

# Formation Les fondamentaux du Machine Learning

Durée de la formation : 3 jour(s)

## OBJECTIFS

A l'issue de la formation, l'apprenant sera capable de :

Comprendre les différences entre apprentissage automatique supervisé, non supervisé et méta-apprentissage

Savoir transformer un gros volume de données à priori hétérogènes en informations utiles

Maîtriser l'utilisation d'algorithmes d'auto-apprentissage adaptés à une solution d'analyse

Comprendre comment exploiter de gros volumes de données textuelles

Être capable d'appliquer ces différentes techniques aux projets Big Data

## PREREQUIS

Connaître l'utilité du Data Mining et les problématiques du Big Data dans le ciblage économique

**Cette formation ne peut être financée que dans le cadre d'un projet d'entreprise (prise en charge entreprise ou OPCO). Les dossiers à financement personnel et CPF ne sont pas pris en compte.**

## PUBLIC

Ingénieurs, analystes, responsables marketing

Data Analysts, Data Scientists, Data Steward

Toute personne intéressée par les techniques de Data Mining et de Machine Learning

# Formation Les fondamentaux du Machine Learning

Durée de la formation : 3 jour(s)

## PROGRAMME

### L'APPRENTISSAGE MACHINE (INTRODUCTION)

Introduction  
Champs de compétences  
Focus Data Science (Data Mining)  
Focus Machine Learning  
Focus Big Data  
Focus Deep Learning  
Définition de l'apprentissage machine  
Exemples de tâches du machine Learning  
Que peuvent apprendre les machines  
Les différents modes d'entraînement

### LES FONDAMENTAUX DE L'APPRENTISSAGE MACHINE

Préambule : – Un problème d'optimisation – Quête de la capacité optimale du modèle – Relation capacité et erreurs – Un apport philosophique – Cadre statistique – Anatomie d'un modèle d'apprentissage machine  
Jeux de données d'entraînement : – Cadre statistique – Les variables prédictives – Chaîne de traitement des variables prédictives – Les variables à prédire  
Fonctions hypothèses : – Principe : jeux de fonctions hypothèses – Contexte de sélection des fonctions hypothèses – Caractéristiques des fonctions hypothèses – Modèles probabilistes Fréquentistes et Bayésiens  
Fonctions de coûts : – Les estimateurs – Principe du maximum de vraisemblance (MLE\*) – MAP – Maximum A Posteriori – Le biais d'un estimateur – La variance d'un estimateur – Le compromis biais – variance – Les fonctions de coûts – La régularisation des paramètres  
Algorithmes d'optimisations : – Les grandes classes d'algorithmes d'optimisation – La descente de gradient (1er ordre) – Descente de gradient (détails) – Les approches de Newton (2nd ordre) – Optimisation batch et stochastique – Pour aller plus loin  
Lab : Mise en oeuvre de l'environnement de travail machine Learning

### LA CLASSIFICATION

Introduction : – Choisir un algorithme de classification  
La régression logistique : – Du Perceptron à la régression logistique – Hypothèses du modèle – Apprentissage des poids du modèle – Exemple d'implémentation : scikit-learn – Régression logistique – Fiche Synthèse  
SVM : – Classification à marge maximum – La notion de marge souple (soft margin) – Les machines à noyau (kernel machines) – L'astuce du noyau (kernel trick) – Les fonctions noyaux – SVM – Maths – SVM – Fiche Synthèse  
Arbres de décision : – Principe de base – Fonctionnement – Maximisation du Gain Informationnel – Mesure d'impureté d'un noeud – Exemple d'implémentation : scikit-learn - Arbres de décision – Fiche Synthèse  
K plus proches voisins (kNN) : – L'apprentissage à base d'exemples – Principe de fonctionnement – Avantages et désavantages – kNN – Fiche synthèse  
Synthèse  
Lab : Expérimentation des algorithmes de classification sur cas concrets

### LES PRATIQUES

Prétraitement : – Gestion des données manquantes – Transformateurs et estimateurs – Le traitement des données catégorielles – Le partitionnement des jeux de données – Mise à l'échelle des données  
Ingénierie des variables prédictives (Feature Engineering) : – Sélection des variables prédictives – Sélection induite par régularisation L1 – Sélection séquentielle des variables – Déterminer l'importance des variables – Réduction dimensionnelle

# Formation Les fondamentaux du Machine Learning

**Durée de la formation : 3 jour(s)**

par Compression des données – L'extraction de variables prédictives – Analyse en composante principale (ACP) – Analyse linéaire discriminante (ADL) – l'ACP à noyau (KPCA)

