

<b>Stage pratique de 4 jour(s)</b> <b>Réf : MLB</b>	
<b>Participants</b> Ingénieurs/chefs de projet souhaitant considérer les techniques d'apprentissage automatique dans la résolution de problèmes industriels.	
<b>Pré-requis</b> Bonnes connaissances en statistiques de base, ou connaissances équivalentes à celles apportées par le stage "Statistiques, maîtriser les fondamentaux" (Réf. STA).	
<b>Prix 2019 : 2540€ HT</b>	
<b>Dates des sessions</b>	
<b>AIX</b> 21 jan. 2020, 26 mai 2020	
<b>BORDEAUX</b> 14 jan. 2020, 26 mai 2020	
<b>BRUXELLES</b> 07 jan. 2020, 12 mai 2020	
<b>GRENOBLE</b> 07 jan. 2020, 27 avr. 2020	
<b>LILLE</b> 07 jan. 2020, 12 mai 2020	
<b>LYON</b> 28 jan. 2020, 02 juin 2020	
<b>MONTPELLIER</b> 03 mar. 2020, 21 juil. 2020	
<b>NANTES</b> 07 jan. 2020, 27 avr. 2020	
<b>ORLEANS</b> 03 mar. 2020, 21 juil. 2020	
<b>PARIS</b> 19 nov. 2019, 10 déc. 2019 07 jan. 2020, 03 mar. 2020 12 mai 2020, 21 juil. 2020	
<b>RENNES</b> 21 jan. 2020, 26 mai 2020	
<b>SOPHIA-ANTIPOLIS</b> 07 jan. 2020, 27 avr. 2020	
<b>STRASBOURG</b> 21 jan. 2020, 26 mai 2020	
<b>TOULOUSE</b> 14 jan. 2020, 26 mai 2020	
<b>TOURS</b> 14 jan. 2020, 26 mai 2020	
<b>Modalités d'évaluation</b> L'évaluation des acquis se fait tout au long de la session au travers des multiples exercices à réaliser (50 à 70% du temps).	
<b>Compétences du formateur</b>	

## Machine learning, méthodes et solutions

*Le Machine Learning couvre l'ensemble des méthodes et concepts qui permettent d'extraire automatiquement à partir de données, des modèles de prédiction et de prise de décision. Vous mettrez en œuvre toute la chaîne de conception appliquée au Machine Learning dans un contexte Big Data batch et streaming.*

### OBJECTIFS PEDAGOGIQUES

Comprendre les différents modèles d'apprentissage  
 Modéliser un problème pratique sous forme abstraite  
 Identifier les méthodes d'apprentissage pertinentes pour résoudre un problème  
 Appliquer et évaluer les méthodes identifiées sur un problème  
 Faire le lien entre les différentes techniques d'apprentissage

[1\) Introduction au Machine Learning](#)  
[2\) Procédures d'évaluation de modèles](#)  
[3\) Les modèles prédictifs, l'approche fréquentiste](#)

[4\) Les modèles et apprentissages bayésiens](#)  
[5\) Machine Learning en production](#)

### Méthodes pédagogiques

*Chaque algorithme de la formation est d'abord présenté théoriquement suivi d'une démonstration dans chacun des trois langages R, Python et Scala.*

### Exercice

*Les participants sont libres de choisir R, Python ou Scala pour effectuer les exercices sur plusieurs jeux de données d'une très grande variété.*

### 1) Introduction au Machine Learning

- Le Big Data et le Machine Learning.
- Les algorithmes d'apprentissage supervisés, non supervisés et par renforcement.
- Les étapes de construction d'un modèle prédictif.
- Détecter les valeurs aberrantes et traiter les données manquantes.
- Comment choisir l'algorithme et les variables de l'algorithme ?

### Démonstration

*Prise en main de l'environnement Spark avec R, Python et Scala à l'aide de Jupyter Notebook. Visualiser plusieurs exemples de modèles fournis.*

### 2) Procédures d'évaluation de modèles

- Les techniques de ré-échantillonnage en jeu d'apprentissage, de validation et de test.
- Test de représentativité des données d'apprentissage.
- Mesures de performance des modèles prédictifs.
- Matrice de confusion, de coût et la courbe ROC et AUC.

### Travaux pratiques

*Evaluation et comparaison des différents algorithmes sur les modèles fournis.*

### 3) Les modèles prédictifs, l'approche fréquentiste

- Apprentissage statistique.
- Conditionnement des données et réduction de dimension.
- Machines à vecteurs supports et méthodes à noyaux.
- Quantification vectorielle.
- Réseaux de neurones et Deep Learning.
- Ensemble learning et arbres de décision.
- Les algorithmes de Bandits, optimisme face à l'incertitude.

### Travaux pratiques

*Mise en œuvre des familles d'algorithmes en utilisant des jeux de données variés.*

### 4) Les modèles et apprentissages bayésiens

- Principes d'inférence et d'apprentissage bayésiens.
- Modèles graphiques : réseaux bayésiens, champs de Markov, inférence et apprentissage.
- Méthodes bayésiennes : Naive Bayes, mélanges de gaussiennes, processus gaussiens.
- Modèles markoviens : processus markoviens, chaînes de Markov, chaînes de Markov cachées, filtrage bayésien.

### Travaux pratiques

*Mise en œuvre des familles d'algorithmes en utilisant des jeux de données variés.*

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.

### 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.
- A 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.

## 5) Machine Learning en production

- 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.

### Travaux pratiques

*Mise en production d'un modèle prédictif avec l'intégration dans des processus de batch et dans des flux de traitements.*