

Machine learning, méthodes et solutions

Cours Pratique de 4 jours

Réf : MLB - Prix 2021 : 2 710€ HT

Le Machine Learning couvre l'ensemble des méthodes et concepts qui permettent d'extraire automatiquement à partir de données, des modèles de prédiction et de prise de décision. Durant ce cours, vous mettrez en œuvre les différents algorithmes du domaine et appréhendez les bonnes pratiques d'un projet Machine Learning.

OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation l'apprenant sera en mesure de :

Comprendre les différents modèles d'apprentissage

Modéliser un problème pratique sous forme abstraite

Identifier les méthodes d'apprentissage pertinentes pour résoudre un problème

Appliquer et évaluer les méthodes identifiées sur un problème

Faire le lien entre les différentes techniques d'apprentissage

LE PROGRAMME

dernière mise à jour : 06/2021

1) Introduction au Machine Learning

- Le Big Data et le Machine Learning.
- Les algorithmes d'apprentissage supervisés, non supervisés et par renforcement.
- Les étapes de construction d'un modèle prédictif.
- Détecter les valeurs aberrantes et traiter les données manquantes.
- Comment choisir l'algorithme et les variables de l'algorithme ?

Démonstration : Prise en main de l'environnement Spark avec Python à l'aide de Jupyter Notebook. Visualiser plusieurs exemples de modèles fournis.

2) Procédures d'évaluation de modèles

- Les techniques de ré-échantillonnage en jeu d'apprentissage, de validation et de test.
- Test de représentativité des données d'apprentissage.
- Mesures de performance des modèles prédictifs.
- Matrice de confusion, de coût et la courbe ROC et AUC.

Travaux pratiques : Evaluation et comparaison des différents algorithmes sur les modèles fournis.

3) Les modèles prédictifs, l'approche fréquentiste

- Apprentissage statistique.
- Conditionnement des données et réduction de dimension.
- Machines à vecteurs supports et méthodes à noyaux.
- Quantification vectorielle.
- Réseaux de neurones et Deep Learning.
- Ensemble learning et arbres de décision.
- Les algorithmes de Bandits, optimisme face à l'incertitude.

Travaux pratiques : Mise en œuvre des familles d'algorithmes en utilisant des jeux de données variés.

PARTICIPANTS

Ingénieurs/chefs de projet souhaitant considérer les techniques d'apprentissage automatique dans la résolution de problèmes industriels.

PRÉREQUIS

Connaissances de base en Python et en statistiques de base (ou connaissances équivalentes à celles apportées par le stage "Statistiques, maîtriser les fondamentaux" (Réf. STA)).

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la pédagogie, et ce pour chaque cours qu'ils enseignent. Ils ont au minimum cinq à dix années d'expérience dans leur domaine et occupent ou ont occupé des postes à responsabilité en entreprise.

MODALITÉS D'ÉVALUATION

Les apports théoriques et les panoramas des techniques et outils ne nécessitent pas d'avoir recours à une évaluation des acquis.

MOYENS PÉDAGOGIQUES ET TECHNIQUES

- Les moyens pédagogiques et les méthodes d'enseignement utilisés sont principalement : aides audiovisuelles, documentation et support de cours, exercices pratiques d'application et corrigés des exercices pour les stages pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
- À l'issue de chaque stage ou séminaire, ORSYS fournit aux participants un questionnaire d'évaluation du cours qui est ensuite analysé par nos équipes pédagogiques.
- Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le stagiaire a bien assisté à la totalité de la session.

4) Les modèles et apprentissages bayésiens

- Principes d'inférence et d'apprentissage bayésiens.
- Modèles graphiques : réseaux bayésiens, champs de Markov, inférence et apprentissage.
- Méthodes bayésiennes : Naive Bayes, mélanges de gaussiennes, processus gaussiens.
- Modèles markoviens : processus markoviens, chaînes de Markov, chaînes de Markov cachées, filtrage bayésien.

Travaux pratiques : Mise en œuvre des familles d'algorithmes en utilisant des jeux de données variés.

5) Machine Learning en production

- Les spécificités liées au développement d'un modèle en environnement distribué.
- Le déploiement Big Data avec Spark et la MLlib.
- Le Cloud : Amazon, Microsoft Azure ML, IBM Bluemix...
- La maintenance du modèle.

Travaux pratiques : Mise en production d'un modèle prédictif avec l'intégration dans des processus de batch et dans des flux de traitements.

LES DATES

AIX

2021 : 24 août, 19 oct., 02 nov.

ANGERS

2021 : 31 août, 19 oct.

BORDEAUX

2021 : 05 oct., 02 nov.

BRUXELLES

2021 : 24 août, 05 oct., 02 nov.

DIJON

2021 : 31 août, 12 oct.

GRENOBLE

2021 : 12 oct., 07 déc.

LILLE

2021 : 05 oct., 02 nov.

LIMOGES

2021 : 31 août, 12 oct.

LUXEMBOURG

2021 : 05 oct.

LYON

2021 : 28 sept., 30 nov.

MONTPELLIER

2021 : 24 août, 07 sept., 14 déc.

NANCY

2021 : 17 août, 07 déc.

NANTES

2021 : 24 août, 26 oct., 02 nov.

NIORT

2021 : 31 août, 19 oct.

ORLEANS

2021 : 17 août, 28 sept., 02 nov.

PARIS

2021 : 24 août, 28 sept., 26 oct.,
02 nov., 14 déc.

REIMS

2021 : 17 août, 07 déc.

RENNES

2021 : 26 oct., 07 déc.

ROUEN

2021 : 07 sept., 05 oct.

SOPHIA

2021 : 24 août, 19 oct., 02 nov.

STRASBOURG

2021 : 05 oct., 02 nov.

SUISSE

2021 : 14 sept., 02 nov.

TOULON

2021 : 07 sept., 05 oct.

TOULOUSE

2021 : 24 août, 26 oct., 02 nov.

TOURS

2021 : 17 août, 28 sept., 02 nov.

CLASSES À DISTANCE

2021 : 24 août, 28 sept., 26 oct.,
26 oct., 02 nov., 02 nov., 14 déc.