

AI: Machine Learning for Geophysical Applications

Introduction

More and more Artificial Intelligence (AI) will play a role not only in society in general but also in the geosciences. Deep Learning resorts under the overall heading of Machine Learning / Artificial Intelligence. In this domain often the word “Algorithms” is used to indicate that computer algorithms are used to obtain results. Also, “Big Data” is mentioned, indicating that these algorithms need a large amount of training data to produce useful results.

Many scientists mention “Let the data speak for itself” when referring to Deep Learning, indicating that hidden or latent relationships between observations and classes or values of (desired) outcomes can be derived using these algorithms. Examples are in the field of seismic processing (first arrival picking), interpretation (facies prediction), etc. Often, we resort to statistical relationships. Then Deep Learning enters the game. From a range of labelled data (called instances) we can derive a linear/nonlinear relationship (model in DL terminology) that predicts the label or value (supervised learning) of new data (instances in DL terminology). But sometimes it is already useful if an algorithm can define separate groupings / clusters, which then still need to be interpreted (unsupervised learning). Even more sophisticated is Semi-supervised learning: labelled and unlabelled data together are clustered whereby the unlabelled data receives the label of the dominant class present in the cluster.

The Course

The aim of the course is to introduce how Machine Learning (ML) can be applied in geophysics. First, we will use a user-friendly package, called Weka. In that part of the course you will acquaint yourself with the workflows and algorithms used in Artificial Intelligence /Machine Learning. In the Deep Learning part, we will use open-source software: TensorFlow and Keras. Power-point presentations and videos will introduce various aspects of DL, but the emphasis is on computer-based exercises. The exercises deal with pre-conditioning the datasets (balancing the input classes, standardization & normalization of data) and applying several methods to classify the data: Multilayer Perceptron, Support Vector, Nearest Neighbour, AdaBoost, Trees. Non-linear Regression is used to predict porosity. Use will be made of Google Colab and Scikit-Learn. It runs on the Cloud and allows use of a GPU. It is “the way” to learn using a whole range of open-source Deep Learning algorithms. In the exercises you will get acquainted with using interactive python notebooks, how to get algorithms using Scikit-Learn and if you restrain yourself from using it on very large datasets, Google Colab is free. The course consists of many presentations and exercises as I am a strong believer in the paradigm: Tell me and I will forget, show me and I might remember, involve me (through exercises) and I will truly learn.

Learning methods and tools

At the end of the course participants will have a clear idea how Deep Learning, being part of Machine Learning / Artificial Intelligence will impact the future of Geosciences. This will be evident from the examples discussed and applied to

many data sets, including the prediction of lithology, porosity, pore fluids, saturations and facies. Quizzes are provided to enhance the learning.

Intended Audience

All those interested in understanding the impact Artificial Intelligence will have on the Geosciences. Hence, not only those involved in exploration and development of hydrocarbons or mineral resources, but also those involved in geothermal and CO₂ storage, where the relationship between data and targets are often difficult to establish.

Pre-requisites

A good understanding of geophysics and statistics.