



Réf. TKBL031

Machine Learning avec Python

Développez des solutions de Machine Learning efficaces et robustes en utilisant Python et ses bibliothèques dédiées.

A propos de cette formation

Le Machine Learning (apprentissage automatique) constitue aujourd'hui l'une des techniques les plus efficaces pour extraire la connaissance et la valeur de la masse croissante des données collectées au sein des entreprises.

Découvrez avec un expert Quantmetry enseignant expérimenté comment Python, grâce à de nombreuses bibliothèques spécialisées et une communauté active vous permet d'implémenter les solutions de machine learning répondant à vos besoins.

Objectifs pédagogiques

- Découvrir les grands modèles utilisés en Machine Learning
- Comprendre les problèmes de classification, de régression et de clustering
- Savoir implémenter, mesurer et optimiser une solution de machine learning

Qu'allez-vous apprendre ?

A l'issue de cette formation mêlant théorie, exercices interactifs et études de cas guidés vous serez en mesure de :

Connaître les différentes typologies d'algorithmes de Machine Learning

Assimiler la démarche de mise en place d'un projet machine learning

Implémenter un modèle complexe de machine learning avec des outils

Mesurer et optimiser les grands algorithmes de machine learning

Faire du Textmining

Faire du Webscraping

Public & Prérequis

Toute personne ayant un profil technique souhaitant découvrir, implémenter et optimiser une solution de machine learning avec Python

Prérequis : Notions de Python

Bonus participants

- Fiches mémoires remises en séance résumant les points clés de la formation
- Ensemble de codes présentés lors de la séance

TARIFS

Durée : 3 jours - 21 h

Inter | 2 997 €

Intra | Sur devis

Prix HT par stagiaire. Petit-déjeuner d'accueil et repas du midi inclus

DATES

PARIS - ST AUGUSTIN



[Consulter les dates disponibles](#)

MODALITÉS



Présentiel

Exposés : 40%
Exercices : 60%

INTERVENANT



Guillaume
Expert IA / DATA

#tropicalstorm
#diveinstructor
#wingsuit

Docteur en imagerie satellitaire, expert en traitement de séries temporelles. Pédagogue expérimenté, il a enseigné en L3 et M2 en école d'ingénieur et pendant plus de 10 ans en cours particuliers



+20 Projets IA / DATAS à son actif



Programme de la journée

JOUR 1

Matin [3h] : Le machine learning, premières notions

Introduction au Machine Learning

Qu'est-ce que le Machine Learning ?

1. Cas d'usages du quotidien, utilisation du ML par les géants industriels d'aujourd'hui
2. Qu'est-ce que le Machine Learning ? Définitions et résumé des étapes clés
3. Le cas d'usages possibles dans les grands secteurs types, leurs spécificités et points en commun
4. Compromis performance vs. interprétabilité : faut-il toujours choisir entre les deux ?

M1 - Apprentissage supervisé – Processus de Data Science

Focus sur la typologie d'algorithme la plus fréquente : l'apprentissage supervisé

1. Définition de la cible et des variables explicatives. Construction du jeu de données et préparation des échantillons.
 2. Choix et entraînement du modèle : importance des statistiques descriptives et prévision de variables.
 3. Comment quantifier et présenter la performance d'un modèle tout en choisissant la métrique adaptée à notre besoin
- Travaux pratiques

Après-Midi [4h] : Apprentissage supervisé

M1 - Apprentissage supervisé – Gestion du sur-apprentissage

Comprendre et éviter le sur-apprentissage

1. Apprendre les tendances identifiées dans les données, mais pas par cœur : les problèmes de sous et sur-apprentissage
2. Échantillonnage des données : bien choisir le set d'entraînement et de test
3. Garder les caractéristiques des données dans ses échantillons : l'échantillonnage stratifié
4. Gérer la présence dans le test de valeurs jamais observées en phase d'entraînement
5. Mieux utiliser les données à disposition grâce à la validation croisée
6. Cas particulier : les séries temporelles

M2 - Apprentissage supervisé – Algorithmes classiques

Les grands algorithmes de l'apprentissage supervisé

1. Les arbres de décisions, et les modèles ensemblistes qui sont basés sur eux : forêts aléatoires et gradient boosting
2. Les support vector machines
3. Optimisation des hyperparamètres : grid search et randomized search
4. Quel rôle pour un set de validation

**Programme de la journée**

JOUR 2

Matin [3h] : Apprentissage non supervisé**M1 - Apprentissage non supervisé**

Que faire en l'absence d'une variable cible ?

