Ch 1

Chapter 1 introduces the basics of Machine Learning to geologists. To do so, it first provides some fundamental definitions and common terms. Then it discusses the learning process and defines the different types of learning paradigms, i.e., supervised, unsupervised, and semisupervised.

Ch 2

Chapter 2 describes how to prepare a Python environment to start working with Machine Learning in Earth Sciences. First, it shows how to set up a local Python environment. Then, it moves to the creation of a remote Linux instance. Finally, it describes how to start working with cloud-based machine learning environments.

Ch3

Chapter 3 deals with the description of Machine Learning workflows. It starts introducing a typical five-step workflow made of 1) data acquisition, 2) pre-processing, 3) model training, 4) model validation, and 5) model deployment. Each step is described in detail also providing examples in Python.

Ch4

Chapter 4 introduced unsupervised machine learning methods. I start by describing the algorithms for dimensionality reduction, which include principal component analysis and manifold learning. Then it describes clustering methods, such as hierarchical clustering, DBSCAN, mean shift, K-means, spectral clustering, and Gaussian mixtures models.

Ch5

Chapter 5 describes how to apply unsupervised machine learning methods in Petrology. It focuses on clinopyroxene analyses erupted by Mt. Etna during the sequence of lava fountains that occurred between February and April of 2021. The application of clustering and dimensionality reduction techniques are described in detail.

Ch6

Chapter 6 deals with the application of unsupervised machine learning methods to multi-spectral images deriving from Earth-observing satellite missions. It describes how to import, pre-process, describe, and analyze multi-spectral data that can be downloaded by access points like USGS Earth Explorer, the Copernicus Open Access Hub, and Theia.

Ch7

Chapter 7 describes supervised learning methods for regression and classification tasks. They include Naive Bayes, Ridge Regression, Least Absolute Shrinkage and Selection Operator (LASSO), Elastic Net, Support Vector Machines, Supervised Nearest Neighbors, and Trees-Based Methods.

Ch8

Chapter 8 focuses on the classification of facies in well-log data by Machine Learning. It progressively develops a Machine Learning workflow that includes: descriptive statistics, algorithm selection, model optimization, model training, and the application to blind observations. Each step is discussed in detail.

Ch9

Chapter 9 deals with the application of Machine Learning regression in Petrology. In detail, it discusses how to calibrate Machine Learning thermo-barometers based on orthopyroxene crystals in equilibrium with the melt in a volcanic plumbing system. Also, it describes the calibration of a thermo-barometer based on orthopyroxenes crystals alone.

Ch10

Chapter 10 introduces basic concepts and definitions of parallel computing and model scaling. It starts by providing basic definitions and terminology. Then, it introduces Daks, a Python library that aims to provide object scalability to Python scientific libraries such as pandas, NumPy, and scikit-learn.

Ch11

Chapter 11 shows how to scale Machine Learning models in the cloud. In the context of cloud computing, the term scaling refers to the ability to quickly and efficiently increase (or decrease) the capability of a computational resource to handle a model that no longer fits the available resources.

Ch12

Chapter 12 is about deep learning. It starts by introducing the basics of deep learning. Then, it introduces PyTorch, a library for deep learning in Python. Also, it describes how to set up and train feedforward networks. Finally, it provides an example application dealing with deep learning potentials in the Earth Sciences.