

« Master Class Machine Learning »

Programme de formation

Description :

Cette formation vous permettra d'acquérir les bases du Machine Learning (apprentissage automatique à partir de données).

Vous y apprendrez les fondamentaux théoriques, les cas d'usages ; vous saurez quand le Machine Learning est une solution pertinente.

Et vous mettrez en pratique ces aspects théoriques pour réaliser vos propres modèles prédictifs en python, sur des exemples concrets (détection de fraudes aux examens, détection de spam, etc.).

La formation s'appuiera sur le livre "Data Science : fondamentaux et études de cas"

Public visé :

- Le statisticien souhaitant comprendre comment le Machine Learning peut compléter et augmenter les études statistiques classiques
- L'informaticien appelé à programmer des applications prédictives
- Le décideur désirant comprendre, par la mise en pratique, les enjeux et les possibilités du Machine Learning

Prérequis :

Pour tirer parti de cette formation, quelques savoir-faire préalables sont souhaitables :

- Pour les exercices pratiques, une connaissance préalable de python : programmation élémentaire de scripts, variables, boucles, etc.
- De façon générale, une habitude de l'utilisation de l'informatique (tableurs, manipulation de données, fichiers, etc.) est indispensable.
- Des connaissances élémentaires en mathématiques sont un plus pour tirer profit des parties théoriques de la formation : notamment statistiques, calcul, ainsi que des rudiments d'analyse (dérivée, équations, etc.)

Objectifs pédagogiques :

- Comprendre les cas d'application et les possibilités du Machine Learning
- Savoir construire ses propres modèles prédictifs
- Savoir évaluer la pertinence d'un modèle prédictif
- Comprendre et savoir choisir le bon modèle prédictif et ses paramètres
- Savoir intégrer de tels modèles prédictifs à un système existant

Matériel :

- Salle de travail : connexion internet, vidéo projecteur, écran

Plan de formation

Intervenant : Christophe LEGAL

Journée 1

9h00 à 17h30

Description :

Introduction théorique

- Machine Learning et optimisation
- Machine Learning et corrélations statistiques
- Machine Learning et data science. Cas d'applications concrets
- Introductions aux réseaux de neurones, aux algorithmes génétiques, et aux arbres de décisions (XGBoostedTree)

Premiers exemples pratiques

- Introduction à python scikit-learn et DataRobot
- Programmation de notre propre réseau de neurone
- Utilisation de bibliothèques dédiées
- Méthodes d'optimisation de réseaux de neurones
- Apprentissages de premiers modèles prédictifs

Journée 2

9h00 à 17h30

Description :

Performances et validité de l'apprentissage

- Mesure de précision, de discrimination, etc.
- Risque de sur-apprentissage
- Base d'apprentissage et base de validation
- Application pratique : détection de fraude de transactions bancaires

Éléments influençant les performances et la validité

- Complexité du modèle
- Rugosité du modèle
- Hyper-paramétrisation
- Meta-learning ; risque de meta-surapprentissage

Description :**Application : détection automatique de spam**

- Traitement de la langue naturelle
- Analyse de fréquence
- Mise en pratique sous python

Intégration d'un module de machine-learning à un système existant

- Flux de fichiers
- Utilisation d'API
- Application : intégration de la détection du spam à une application de messagerie

Pour aller plus loin

- Préparation des données
- Pré-filtrage des entrées
- Réseaux convolutifs
- Apprentissage semi-supervisé et non-supervisé

Coût de la formation : 2 390 euros HT par personne

LE FORMATEUR**Christophe LEGAL**

Consultant en Intelligence Artificielle chez Kereon Intelligence,

Christophe LE GAL est docteur en intelligence artificielle depuis 1999 (INPG) et ingénieur en calcul scientifique (ENSIMAG).

Sa carrière s'est déroulée à la fois dans le secteur privé, où il a été directeur général pendant une décennie d'une entreprise spécialisée dans la simulation numérique, et dans le milieu académique.

Son activité actuelle prend place dans trois contextes : il enseigne l'intelligence artificielle, la recherche opérationnelle et l'algèbre à l'École nationale d'ingénieur de Brest (ENIB). Il est chercheur associé au LATIM, où il développe des méthodes mêlant intelligence artificielle, machine learning et statistiques pour l'amélioration du diagnostic de maladies auto-immunes. Et enfin, il met en œuvre ces méthodes d'analyse pour divers clients, par exemple pour créer des modèles prédictifs de comportement.

