Deep Learning (& classification models) performance in aftershock prediction

*Marco Broccardo[1,2], Arnaud Mignan[2,3]

1. Institute of Risk and Uncertainty, University of Liverpool , 2. Swiss Seismological Service, ETH Zürich, Switzerland, 3. Institute of Risk Analysis, Prediction and Management, Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology, Shenzhen, China

Deep Learning is growing rapidly, as stand-alone research not only in data science but also in classical natural sciences, where it has been shown to have considerable predictive power in many highly complex phenomena. This is supported by the superior performance of Deep Learning in identifying and recognizing patterns from large data sets. A key strength of Deep Neural Networks (DNN) is that there is no requirement for feature extraction, effectively eliminating any subjective feature engineering. However, it is still a challenge (and generally a subjective choice) to design the architecture and topology of the network (i.e., the number of nodes and layers). Given the superior predictive power and flexibility of the method, it is not surprising that DNNs are also becoming popular in statistical seismology. In fact, the combination of the latest data acquisition techniques, together with Deep Learning, is giving new hope for earthquake prediction, a challenge that is still considered impossible by the majority of the scientific community.

Recently, there has been growing interest in using DNN to predict the spatial distribution of aftershocks following a major seismic event. There are two main reasons for this interest: first, it is a problem of great importance both from a scientific point of view and for seismic risk assessment and management; second, there are extensive collections of aftershock events that make the use of a data-driven approach attractive. In this perspective, the most widespread trend is to turn the problem of aftershock prediction into a classic classification problem. In particular, the volume surrounding the main aftershock event is divided into cells. Next, the input is defined by stress or kinematic variables for each single cell, and the output is either 1 or 0. In particular, a value of 1 is given (only) if there is at least one aftershock inside the cell. In these classification problems, performance is usually measured by the Area under the Receiver Operational Characteristics (AUC-ROC) curves. This approach has been used to define DNN even though it has been shown that their performance is not superior to a simple logistic regression. However, this binary aftershock classification generates unbalanced data sets with the vast majority of cells containing zeros. The problem, therefore, becomes a small event detection, and, in these cases, AUC-ROC curves may not be the appropriate metric to evaluate the performance of a classifier.

This talk examines the area under the Precision-Recall Curves (which proved to be more appropriate for imbalanced data sets) to assess the performance of the different classifiers in the context of aftershock forecasts. Then, given these evaluations, we carefully examine the use of different stress-based metrics for aftershock classification, showing that their predictive power is related to convex combinations of tensor stress elements and not to specific physical patterns.

Keywords: Deep Learning, Aftershock prediction, Machine Learning

