ABSTRACT

Eruptive activity that started in June 2018 caused a landslide on Anak Krakatau volcano in December 2018. Volcanic material was released into the sea by the landslide and caused a tsunami that killed 437 people, making it one of the deadliest volcano-generated tsunamis. The tsunami early warning system deployed in Indonesia was ineffective since it could only detect tsunami based on earthquake data. Therefore, a tsunami early warning system based on sea level data is needed. In this paper, we designed a machine-learning-based tsunami early warning system using the Extreme Gradient Boosting (XGBoost) method to classify tsunami and non-tsunami signals based on sea level data. As a study case, we use sea level data obtained from the Inexpensive Device for Sea Level Measurement in Marina Jambu that have been artificially added with tsunami signals from the 2018 tsunami caused by the landslide of Mount Anak Krakatau to increase the number of tsunami signals. After performing feature engineering and data imbalanced handling techniques such as adding lag and time features and applying class weight, we obtained good results from the XGBoost with a macro average F1-score of 0.76. The XGBoost outperformed other machine-learning methods, such as Support Vector Machine (SVM) and Random Forest.

Keywords: tsunami early warning system, classification, machine learning, xgboost