singham, D.I. Slope stability assessment using stochastic rainfall simulation. *Procedia Comput. Sci.* **2012**, *9*, 699–706. [CrossRef]
5. Azmoon, B.; Biniyaz, A.; Liu, Z.; Sun, Y. Image-data-driven slope stability analysis for preventing landslides using deep learning. *IEEE Access* **2021**, *9*, 150623–150636. [CrossRef]
6. Ma, W.; Dong, J.; Wei, Z.; Peng, L.; Wu, Q.; Chen, C.; Xie, F. Landslide displacement prediction with gated recurrent unit and spatial-temporal correlation. *Front. Earth Sci.* **2022**, *10*, 1182. [CrossRef]
7. Kendall, A.; Badrinarayanan, V.; Cipolla, R. Bayesian segnet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding. *arXiv* **2015**, arXiv:1511.02680.
8. Yoo, B.S. Study of Failure Analysis Methods Based on Real-Time Monitoring Data for Landslide Warning System. Doctoral Dissertation, Kumoh National Institute of Technology, Gumi-si, Republic of Korea, 2006.
9. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
10. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
11. Yao, L.; Guan, Y. An improved LSTM structure for natural language processing. In Proceedings of the IEEE International Conference of Safety Produce Informatization (IICSPI), Chongqing, China, 10–12 December 2018.
12. Graves, A.; Jaitly, N.; Mohamed, A.R. Hybrid speech recognition with deep bidirectional LSTM. In Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding, Olomouc, Czech Republic, 8–12 December 2013.
13. Li, Y.; Zhu, Z.; Kong, D.; Han, H.; Zhao, Y. EA-LSTM: Evolutionary attention-based LSTM for time series prediction. *Knowl.-Based Syst.* **2019**, *181*, 104785. [CrossRef]