



Machine Learning

Mise à jour nov. 2023

Durée 2 jours (14 heures)

« Délai d'accès maximum 1 mois »

OBJECTIFS PROFESSIONNELS

- Comprendre comment organiser le traitement des données et structurer les processus de Machine Learning.

PARTICIPANTS

- Chefs de projet, développeurs, data scientists, architectes

PRE-REQUIS

- Posséder une culture informatique de base.

MOYENS PEDAGOGIQUES

- Réflexion de groupe et apports théoriques du formateur
- Travail d'échange avec les participants sous forme de
- Utilisation de cas concrets issus de l'expérience professionnelle
- Validation des acquis par des questionnaires, des tests d'évaluation, des mises en situation et des jeux pédagogiques.
- Remise d'un support de cours.

MODALITES D'EVALUATION

- Feuille de présence signée en demi-journée,
- Evaluation des acquis tout au long de la formation,
- Questionnaire de satisfaction,
- Positionnement préalable oral ou écrit,
- Evaluation formative tout au long de la formation,
- Evaluation sommative faite par le formateur ou à l'aide des certifications disponibles,
- Sanction finale : Certificat de réalisation, certification éligible au RS selon l'obtention du résultat par le stagiaire

MOYENS TECHNIQUES EN PRESENTIEL

- Accueil des stagiaires dans une salle dédiée à la formation, équipée d'ordinateurs, d'un vidéo projecteur d'un tableau blanc et de paperboard. Nous préconisons 8 personnes maximum par action de formation en présentiel

MOYENS TECHNIQUES DES CLASSES EN CAS DE FORMATION DISTANCIELLE

- A l'aide d'un logiciel comme Teams, Zoom etc... un micro et éventuellement une caméra pour l'apprenant,
- suivez une formation uniquement synchrone en temps réel et entièrement à distance. Lors de la classe en ligne, les apprenants interagissent et communiquent entre eux et avec le formateur.
- Les formations en distanciel sont organisées en Inter-Entreprise comme en Intra-Entreprise.
- L'accès à l'environnement d'apprentissage (support de cours, labs) ainsi qu'aux preuves de suivi et d'assiduité (émargement, évaluation) est assuré. Nous préconisons 4 personnes maximum par action de formation en classe à distance

ORGANISATION

- Les cours ont lieu de 9h à 12h30 et de 14h à 17h30.

PROFIL FORMATEUR

- Nos formateurs sont des experts dans leurs domaines d'intervention
- Leur expérience de terrain et leurs qualités pédagogiques constituent un gage de qualité.

A L'ATTENTION DES PERSONNES EN SITUATION DE HANDICAP

- Les personnes atteintes de handicap souhaitant suivre cette formation sont invitées à nous contacter directement, afin d'étudier ensemble les possibilités de suivre la formation.

Programme de formation

Introduction au machine learning (03h00)

- Le Big Data et le machine learning.
- Zoom sur les données et les requêtes, attentes, des utilisateurs
- Les étapes de la préparation des données.
- le data munging/wrangling
- Le rôle du data scientist

Le Machine Learning (02h30)

- Définition, les attentes par rapport au Machine Learning
- Les valeurs d'observation, et les variables cibles.
- Ingénierie des variables
- Comment choisir l'algorithme et les variables de l'algorithme.

Apprentissage automatique (03h00)

- Les algorithmes d'apprentissage supervisés, non supervisés
- Classification des données
- Les étapes de construction d'un modèle prédictif.
- Détecter les valeurs aberrantes et traiter les données manquantes.
- Algorithmes : régression linéaire, k-voisins, classification naïve bayésienne, arbres de décision, etc ..

Les risques et écueils (01h45)

- Importance de la préparation des données.
- L'écueil du « surapprentissage ».
- Les limites du Machine Learning

La visualisation des données (01h15)

- L'intérêt de la visualisation.
- Outils disponibles,

Machine learning en production (02h30)

- 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

Exemples de visualisation avec R et Python (00h00)