Machine learning

ECTS

Volume horaire
55h

Présentation

Description

Programme (contenu détaillé):

- Introduction à l'apprentissage machine
- Estimation du risque (ou erreur de généralisation), Optimisation du compromis biais / variance
- Sélection de modèles et sélection de variables via des critères pénalisés : CP de Mallows, BIC, Ridge, Lasso
- Analyse discriminante linéaire et quadratique, k plus proches voisins.
- Arbres binaires de classification et de régression (CART)
- Agrégation d'arbres, forêts aléatoires
- Support Vector Machine et Support Vector Regression
- Réseaux de neurones, perceptron multicouches, algorithme de rétropropagation du gradient, algorithmes d'optimisation, introduction à l¿apprentissage profond.
- Algorithmes de boosting
- Imputation de données manquantes.
- Biais des algorithmes
- Cadre juridique et impacts sociétaux de l'IA.

Objectifs

A la fin de ce module, l'étudiant.e devra avoir compris et pourra expliquer (principaux concepts) :

- Les propriétés des principales méthodes d'apprentissage et leurs limites.
- Le principe du compromis biais-variance, de la sélection de modèles
- Les algorithmes et les méthodes d'estimation d'un risque (validation croisée, bootstrap..)
- L'optimisation et l'implémentation des principales méthodes en R et Python (Scikit-learn).
- Les principes éthiques et juridiques de l'Intelligence Artificielle.

L'étudiant.e devra être capable de :

- Analyser des jeux de données issus de divers domaines (météorologie, industrie, IoT ...) avec les librairies de R et Python.
- Mettre en œuvre les principales méthodes et exécuter les algorithmes d'apprentissage suivants : analyse discriminante, k plus proches voisins, arbres de classification et de régression, forêts aléatoires, réseaux de neurones, SVM, boosting.
- Optimiser les valeurs des hyper-paramètres, automatiser la chaîne des traitements.
- Optimiser la gestion des données manquantes.
- Détecter les failles légales ou éthiques (biais, discrimination, opacité) des algorithmes d'apprentissage automatique.

Pré-requis nécéssaires

Éléments de Modélisation statistique

Introduction aux langages R et Python

Informations non contractuelles

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...

Infos pratiques

Lieu(x)

Toulouse