

Mis à jour le 24/09/2026

S'inscrire

Formation Altair AI Studio

3 jours (21 heures)

Présentation

Altair AI Studio, anciennement RapidMiner Studio, est une plateforme de Data Science permettant de préparer les données et de construire des modèles de Machine Learning à partir de workflows visuels, tout en limitant le recours au développement.

Cette formation Altair AI Studio vous permettra de maîtriser l'ensemble du cycle d'un projet de Data Mining selon la méthodologie CRISP-DM. Vous apprendrez à importer des données provenant de différentes sources, à les nettoyer, les transformer et les enrichir afin de construire des datasets fiables et exploitables.

Vous serez en mesure de développer des modèles de classification et de régression, mais également d'explorer vos données avec des techniques de clustering et de réduction dimensionnelle. Vous apprendrez à comparer les algorithmes et à interpréter leurs résultats afin de sélectionner les approches adaptées à vos problématiques métier.

La formation aborde également l'évaluation et l'optimisation des modèles. Vous apprendrez à utiliser la cross-validation, les principales métriques de performance et l'optimisation des hyperparamètres afin de limiter le sur-apprentissage et d'améliorer la qualité de vos prédictions.

Enfin, vous apprendrez à construire des workflows analytiques réutilisables, à intégrer AI Studio avec vos sources et systèmes externes et à étendre vos traitements avec Python. Vous découvrirez également comment intégrer des LLM et modèles génératifs à vos processus afin d'enrichir vos cas d'usage analytiques.

À l'issue de cette formation, vous serez capable de construire un workflow Data Science complet avec Altair AI Studio, depuis la préparation des données jusqu'à l'évaluation et l'intégration d'un modèle Machine Learning dans vos processus métier.

Comme toutes nos formations, celle-ci vous présentera **la dernière version stable** de la technologie et ses nouveautés.

Objectifs

- Comprendre l'environnement Altair AI Studio et la méthodologie CRISP-DM
- Importer, nettoyer, transformer et enrichir des données provenant de différentes sources
- Construire des modèles de classification, régression et clustering
- Évaluer les performances et optimiser les hyperparamètres des modèles Machine Learning
- Construire des workflows analytiques réutilisables et automatisés
- Étendre les traitements avec Python et découvrir l'intégration des LLM et modèles génératifs

Public visé

- Data Analysts
- Data Scientists
- Business Analysts
- Professionnels de la donnée souhaitant développer des modèles sans multiplier le code

Pré-requis

- Connaissances de base en analyse de données
- Notions de statistiques descriptives recommandées
- Connaissances de base en SQL pour travailler avec des sources de données relationnelles
- Aucune maîtrise avancée de Python ou d'un langage de programmation n'est nécessaire

Programme de notre formation Altair AI Studio : Data Mining et Machine Learning

[Jour 1 - Matin]

Prendre en main Altair AI Studio et les workflows Data Mining

- Comprendre l'évolution de RapidMiner Studio vers Altair AI Studio et son positionnement dans l'écosystème Data Science
- Découvrir la méthodologie CRISP-DM et les étapes d'un projet de Data Mining
- Prendre en main l'interface, les repositories, processus, opérateurs et paramètres
- Construire et exécuter un premier workflow visuel de traitement des données
- Importer des données depuis des fichiers, bases de données et autres sources disponibles
- Atelier pratique : créer un premier processus Altair AI Studio pour importer, explorer et transformer un jeu de données.

[Jour 1 - Après-midi]

Préparer, transformer et explorer les données

- Nettoyer les données et traiter les valeurs manquantes, doublons et valeurs aberrantes
- Transformer les variables avec normalisation, discrétisation, encodage, filtrage et échantillonnage
- Fusionner, joindre et agréger plusieurs sources de données dans un même workflow
- Créer de nouvelles variables et mettre en œuvre des techniques de feature engineering
- Analyser les distributions, corrélations et outliers avec les outils de visualisation et d'exploration
- Atelier pratique : préparer un dataset brut, construire de nouvelles variables et réaliser une analyse exploratoire avant modélisation.

[Jour 2 - Matin]

Construire des modèles supervisés de classification et de régression

- Comprendre les différences entre classification et régression et sélectionner une approche selon le problème métier
- Construire des modèles avec Decision Tree, Random Forest, SVM et k-NN
- Mettre en œuvre des modèles de régression et interpréter leurs prédictions
- Structurer les étapes d'entraînement, d'application du modèle et de prédiction dans un workflow
- Mesurer les performances avec accuracy, precision, recall, F1-score et métriques de régression
- Atelier pratique : entraîner plusieurs modèles supervisés, comparer leurs résultats et sélectionner le modèle le plus adapté au cas d'usage.

[Jour 2 - Après-midi]

Explorer les données avec clustering et réduction dimensionnelle

- Comprendre les principes du Machine Learning non supervisé et ses principaux cas d'usage
- Construire des clusters avec k-Means, k-Medoids et DBSCAN
- Choisir les mesures de distance et de similarité adaptées aux données analysées
- Évaluer les clusters avec l'indice de Silhouette et le Davies-Bouldin Index
- Réduire la dimension et sélectionner les variables avec PCA, pondération et techniques de sélection de caractéristiques
- Atelier pratique : segmenter un jeu de données, comparer plusieurs stratégies de clustering et interpréter les groupes obtenus.

[Jour 3 - Matin]

Évaluer et optimiser les modèles Machine Learning

- Mettre en œuvre le split training/test et la cross-validation pour évaluer les modèles
- Interpréter la matrice de confusion, les courbes ROC, l'AUC et les principales métriques de performance
- Identifier le sur-apprentissage et le sous-apprentissage et appliquer les stratégies correctives
- Comparer plusieurs algorithmes et configurations afin de sélectionner le meilleur modèle
- Optimiser les hyperparamètres et automatiser la recherche des configurations performantes
- Atelier pratique : évaluer plusieurs modèles par cross-validation puis optimiser les hyperparamètres du meilleur candidat.

Industrialiser et enrichir les workflows Altair AI Studio

- Structurer des workflows réutilisables et organiser les processus pour faciliter leur maintenance
- Exporter les résultats et intégrer les traitements avec des fichiers, bases de données, API et systèmes externes
- Étendre les capacités d'Altair AI Studio avec Python et combiner approche visuelle et développement
- Automatiser les traitements et préparer l'intégration des workflows analytiques dans les processus de l'entreprise
- Découvrir l'intégration des LLM et Generative Models dans AI Studio : prompting, modèles externes et nouveaux cas d'usage
- Atelier pratique : construire un workflow de bout en bout intégrant préparation des données, modèle ML, automatisation et enrichissement par un composant externe.

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.