1. Méthodologie et exemple d'apprentissage non supervisé : clustering avec k-means et DbSCAN
2. Coupler les approches supervisées et non supervisées afin d'obtenir des résultats plus exploitables.
3. Un problème non supervisé classique : la réduction de dimension. Analyse en composantes principales et feature selection

Après-Midi [4h] : Exploration de texte, web scraping**M1 - Text mining**

Exploiter des données textuelles

1. Clarification de la terminologie et des cas d'usage possibles : NLP, NLU, NLG, NER
2. Les sources de données possibles à utiliser dans un projet
3. Les étapes d'un prétraitement de texte : nettoyage, gestion des signes spéciaux, lemmatisation et racinisation
4. Principaux algorithmes : TF-IDF, Latent Dirichlet Analysis, Word2Vec, Doc2Vec, ...

M2 - Web scraping

Extraire les données du web, différentes approches et mise en pratique

1. Collecte des données web : utilisation des API vs. web scraping. Coûts et limites des deux approches.
2. Packages utilisés : urllib et BeautifulSoup. TP associé
3. Contraintes légales : on n'a pas le droit de faire tout ce qu'on est en mesure de faire !

**Programme de la journée**

JOUR 3

Matin [4h] : Étude de cas - Explorer la donnée**M1 - Appropriation du jeux de données**

Mise en condition sur un projet de Data Science

1. Présentation de la plateforme Kaggle et extraction d'un jeu de données
2. Exploration des données : types des colonnes / valeurs manquantes ou aberrantes, distribution des valeurs.
3. Visualisation des données : graphiques univariés et bivariés, corrélation entre features
4. Manipulation des variables : valeurs manquantes, outliers, gérer les variables catégorielles et feature engineering

Après-Midi [3h] : Étude de cas - modélisation et prédiction**M1 - Création de modèle**

Calibration d'un modèle

1. Création d'un échantillon d'entraînement et de validation, définition de la métrique de score.
2. Choix des modèles à essayer et leur entraînement. Mesures de performances sur le train set.
3. Calibration du modèle : validation croisée, hyperparamétrage (grid search).

M2 - Prédiction

Mesurer les performances du modèle

1. Prédiction sur le set de test. Bug-fixing en cas de problèmes (ex. variable catégorielle inconnue)
2. Mesures de performances des prédictions sur le set de test, et comparaisons avec les métriques sur le training set.
3. Étude des résultats, identification du modèle plus performant, discussion des possibles pistes d'amélioration.

Bulletin d'inscription



Présentiel

Réf. TKBL031

Machine Learning avec Python

3 jours - 21 h | 2 997 € HT

Petit-déjeuner d'accueil et repas du midi inclus

PARTICIPANT

Prénom - Nom*

Fonction*

Adresse email*

Téléphone mobile

En communiquant votre numéro vous recevrez gratuitement par SMS des informations et rappels relatifs à votre formation.


[Consulter les dates disponibles](#)

SOCIÉTÉ

Raison Sociale*

Si vous n'exercez pas en société saisissez votre statut (exemple : Indépendant)

SIRET / SIREN *

N° de TVA intracommunautaire

Adresse du siège social*

Adresse de facturation (si différente du siège)

Code Postal*

Ville*

Code Postal

Ville

Gestion administrative de la formation

Cadre réservé aux entreprises

Date et signature précédée de la mention « Bon pour commande »

Correspondant formation : Prénom - Nom *

Ligne directe*

Adresse email*

Cochez les cases correspondantes à votre situation :

La convention doit être établie à un autre nom que le mien

La formation fait l'objet d'une prise en charge par un organisme (OPCA)

La formation fait l'objet d'une subrogation de paiement

Bulletin à remplir et à retourner par email à l'adresse suivante :
inscription@institut.quantmetry.com

En soumettant ce formulaire par email vous acceptez que vos informations soient enregistrées et exploitées dans le cadre de votre demande d'inscription et de la relation commerciale qui peut en découler.

La signature de ce bulletin entraîne l'acceptation des conditions générales de vente disponibles en consultation sur le site Internet de l'Institut. L'inscription est effective dès réception de la confirmation d'inscription envoyée par l'Institut.

Institut Quantmetry est une marque de Quantmetry, SAS au capital de 1 024 € - 52 rue d'Anjou, 75008 Paris
RCS Paris n° 531172393 - TVA : FR27531172393 - N° d'agrément de formation : 11755601075