

# Microsoft Azure Machine Learning, développer et exploiter des algorithmes

Cours Pratique de 3 jours - 21h

Réf : AZL - Prix 2024 : 2 280CHF HT

Les algorithmes s'imposent comme l'un des sujets prédominants du Big Data. Ce sont les outils des méthodes exploratoires, explicatives ou prédictives que l'on applique aux données, dans le cadre du Machine Learning. Ce cours vous permettra d'acquérir les compétences nécessaires à l'utilisation d'Azure Machine Learning.

## OBJECTIFS PÉDAGOGIQUES

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

Prendre en main l'interface d'Azure Machine Learning

Choisir parmi plusieurs algorithmes équivalents selon une problématique

Découvrir les bases des langages R et Python pour augmenter les capacités d'Azure Machine Learning

Exploiter une expérience au travers d'un Web Service

## EXERCICE

Cas pratiques sur des données réalistes et volumineuses

## LE PROGRAMME

dernière mise à jour : 04/2020

### 1) Prise en main de l'interface Azure Machine Learning

- L'offre Azure. Facturation à l'usage.
- Prise en main de l'interface Machine Learning Studio.
- Créer un dataset. Se connecter à une source de données.
- Construire une expérience de ML.
- Définir un Web Service prédictif.
- La Gallery Cortana Intelligence.

*Travaux pratiques : Prise en main de l'interface Azure ML. Création d'un dataset. Définition d'un Web Service prédictif.*

### 2) Créer une expérience de Machine Learning

- Utiliser l'arbre de choix des algorithmes.
- Détecter les valeurs aberrantes.
- Choisir les variables de l'algorithme (features sélection).
- Initialiser le modèle, entraîner le modèle, évaluer le modèle.
- Reformuler un modèle prédictif.
- Transformer les variables de l'algorithme (features engineering).
- Limiter les lignes d'un jeu de données.

*Travaux pratiques : Évaluer des différents algorithmes à l'aide de la courbe ROC.*

### 3) Savoir paramétrer les grandes familles d'algorithmes

- Algorithmes de clustering (approche non supervisée).
- Algorithmes de régression linéaire.
- Algorithmes de régression logistique ou ordinaire.

## PARTICIPANTS

Data Scientists, data miners, statisticiens, développeurs en charge de la mise en production des modèles.

## PRÉREQUIS

Connaissances de base en statistiques (centrage, dispersion, corrélation, tests d'hypothèses). Des notions de programmation ou d'algorithmique peuvent être utiles.

## 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

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

## 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.

## MODALITÉS ET DÉLAIS D'ACCÈS

L'inscription doit être finalisée 24 heures avant le début de la formation.

## ACCESSIBILITÉ AUX PERSONNES HANDICAPÉES

Vous avez un besoin spécifique d'accessibilité ? Contactez Mme FOSSE, référente handicap, à l'adresse suivante psh-accueil@orsys.fr pour étudier au mieux votre demande et sa faisabilité.

- Algorithmes de classification (approche supervisée) binaire ou one-versus-all.
- Méthodes ensemblistes (forêt, jungle...).
- Packages R et Python. Le framework Vowpal Wabbit.
- Paramétrage des algorithmes.

*Travaux pratiques : Paramétrer des familles d'algorithmes avec R/Python.*

#### 4) Traiter d'autres types de données

- Analyser les séries temporelles, détecter les anomalies.
- Analyse de données textuelles avec les packages R.
- Appliquer un algorithme Vowpal Wabbit (Latent Dirichlet Analysis).
- Exploiter les images avec notebooks Jupyter.

*Travaux pratiques : Traitement des données texte ou image.*

#### 5) Découvrir les nouveaux outils autour d'Azure Machine Learning

- Nouvelles briques Azure pour le ML (Experimentation / Model Management).
- Inspection et préparation des données (transformations par exemple, transformations avancées).
- Mise en œuvre des instances Azure Machine Learning.
- Suivi des exécutions et des métriques d'évaluation.
- Scénarios de déploiement (local/Spark/Docker/AKS).

*Travaux pratiques : Préparation de données et transformations avancées.*

## LES DATES

---

CLASSE À DISTANCE

2024 : 27 mai, 02 oct., 11 déc.