Machine learning: Methods and solutions

Hands-on course of 4 days - 28h
Ref.: MLB - Price 2024: €2 860 (excl. taxes)

EDUCATIONAL OBJECTIVES
At the end of the training, the trainee will be able to:

Understand the different learning models

Model a practical problem in abstract form

Identify relevant learning methods to solve a problem

Apply and evaluate the identified methods for a problem

Make the connection between different learning techniques

THE PROGRAMME

last updated: 06/2022

1) Introduction to Machine Learning
- Big Data and Machine Learning.
- Supervised, unsupervised and reinforcement learning algorithms.
- Steps for building a predictive model.
- Detecting outliers and handling missing data.
- How to choose the algorithm and its variables

Demonstration : Getting started in the Spark environment with Python using Jupyter Notebook. View several examples of the models provided.

2) Model evaluation procedures
- Techniques for resampling in training, validation and testing sets.
- Learning data representativeness test.
- Predictive model performance measurements.
- Confusion and cost matrix and AUC-ROC curve.

Hands-on work : Evaluation and comparison of different algorithms on the provided models.

3) Predictive models, the frequentist approach
- Statistical learning.
- Data conditioning and dimensionality reduction.
- Support vector machines and kernel methods.
- Vector quantization.
- Neural nets and Deep Learning
- Ensemble learning and decision trees.
- Bandits' algorithms, optimism in the face of uncertainty.

Hands-on work : Implementing algorithm families using various data sets.

4) Bayesian models and learning
- Principles of Bayesian inference and learning.
- Graphical models: Bayesian networks, Markov fields, inference and learning.
- Bayesian methods: Naive Bayes, mixtures of Gaussians, Gaussian processes.
- Markov models: Markov processes, Markov chains, hidden Markov chains, Bayesian filtering.

Hands-on work : Implementing algorithm families using various data sets.

5) Machine Learning in live environments

- Features related to the development of a model in a distributed environment.
- Big Data deployment with Spark and MLlib.
- The Cloud: Amazon, Microsoft Azure ML, IBM Bluemix, etc.
- Maintenance of the model.

Hands-on work : Taking a predictive model live, with integration into batch processes and processing flows.

DATES

REMOTE CLASS
2025 : 01 Apr, 03 Jun, 02 Sep, 30
Dec