EGU22-7363
https://doi.org/10.5194/egusphere-egu22-7363
EGU General Assembly 2022

SeisBench - A Toolbox for Machine Learning in Seismology

Jack Woollam[1], Jannes Münchmeyer[2,3], Frederik Tilmann[2,4], Andreas Rietbrock[1], Dietrich Lange[5], Thomas Bornstein[3], Tobias Diehl[6], Carlo Giunchi[7], Florian Haslinger[6], Dario Jozinovic[8], Alberto Michelini[8], Joachim Saul[2], and Hugo Soto[2]

[1]Karlsruhe Insitute of Technology, Geophysical Institute, Germany
[2]Deutsches GeoForschungZentrum (GFZ), Potsdam, Germany
[3]Humboldt-Universität zu Berlin, Berlin, Germany
[4]Freie Universität Berlin, Berlin, Germany
[5]GEOMAR Helmholtz Center for Ocean Research, Kiel, Germany
[6]ETH Zurich, Zurich, Switzerland
[7]Istituto Nazionale di Geofisica e Vulcanologia (INGV), Pisa, Italy
[8]Istituto Nazionale di Geofisica e Vulcanologia (INGV), Roma, Italy

Machine Learning (ML) methods have seen widespread adoption in seismology in recent years. The ability of these techniques to efficiently infer the statistical properties of large datasets often provides significant improvements over traditional techniques when the number of data are large (>millions of examples). With the entire spectrum of seismological tasks, e.g., seismic picking and detection, magnitude and source property estimation, ground motion prediction, hypocentre determination; among others, now incorporating ML approaches, numerous models are emerging as these techniques are further adopted within seismology.

To evaluate these algorithms, quality controlled benchmark datasets that contain representative class distributions are vital. In addition to this, models require implementation through a common framework to facilitate comparison. Accessing both benchmark datasets, and integrating models built in such varying frameworks is currently a time-consuming process, hindering further advancement of ML techniques within seismology. These development bottlenecks also affect 'practitioners' seeking to deploy the latest models on seismic data, who may not want to necessarily learn entirely new ML frameworks to perform this task.

We present SeisBench as a software package to tackle these issues. SeisBench is an open-source framework for deploying ML in seismology. SeisBench standardises access to both models and datasets, whilst also integrating a range of common processing and data augmentation operations through the API. Through SeisBench, users can access several seismological ML models and benchmark datasets available in the literature via a single interface. SeisBench is built to be extensible, with community involvement encouraged to expand the package. Having such frameworks available for accessing leading ML models forms an essential tool for seismologists seeking to iterate and apply the next generation of ML techniques to seismic data.