

Machine learning

Introducing

Description

Program (detailed contents):

- Introduction to machine learning
- Risk estimation (or generalization error), Optimization of the bias/variance trade-off
- Model selection and variable selection via penalized criteria: Mallows CP, BIC, Ridge, Lasso
- Linear and quadratic discriminant analysis, k nearest neighbors.
- Binary classification and regression trees (CART)
- Tree aggregation, random forests
- Support Vector Machine and Support Vector Regression
- Neural networks, multilayer perceptron, gradient backpropagation algorithm, optimization algorithms, introduction to deep learning.
- Boosting algorithms
- Missing data imputation.
- Bias of algorithms
- Legal framework and societal impacts of AI.

Objectives

At the end of this module, the student will have understood and be able to explain (main concepts):

- The properties of the main learning methods and their limitations.
- The principle of bias-variance trade-off, model selection
- Algorithms and methods for estimating a risk (cross

validation, bootstrap..)

- Optimization and implementation of the main methods in R and Python (Scikit-learn).
- Ethical and legal principles of Artificial Intelligence.

The student will be able to:

- Analyze datasets from various domains (meteorology, marketing, industry, ...) with R and Python libraries.
- Implement the main methods and run the following learning algorithms: discriminant analysis, k-nearest neighbors, classification and regression trees, random forests, neural networks, SVM, boosting.
- Optimize the values of the hyper-parameters, automate the processing chain.
- Optimize the management of missing data.
- Detect legal or ethical flaws (bias, discrimination, opacity) in machine learning algorithms.

Necessary prerequisites

Statistical modelling

Introduction to R and Python languages

Évaluation

L'évaluation des acquis d'apprentissage est réalisée en continu tout le long du semestre. En fonction des enseignements, elle peut prendre différentes formes : examen écrit, oral, compte-rendu, rapport écrit, évaluation par les pairs...

Practical info

Location(s)

 Toulouse