Réglages des hyper-paramètres et évaluation des modèles : – Bonnes pratiques – La notion de Pipeline – La validation croisée (cross validation) – Courbes d'apprentissage – Courbes de validation – La recherche par grille (grid search) – Validation croisée imbriquée (grid searchcv) – Métriques de performance

Synthèse

Lab : Expérimentation des pratiques du machine learning sur cas concrets

## L'APPRENTISSAGE D'ENSEMBLES (ENSEMBLE LEARNING)

Introduction

L'approche par vote

Une variante : l'empilement (stacking)

Le bagging

Les forêts aléatoires

Le boosting

La variante Adaboost

Gradient Boosting

Fiches synthèses

Lab : L'apprentissage d'ensemble sur un cas concret

## LA RÉGRESSION

Régression linéaire simple

Régression linéaire multi-variée

Relations entre les variables

Valeurs aberrantes (RANSAC)

Évaluation de la performance des modèles de régression

La régularisation des modèles de régression linéaire

Régression polynomiale

La régression avec les forêts aléatoires

Synthèse

Lab : La régression sur un cas concret

## LE CLUSTERING

Introduction

Le regroupement d'objets par similarité avec les k-moyens (k-means)

k-means : algorithme

L'inertie d'un cluster

Variante k-means ++

Le clustering flou

Trouver le nombre optimal de clusters avec la méthode Elbow

Appréhender la qualité des clusters avec la méthode des silhouettes

Le clustering hiérarchique

Le clustering par mesure de densité DBSCAN

Autres approches du Clustering

Synthèse

Lab : Le clustering sur un cas concret

## LES PLUS DE CETTE FORMATION

# Formation Les fondamentaux du Machine Learning

**Durée de la formation : 3 jour(s)**

Une formation très pratique : 70% du temps de la formation est dédié à la mise en pratique pour une meilleure assimilation de notions de base.

Cette formation est basée sur des exercices principalement proposés par le formateur et tirés de l'ouvrage qui sert de support pour la formation.

Les travaux pratiques sont principalement réalisés avec R et Python.

Des consultants expérimentés partagent leur savoir-faire avec les participants.

# Formation Les fondamentaux du Machine Learning

## Durée de la formation : 3 jour(s)

### Prise en compte du handicap

Pour les personnes en situation de handicap, afin de nous permettre d'organiser le déroulement de la formation dans les meilleures conditions possibles, contactez-nous via notre formulaire de contact ou par mail (formation@access-it.fr) ou par téléphone (0320619506). Un entretien avec notre référente handicap pourra être programmé afin d'identifier les besoins et aménagements nécessaires.

### Modalités et moyens Pédagogiques, techniques et d'encadrement mis en œuvre

Répartition théorie/pratique : 45%/55%. Cette formation se compose d'une alternance d'apports théoriques, de travaux pratiques s'articulant autour d'une application fil rouge, de démonstrations, de phases d'échanges entre participants et de synthèses de la part du formateur.

Formation accessible à distance de n'importe où et n'importe quand, via un ordinateur type PC disposant d'une connexion à Internet à haut débit (ADSL ou plus).

Pour assurer un démarrage dans les meilleures conditions au premier jour de la formation, notre service logistique se met systématiquement en relation, en amont, avec vous afin de réaliser un test de validation technique et de vous présenter l'environnement de formation.

Pendant toute la durée de la formation, le stagiaire dispose d'une assistance technique et pédagogique illimitée, par e-mail, avec un délai de prise en compte et de traitement qui n'excède pas 24h. En complément, le stagiaire peut planifier un rendez-vous pédagogique avec un formateur expert afin d'échanger sur des éléments de la formation.

La durée de la formation affichée sur cette page est une durée estimée qui peut varier en fonction du profil du stagiaire et de ses objectifs (notamment s'il souhaite valider sa formation par le passage d'un examen de certification).

Durant la formation, le formateur prévoit :

Des démonstrations organisées en modules et en séquences découpées le plus finement possible, en suivant le programme pédagogique détaillé sur cette page ;

Des énoncés et corrigés de travaux pratiques à réaliser tout au long de la formation ;

Des travaux pratiques sont proposés ; la plateforme prévoit l'environnement technique nécessaire à la réalisation de l'ensemble des travaux pratiques ;

Le formateur valide les connaissances acquises après chaque TP ;

Il est proposé un ou plusieurs livres numériques faisant office d'ouvrage(s) de référence sur le thème de la formation.

### Validation et sanction de la formation

Une attestation mentionnant les objectifs, la nature et la durée de l'action et les résultats de l'évaluation des acquis de la formation sera remise au stagiaire à l'issue de sa formation par courrier électronique.

A la demande, il sera délivré un certificat de réalisation.

