

Data Mining et Machine Learning

Présentation

Cette formation vous permettra d'acquérir les compétences clés du Machine Learning et ainsi être en mesure d'obtenir des informations utiles pour prendre de meilleures décisions.

L'apprentissage automatique est la science de l'utilisation des ordinateurs pour exécuter des modèles prédictifs qui apprennent à partir des données existantes afin de prévoir les comportements, les résultats et les tendances futurs. Et avec le marché mondial du Big Data en plein essor, il y a une énorme demande de scientifiques des données qualifiés.

Durée : 21,00 heures (3 jours)

Tarif INTRA : Nous consulter

Objectifs de la formation

- Comprendre la manipulation et la visualisation de données.
- Acquérir une connaissance approfondie de cette intelligence artificielle qu'est le Machine Learning
- Acquérir les connaissances du Text Mining et les principales techniques utilisées pour l'extraction et l'analyse des données de texte.

Prérequis

- Avoir connaissance de l'utilité du Data Mining en rapport avec les problématiques Big Data
- Comprendre l'intérêt de ces concepts pour l'aide à la décision et le ciblage économique

Public

- Tous les professionnels qui aspirent à faire carrière dans l'analyse commerciale en utilisant le langage R dans l'exploration de texte et l'apprentissage automatique peuvent choisir ce cours.
- Même en tant que débutant, vous pouvez vous spécialiser en apprenant le langage R avec le Text Mining et le Machine Learning avec cette formation.



Programme de la formation

Introduction au Data mining et concepts généraux

- Les défis des grands ensembles de données
- Le Data Mining parmi l'IA, l'apprentissage automatique et l'apprentissage profond.
- Comparaison pilotée par les cas d'utilisation de Data Mining, ML et DL
- Apprentissages automatisés: supervisé, non supervisé et méta-apprentissage
- Ingénierie de la décision et analyse procédurale hiérarchique
- Algorithme de Lloy Shaply et chaîne de Markov
- Echantillonnage

Data Mining avec R

- Installer R et R Studio
- Création d'objets, de tableaux et de trames de données dans R
- Utilisation de la structure, dimensions, chargement de fichiers CSV et packages
- Manipuler les données avec R
- Chargement et combinaison de vecteurs
- Retravailler les données: échanges, tris, conversions, usage des commandes sub, gsub, regexr, gregexpr, apply, lapply, sapply
- Visualisation de données, usage des commandes plot, lines, boxplot, stars, barplot, pie, hist, rug, sunflowerplot, tabplot, ggplot2, maptools.
- Validation du modèle et interprétation
- Analyse de régression

Machine Learning

- Apprentissages supervisé et non supervisé
- Régression linéaire multiple
- Algorithmes de classification
- Arbres de décision et classifieur naïf de Bayes
- K-voisins les plus proches (K-NN)
- Règles de regroupement et systèmes de recommandation
- Analyse de réseau
- Analyse de cluster: hiérarchique et K-means
- Régression logistique
- SVM Support Vector Machines (SVM)
- Les réseaux de neurones

Text Mining

- Introduction aux concepts de Text Mining
- Analyse de sentiment
- Nuage de mots positifs et négatifs
- Théorie derrière la régression avancée
- Régression avancée avec R
- Analyses d'audience Internet: théorie sous-jacente aux Web Analytics

Big Data

- Data Mining et Machine Learning en Big Data Analytics
- Gérer de gros volumes de données

Organisation**Formateur**

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes.

Moyens pédagogiques et techniques

- Apports didactiques pour apporter des connaissances communes.
- Mises en situation de réflexion sur le thème du stage et des cas concrets.
- Méthodologie d'apprentissage attractive, interactive et participative.
- Equilibre théorie / pratique : 60 % / 40 %.
- Supports de cours fournis au format papier et/ou numérique.
- Ressources documentaires en ligne et références mises à disposition par le formateur.
- Pour les formations en présentiel dans les locaux de Softeam, les stagiaires sont accueillis dans une salle de cours équipée d'un réseau Wi-Fi, d'un tableau blanc ou paperboard. Un carnet de notes est offert. Un ordinateur avec les logiciels appropriés est mis à disposition (le cas échéant).

Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation :

- Recueil des besoins : permet de récolter des informations sur le stagiaire (profil, formation, attentes particulières, ...).
- Auto-positionnement des stagiaires afin de mesurer le niveau de départ.

Tout au long de la formation :

- Évaluation continue des acquis via des questions orales, exercices / projet fil rouge, des QCM, des cas pratiques et mises en situation.

A la fin de la formation :

- Auto-positionnement des stagiaires afin de mesurer l'acquisition des compétences.
- Evaluation du formateur des compétences acquises par les stagiaires.
- Questionnaire de satisfaction à chaud : permet de connaître le ressenti des stagiaires à l'issue de la formation.
- Questionnaire de satisfaction à froid : permet d'évaluer les apports réels de la formation et leurs mises en application au quotidien.

Accessibilité

Cette formation est **accessible** aux personnes en situation de handicap, consultez-nous pour plus d'informations.