**Machine Learning for Fully Automated Detection of Volcanic Seismic Signals in Real Seismic Records at Sinabung Volcano, North Sumatra**

**Bagas Anwar Arif Nur**1,2\* **, Mohammad Hasib**2 , **Estu Kriswati**2 

1 Department of Physics, State University of Jakarta, Jakarta.

2 Research Center for Geological Disaster, National Research and Innovation Agency (BRIN), Bandung.

1. **Supplementary data**
	1. **Non-Normalized Test**
		1. *PCA Visualization without normalization*



**Figure S1**. Distribution of dataset using PCA with 2 principal components without normalization. The red star, green triangle, and blue asterisk show the distribution of LF, VT, and Explosion

* 1. **Performance of MLP on Non-normalized data**

|  |
| --- |
| **Table S1.** Confusion Matrix of MLP on Non-normalized data |
|  |  | Predicted Label |
|  |  | LF | Explosion | VT |
| True Label | LF | 13 | 0 | 5 |
| Explosion | 6 | 17 | 0 |
| VT | 4 | 0 | 15 |
| Accuracy |  | 75% |
| Precision |  | 79.04% |

* 1. **Performance of RF on Non-normalized data**

|  |
| --- |
| **Table S2.** Confusion Matrix of RF on Non-normalized data |
|  |  | Predicted Label |
|  |  | LF | Explosion | VT |
| True Label | LF | 16 | 0 | 2 |
| Explosion | 1 | 22 | 0 |
| VT | 3 | 0 | 16 |
| Accuracy |  | 90% |
| Precision |  | 90.48% |

* 1. **Evaluation using only 2 class (VT and LF)**
		1. *Performance of L1 using only 2 class*

|  |  |
| --- | --- |
|  | **Table S3.** Confusion Matrix of L1 using only 2 class |
|  |  | Prediction |
|  | LF | VT |
| True Label | LF | 2 | 14 |
| VT | 0 | 24 |
| Accuracy |  | 65.00% |
| Precision |  | 77.89% |

* + 1. *Performance of L2 using only 2 class*

|  |
| --- |
| **Table S4.** Confusion Matrix of L2 using only 2 class |
|  |  | Prediction |
|  | LF | VT |
| True Label | LF | 16 | 0 |
| VT | 3 | 21 |
| Accuracy |  | 92.50% |
| Precision |  | 93.68% |

* 1. **Comparison between prediction and catalogue**
		1. *Comparison between prediction and catalogue for RF*



**Figure S2**. The comparisson between prediction result using RF and earthquake catalogue and for each day. The top panel for each color shows the earthquake catalogue and the bottom panel shows the total number of earthquake predictions.

* + 1. *Comparison between prediction and catalogue for L1*



**Figure S3**. The comparisson between prediction result using L1 and earthquake catalogue and for each day. The top panel for each color shows the earthquake catalogue and the bottom panel shows the total number of earthquake predictions.

* + 1. *Comparison between prediction and catalogue for L2*

**

**Figure S4**. The comparisson between prediction result using L2 and earthquake catalogue and for each day. The top panel for each color shows the earthquake catalogue and the bottom panel shows the total number of earthquake predictions.