



Introduction au machine learning et au deep learning, mise en oeuvre en Python

Objectifs et contenu

- Maîtriser les principaux algorithmes de machine learning et deep learning pour l'apprentissage supervisé
- Comprendre les concepts et fonctionnements des algorithmes
- Être capable de les mettre en oeuvre avec Python
- Être capable de choisir les algorithmes de machine learning selon les cas d'usages
- Savoir évaluer les performances des algorithmes de machine learning

L'apprentissage machine (machine learning) est une discipline scientifique qui s'intéresse à la conception et au développement d'algorithmes permettant aux ordinateurs d'apprendre à prendre des décisions à partir de données. L'ensemble des données possibles qui alimentent une tâche d'apprentissage peut être très vaste et varié, ce qui rend la modélisation et les hypothèses préalables critiques pour la conception d'algorithmes pertinents.

Ce stage se concentre sur la méthodologie sous-jacente à l'apprentissage supervisé avec un accent particulier sur la formulation mathématique des algorithmes et la façon dont ils peuvent être mis en oeuvre et utilisés dans la pratique.

- Introduction à l'apprentissage supervisé : régression et classification binaire, métriques d'évaluation classiques et quelques « recettes de cuisine » (cross-validation, overfitting)
- Méthodes linéaires : LDA, modèles linéaires généralisés, régression logistique. SVM linéaire : hinge loss, méthodes de régularisation ridge, lasso, problèmes en grande dimension
- Méthodes non-linéaires : arbres de décision, CART, boosting, XGBoost, LightGBM, random forests, méthodes à noyaux
- Méthodes d'optimisation pour le machine learning : coordinate gradient descent, descente de gradient stochastique et améliorations
- Deep learning : feed-forward neural networks, convolutional neural networks, back-propagation, algorithmes stochastiques pour l'entraînement, early stopping, méthodes d'initialisation, régularisation dropout, mise en oeuvre avec tensorflow et keras

Alternance de cours (50 %) et d'ateliers de mise en pratique (50 %)

Intervenant(s)

Stéphane GAÏFFAS (Professeur au LPSM, Université Paris Diderot et à l'école Polytechnique)

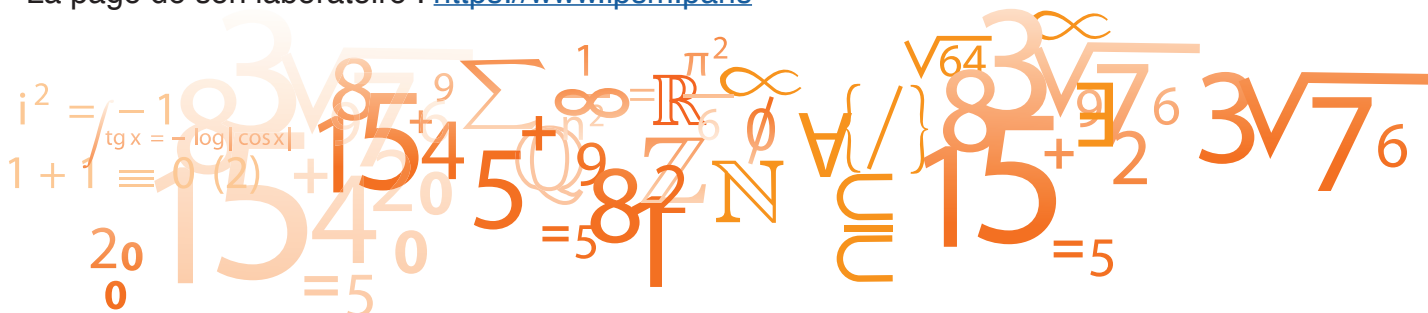
Sa page personnelle : <http://www.cmap.polytechnique.fr/~gaiffas/>

La page de son laboratoire : <https://www.lpsm.paris>

Karine TRIBOULEY (Professeur au LPSM, Université Paris Diderot)

Sa page personnelle : <https://www.lpsm.paris/laboratoire/annuaire/tribouley/>

La page de son laboratoire : <https://www.lpsm.paris>



Public visé

Chercheurs et ingénieurs en lien avec des problématiques liées à la science des données, allant de la science expérimentales aux sciences humaines, data-miners, informaticiens, développeurs et statisticiens intéressés par l'état de l'art en science des données.

Prérequis

Connaissance d'au moins un langage de programmation ; connaissance de bases en probabilités et statistiques.

Dates du stage, durée, emploi du temps et organisation

3 jours du mercredi 18/03/2020 au vendredi 20/03/2020

De 10 à 25 stagiaires

TD encadrés par un intervenant pour 10 stagiaires maximum

Lieu

Paris.

Coût pédagogiques

1800 Euros.

Infrastructure requise

Il est demandé aux stagiaires de venir avec leur propre ordinateur sur lequel les logiciels et outils (libres) nécessaires à la formation (communiqués avant le début du stage) seront préalablement installés.

A l'issue de la formation

Evaluation de la formation par les stagiaires.

Envoi d'une attestation de formation.