### Type de formation

Professionalisante ayant pour objectif le perfectionnement, l'élargissement des compétences

### Moyens permettant de suivre l'exécution de l'action

Le contrôle de la présence des stagiaires sera assuré par la vérification de l'assiduité des participants. Le stagiaire signera une feuille de présence par demi-journée de formation. Celle-ci sera également signée par le formateur.

### Modalité d'évaluation des acquis

Durant la formation, le participant est amené à mettre en pratique les éléments du cours par la réalisation de travaux pratiques réalisés sur PC.

La validation des acquis du stagiaire est faite par le formateur à la fin de chaque atelier. Cette validation individuelle est possible du fait du faible nombre de participants par session de formation (6 personnes maximum).

A la fin de la formation, le stagiaire a donc atteint les objectifs fixés par la formation.

En complément, pour les stagiaires qui le souhaitent, certaines formations peuvent être validées officiellement par l'éditeur en passant un examen de certification.

Access it étant centre d'examen ENI, les examens peuvent être réalisés sur demande à distance ou dans nos locaux de Villeneuve d'Aseq.

Les candidats à la certification doivent produire un travail personnel important en vue de se présenter au passage de l'examen, le seul suivi de la formation ne constitue pas un élément suffisant pour garantir un bon résultat et/ou l'obtention de la certification

### Assistance Post-Formation

Toute personne ayant suivi une formation avec Access it bénéficie d'une assistance post-formation d'une durée de 1 mois. Ce nouveau service d'accompagnement permet aux stagiaires rencontrant des difficultés dans la mise en œuvre des connaissances acquises de solliciter l'aide de nos formateurs sur des aspects relatifs aux programmes de formation suivis. Pour en bénéficier, il suffit de se rendre sur la page contact de notre site web et de remplir le formulaire. Une réponse est apportée par mail ou par téléphone dans un délai de 48 heures.

### Centre d'Examen

Notre centre de formation agréé Qualiopi bénéficie de l'agrément d'ENI et ICDL pour les certifications informatiques. C'est pour nos clients la garantie de pouvoir suivre des formations préparant à des certifications professionnelles.

### Aide à l'orientation

Pour chacune des grandes thématiques couvertes par notre offre de formation, nous proposons via nos spécialistes un rendez-vous physique ou téléphonique qui via un diagnostic permettra aux personnes souhaitant être accompagnées dans le choix d'un programme ou dans la définition d'un parcours de formation une orientation vers les programmes les plus adaptés à leurs besoins et à leur niveau.

### Aspects Pratiques

Dès leur inscription, les participants sont contactés par nos services qui s'assurent que les débits internet constatés sur le lieu depuis lequel ils souhaitent se former sont suffisants pour suivre la formation dans des conditions optimales.

À l'occasion de cet appel, nos experts s'assurent également qu'ils disposent du matériel nécessaire pour suivre la formation (PC Portable, webcam, Micro-casque..).

Avant le début de la formation, les participants reçoivent un lien leur permettant d'accéder à la classe virtuelle et leurs identifiants personnels de connexion. Un aide à l'utilisation de la solution de visioconférence utilisée leur est également proposée.

Le jour de la formation, ils se connectent à la classe virtuelle depuis leur navigateur internet. Ils voient et entendent le formateur ainsi que les autres participants et peuvent à tout moment communiquer avec eux.

Ils participent aux échanges et réalisent les ateliers dans les mêmes conditions que s'ils étaient en salle. Grâce à nos outils de prise en main à distance, les formateurs peuvent à tout moment prendre la main sur leurs postes pour les aider ou vérifier leurs TP.

Tout au long de la formation, les participants peuvent bénéficier de l'assistance immédiate de nos experts en composant le numéro qui leur a été communiqué avant la formation.

Des bilans intermédiaires ont lieu en présence des participants du formateur et du référent pédagogique d'Access it afin de vérifier l'état d'avancement de la session, les difficultés rencontrées et permettre d'éventuels actions correctives.

### Bénéfices pour les participants

Se former depuis leur lieu de travail ou leur domicile,

Accéder sans se déplacer à la qualité d'une formation délivrée par un formateur consultant ayant une expérience probante sur le sujet animé.

Bénéficier à distance de la richesse d'une formation interentreprises : échanges avec le formateur et les autres participants, partages d'expériences, ateliers pratiques...

Pouvoir se former en toutes circonstances et notamment en cas d'imprévu.

### Bénéfices pour l'entreprise

Optimiser ses budgets en limitant les frais de déplacement et d'hébergement.

Proposer à tous ses collaborateurs, quelle que soit leur situation géographique, des formations de qualité (en Inter comme en Intra).

Limiter les temps de déplacement.

Proposer davantage de choix dans les formations à des collaborateurs peu mobiles.

Assurer la montée en compétences de ses collaborateurs quelles que soient les circonstances