

# Formation : Machine learning, l'état de l'art

Séminaire - 2j - 14h00 - Réf. MLE  
Prix : 2020 CHF H.T.

★★★★☆ 4,4 / 5

Blended

Ce séminaire détaille les enjeux liés au traitement de la donnée par l'intelligence artificielle, et en particulier par les algorithmes du machine learning. Il montre aux décideurs, les principaux algorithmes du domaine, les solutions concrètes et la démarche de projet à appliquer selon les cas d'usages en entreprise.

## Objectifs pédagogiques

À l'issue de la formation, le participant sera en mesure de :

- ✓ Positionner le machine learning dans la chaîne de traitement de la donnée
- ✓ Distinguer les compétences nécessaires ou les profils à recruter
- ✓ Identifier les clés de réussite d'un projet autour du machine learning
- ✓ Comprendre les concepts d'apprentissage automatique et l'évolution du big data vers le machine learning
- ✓ Appréhender les enjeux de l'utilisation du machine learning, incluant les bénéfices attendus et des exemples d'usage

## Public concerné

Dirigeants d'entreprise (CEO, COO, CFO, SG, DRH...), DSI, les CDO, responsables informatiques, consultants, responsables de projets big data.

## Prérequis

Posséder une culture informatique générale, des notions de probabilités et statistiques sont recommandés.

Vérifiez que vous avez les prérequis nécessaires pour profiter pleinement de cette formation en faisant [ce test](#).

## Méthodes et moyens pédagogiques

### Méthodes pédagogiques

Illustration par des cas concrets. Présentation des principaux cas d'usage selon les secteurs d'activités (automobile, industrie, biens de consommation, finance, santé, énergie, agriculture, transports, télécommunication...).

### PARTICIPANTS

Dirigeants d'entreprise (CEO, COO, CFO, SG, DRH...), DSI, les CDO, responsables informatiques, consultants, responsables de projets big data.

### PRÉREQUIS

Posséder une culture informatique générale, des notions de probabilités et statistiques sont recommandés.

### 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.

## 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.

## Programme de la formation

### 1 Histoire du machine learning et contexte du big data

- Replacer à leur échelle les concepts d'intelligence artificielle, apprentissage automatique (machine learning)...
- Le lien avec les mathématiques, les statistiques (inférentielles), le data mining et la data science.
- Passer de l'analyse descriptive à l'analyse prédictive puis prescriptive.
- Les applications du Machine Learning (moteurs de recherche, détection des spams, lecture des chèques).
- La typologie des algorithmes de Dominique Cardon.
- La communauté data science et les challenges Kaggle (ex. de Netflix).

#### Etude de cas

Études d'applications concrètes du machine learning (moteurs de recherche, détection des spams, lecture des chèques).

### 2 Les données à disposition : collecte et préparation

- Données structurées, semi-structurées et non structurées.
- Nature statistique des données (qualitatives ou quantitatives).
- Objets connectés (IoT) et streaming.
- Opportunités et limites de l'open data.
- Identification des corrélations, problème de la multicollinéarité.
- Réduction des dimensions par Analyse des composantes principales.
- Détection et correction des valeurs aberrantes.
- Les ETL (Extract Transform Load).
- Le web scraping.

#### Démonstration

Démonstration d'un ETL (Extract Transform Load). Recueil de données web.

### 3 Les outils du marché pour le traitement de la donnée et le machine learning

- Les logiciels traditionnels (SAS, SPSS, Stata...) et leur ouverture à l'open source.
- Choisir entre les deux leaders open source : Python et R.
- Plateformes cloud (Azure, AWS, Google Cloud Platform) et solutions SaaS (IBM Watson, Dataiku).
- Nouveaux postes en entreprises : data engineer, data scientist, data analyst, etc.
- Associer les bonnes compétences à ces différents outils.
- Les API en ligne (IBM Watson, Microsoft Cortana Intelligence...).
- Les chatbots (agents conversationnels).

#### Démonstration

Démonstration d'un chatbot (agent conversationnel) et d'Azure Machine Learning.

#### 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 formations pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
- À l'issue de chaque formation 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 participant 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

Pour toute question ou besoin relatif à l'accessibilité, vous pouvez joindre notre équipe PSH par e-mail à l'adresse [psh-accueil@orsys.fr](mailto:psh-accueil@orsys.fr).

## 4 Les différents types d'apprentissage en machine learning

- Apprentissage supervisé : répéter un exemple.
- Apprentissage non supervisé : découvrir les données.
- Online (machine) learning par opposition aux techniques batch.
- Reinforcement learning : optimisation d'une récompense.
- Autres types d'apprentissage (par transfert, séquentiel, actif...).
- Illustrations (moteurs de recommandation...).

### Démonstration

Démonstrations sur les différents types d'apprentissage machine learning possibles.

## 5 Les algorithmes du machine learning

- Régression linéaire simple et multiple. Limites des approches linéaires.
- Régression polynomiale (LASSO). Séries temporelles.
- Régression logistique et applications en scoring.
- Classification hiérarchique et non hiérarchique (KMeans).
- Classification par arbres de décision ou approche Naïve Bayes.
- Random Forest (développement des arbres de décision).
- Gradient Boosting. Réseaux de neurones. Machine à support de vecteurs.
- Deep learning : exemples et raisons du succès actuel.
- Text mining : analyse des corpus de données textuelles.

### Démonstration

Démonstration des différents algorithmes de base sous R ou Python.

## 6 Procédure d'entraînement et d'évaluation des algorithmes

- Séparation du jeu de données : entraînement, test et validation.
- Techniques de bootstrap (bagging).
- Exemple de la validation croisée.
- Définition d'une métrique de performance.
- Descente de gradient stochastique (minimisation de la métrique).
- Courbes ROC et de lift pour évaluer et comparer les algorithmes.
- Matrice de confusion : faux positifs et faux négatifs.

### Démonstration

Démonstration du choix du meilleur algorithme.

## 7 Mise en production d'un algorithme de machine learning

- Description d'une plateforme big data.
- Principe de fonctionnement des API.
- Du développement à la mise en production.
- Stratégie de maintenance corrective et évolutive.
- Évaluation du coût de fonctionnement en production.

### Démonstration

Démonstration d'API de géolocalisation et d'analyse de sentiments.

## 8 Aspects éthiques et juridiques liés à l'intelligence artificielle

- Missions de la CNIL et évolutions à venir.
- Question du droit d'accès aux données personnelles.
- Question de la propriété intellectuelle des algorithmes.
- Nouveaux rôles dans l'entreprise : chief data officer et data protection officer.
- Question de l'impartialité des algorithmes.
- Attention au biais de confirmation.
- Les secteurs et les métiers touchés par l'automatisation.

### Réflexion collective

Réflexion en commun pour identifier les clés de réussite.

## Options

**Blended : 290 € HT**

Approfondissez les connaissances acquises en formation grâce aux modules e-learning de notre [Chaîne e-learning intelligence artificielle - machine learning, deep learning](#). Un apprentissage flexible et complet, à suivre à votre rythme dès le premier jour de votre présentiel.

## Dates et lieux

### CLASSE À DISTANCE

2026 : 18 juin, 1 oct., 24 nov.

### LAUSANNE

2026 : 18 juin, 1 oct.

### GENÈVE

2026 : 18 juin, 1 oct.