

# INTRODUCTION AU MACHINE LEARNING ET À L'IA

FFCNCERCERXBD04

PRIX : 1 910 €

DURÉE : 2 JOURS

Pauses et déjeuners offerts



AVANCÉ



ATELIER



FAISABLE À DISTANCE



RÉALISABLE EN ANGLAIS

## PRÉSENTATION

Les sciences de l'information transforment aujourd'hui de nombreux domaines tels que la médecine, la finance, le marketing, la culture, la défense ou encore la sécurité. La capacité à traiter et analyser des données massives (big data) est un enjeu majeur dans tous les secteurs d'activités. La discipline qui développe et étudie des méthodes concrètes pour modéliser ce type de données est le machine learning. A l'interface des mathématiques appliquées et de l'informatique, cette discipline vise à produire des outils de prédiction et d'aide à la décision automatisés, à partir d'un apprentissage sur des données et des événements. L'apparition d'algorithmes très performants pour la classification de données en grande dimension, tels que les réseaux de neurones, a progressivement transformé le champ occupé jusqu'alors par la statistique traditionnelle. Celle-ci s'appuyait en grande partie sur le prétraitement réalisé par l'opérateur humain.

La formation présente le domaine et ses fondements. Ses problèmes et ses méthodes les plus récentes sont également étudiés. Les concepts sont illustrés par des applications variées.

## OBJECTIFS

- Identifier les fondements historiques et théoriques de l'apprentissage automatique
- Utiliser les algorithmes majeurs en machine learning et reconnaître les situations dans lesquelles les appliquer

## PROGRAMME

### Introduction

#### Problématiques traitées par le Machine Learning

- Classification
- Régression
- Réduction de dimension
- Critères de choix d'algorithme

#### Introduction à la théorie de l'apprentissage (principe de la minimisation de risque empirique) – Exemples applicatifs :

- LDA (Linear Discriminant Analysis) – supervisé
- K-means – non supervisé

#### Sur-apprentissage, sélection de modèle

#### Optimisation pour le Machine Learning

#### Régression logistique

#### Arbres de classification

#### Méthodes ensemblistes

#### Support Vector Machine (SVM)

#### Introduction aux réseaux de neurones/ deep learning

- Traitement du langage naturel NLP
- Traitement d'images

#### Travaux pratiques avec Python, ScikitLearn

#### Synthèse et conclusion

## PUBLIC/PRÉREQUIS

Ingénieurs et techniciens ayant besoin de méthodes d'apprentissage pour automatiser des tâches (prédiction, décision, etc.) de manière performante. Chefs de projets qui souhaitent mieux identifier les tâches que le machine learning permettrait d'automatiser.

Des bases en Python (opérateurs clés (if, for, etc.), listes, dictionnaires) et des connaissances en mathématiques (fonctions et dérivées, limites, norme et produit scalaire) sont nécessaires afin de tirer pleinement profit de cette formation.

## RESPONSABLE(S)

### Stéphan CLÉMENÇON

Enseignant-chercheur au département « Traitement du Signal et de l'Image » de Télécom Paris, ses recherches portent sur la théorie statistique de l'apprentissage. Il est coordinateur de la chaire « Machine Learning for big data » et enseigne le « machine-learning » à Télécom Paris, à l'ENSAE Paris, à l'Université Paris 7 et à l'ENS Cachan.

## MODALITÉS PÉDAGOGIQUES

Travaux pratiques en Python :

- Initiation à Scikit-Learn
- Application des algorithmes
- Compréhension des mesures de performances et des avantages/inconvénients des différentes méthodes.