Open and Automatic Machine Learning

FACULTY

Dr. ir. J. (Joaquin) Vanschoren

GENERAL DESCRIPTION OF THE MODULE

Building machine learning systems remains something of a (black) art, requiring a lot of prior experience to create appropriate machine learning systems (from machine learning pipelines to deep learning architectures) and their hyperparameters. In this module, we will learn how to turn this black art into a science.

First, we will learn about OpenML, an open science platform allowing anyone to share datasets and train many machine learning models from many software tools in a frictionless yet principled way. It organizes all results online, providing detailed insight into the performance of machine learning techniques, and allowing a more scientific, data-driven approach to building new machine learning systems.

Second, we will learn about automatic machine learning (AutoML), where we use machine learning to automatically build machine learning pipelines. We will cover the main techniques for automated hyperparameter optimization and pipeline synthesis. Finally, we will discuss meta-learning techniques, which *learn how to learn* from prior experiments (on other datasets) and apply that to build better models and pipelines automatically.

LEARNING OUTCOMES

- Using open datasets, pipelines and experiments from OpenML
- Sharing your own datasets, pipelines, and experiments on OpenML
- Solid understanding of the main techniques in automatic machine learning
- Leveraging OpenML and AutoML techniques to build better machine learning models, faster

TOOLS USED IN THIS MODULE

- A modern Python environment (e.g. Anaconda)
- sklearn, OpenML, AutoSKLearn, TPOT
- See the Prerequisites notebook for installation instuctions

PROGRAM

Time	Topic
9.00 - 10.30	Open Machine Learning
10.30 - 10.45	Coffee
10.45 - 12.15	OpenML: exercises
12.15 - 13.00	Lunch
13.00 - 14.30	Automatic Machine Learning
14.30 - 14.45	Coffee
14.45 - 15.45	AutoML: exercises