Geophysical Research Abstracts
Vol. 21, EGU2019-4486-2, 2019
EGU General Assembly 2019
© Author(s) 2019. CC Attribution 4.0 license.

Final magnitude and source-distance estimation using Autoencoders and Recurrent Neural Networks

Payman Janbakhsh, Russell Pysklywec, and Hosein Shahnas
University of Toronto, Earth Sciences, Toronto, Canada (russ@es.utoronto.ca)

Application of machine learning algorithms to improve EEWS performance for on-site approach, where a single station is located near a user-target site is the goal of this research.
Final magnitude and source-distance estimation using first 3 seconds of P-phase is presented here and estimation of peak ground motion at the user-site is currently under work.
To estimate final magnitude, 80186 vertical records of P-wave with catalog final magnitude between 3-4.5 from multiple stations were selected and divided into training/validation and test sets. Subsequently 4 levels of noise were added to extend the size of training set and increase the model's capability to handle noisy dataset. A total of 390000 traces were used for training/validation of the model with 50 epochs, 5 patience tolerance, and 256 batch size. The best performing model and its weights was used to predict the final magnitude on 10000 traces in test set, and the resulting accuracy is reported at 98% with 0.3 error tolerance in estimated magnitude.
To estimate the source-station distance, initial dataset consists of 830000 HHZ records of P-wave arrival from multiple stations with source-distance range 0-120 km. For the purpose of training the algorithm one arbitrary station is selected for training and testing (assuming that the station would be an on-site station near the user location).
The selected records from a particular station is divided into 15000 traces for training/validation set, and 1819 traces for the final testing of the model. 4 levels of noise was added to the training set to expand the set to 75000 traces.
The model is composed of a pre-trained half autoencoder network (where data space is mapped from a 600-dimensional space to 65792 dimensions, i.e. expanding the dimensionality of feature space), stacked dense net, and stacked LSTM neural net. From 75000 records, 15% was extracted randomly for validation and the rest were used to train the model for 100 epochs at patch size of 256 and patience level of 30. The accuracy report on the unseen 1819 test set is 81% with 5 km error tolerance.
Next step would be training the model to predict peak ground motion at the station site (where it is assumed to be near the user-target site), and expand the data set to 0-8 range for predicting final magnitude.
The model's capability to predict final magnitude and source-target distance at the above-mentioned accuracy can be increased with the aid of larger data set and site-specific noise filters.