

Mis à jour le 18/03/2024

S'inscrire

Formation Dataiku Avancé

Présentation

Maîtrisez l'ensemble de votre workflow data grâce à notre formation Dataiku avancé. Que vous soyez Data Engineers, Data Analysts ou Data Scientists, notre programme de perfectionnement vous permettra de concevoir des pipelines de données complexes et scalables.

Nous explorerons les pipelines de données avancés. Vous apprendrez à les créer, les optimiser et les automatiser, en vous appuyant sur des scénarios complexes et des règles précises. Cette première étape est essentielle pour quiconque souhaite se certifier en tant que concepteur avancé sur Dataiku.

Vous découvrirez comment créer, évaluer et affiner des modèles directement dans l'environnement de Dataiku. De la mise en application de vos modèles à leur compréhension profonde grâce aux outils de l'IA responsable.

Enfin, que vous souhaitiez intégrer votre propre code dans vos projets, gérer ce code de manière efficace, ou même créer des plugins pour vos collègues, cette formation vous donnera les clés pour devenir un développeur certifié sur Dataiku.

Pour cette formation, nous utiliserons la dernière version de Dataiku : [DSS 12](#).

Objectifs

- Maîtrisez la création et l'optimisation de modèles ML avec le Visual ML de Dataiku pour des solutions efficaces
- Utilisez l'IA Responsable de Dataiku pour des modèles performants, transparents et éthiques
- Construisez des pipelines de données efficaces avec Dataiku pour une analyse approfondie et une optimisation des projets
- Automatisez les pipelines de données et intégrez du code dans Dataiku pour une personnalisation et adaptabilité avancées
- Déployez et surveillez vos projets en production avec Dataiku pour une transition fluide et un succès à long terme

Public visé

- Data analysts
- Data scientists
- Data engineers
- Analyst SAS

Pré-requis

- Vous avez une compréhension fondamentale de l'environnement et des fonctionnalités de base de Dataiku, idéalement acquise par une expérience pratique ou notre formation initiale sur Dataiku
- Vous êtes familier avec les concepts de data science et d'analyse de données, et vous avez une compréhension des principes de base du machine learning
- Vous avez déjà codé au moins en Python et/ou R, essentielle pour tirer parti des scripts et des notebooks dans Dataiku, personnaliser les analyses, et développer des modèles
- Avoir suivi notre [formation Dataiku](#) est préférable

Pré-requis logiciel

Avoir une licence pour Dataiku.

Programme de notre formation Dataiku Avancé

Jour 1

Introduction

- Présentation des fonctionnalités avancées de Dataiku
- Vue d'ensemble des possibilités offertes par l'outil

Outils visuels avancés de Dataiku

- Recettes visuelles avancées
- Utiliser des variables
- Data Pipelines
- Automation
- Collaboration
- Quiz
- Récapitulatif

Introduction Machine Learning

- Concepts
- Modélisation prédictive

- Prédiction : Régression et classification
- Clustering
- Quiz
- Récapitulatif

Analysez vos données avec l'outil Interactive Stats

- Interface
- Analyse univariée et bivariée
- Ajustement de courbes et distributions
- Matrice de corrélation
- Analyse en composantes principales
- Tests statistiques
- Quiz
- Récapitulatif

Machine Learning

- Mettre en place d'un modèle
- Ajustement du modèle
- Compréhension des prédictions faites par le modèle
- Quiz
- Récapitulatif

Scoring

- Déploiement de votre modèle
- Scoring Data
- Evaluation de votre modèle
- Quiz
- Récapitulatif

Machine Learning avancé

- Stratégies de développement de modèles
- Diagnostic du modèle
- Actionability
- Quiz
- Récapitulatif

Modèles stratifiés ou partitionnés

- Mise en place d'un modèle stratifié
- Quiz
- Récapitulative

TALN - De manière visuelle

- Introduction au TALN
- Préparation des données textuelles
- Gestion des caractéristiques textuelles pour le ML
- Quiz
- Récapitulatif

Analyse et prévision des séries temporelles

- Fondamentaux des séries temporelles
- Analyse des séries temporelles
- Prévision des séries temporelles
- Quiz
- Récapitulatif

Préparation de données temporelles

- Recette Resampling
- Recette Interval extraction
- Recette Windowing
- Recette Externa extraction
- Quiz
- Récapitulatif

Jour 2

Code dans Dataiku

- Notebooks de code
- Recettes de code
- Environnements de code
- Intégrations d'IDE externes
- Quiz
- Récapitulatif

Partagez votre code

- Introduction
- Libraries
- Importation depuis Git
- Échantillons de code
- Pratique
- Quiz
- Récapitulatif

Modèles ML personnalisés

- Prétraitement personnalisé en ML visuel
- Modélisation personnalisée en ML visuel
- Quiz
- Récapitulative

Variables pour les coders

- Définition des variables
- Utilisation des variables dans une recette de code
- Modification de la valeur des variables
- Pratique
- Quiz
- Récapitulatif

Visualisation

- Webapps
- Insights statiques
- Quiz
- Récapitulatif

Managed folders

- Savoir utiliser les Managed folders
- Pratique
- Quiz
- Récapitulatif

APIs de Dataiku

- Introduction
- Le package dataiku et l'API publique
- Pratique
- Quiz
- Récapitulatif

Déploiement et monitoring

- Dataiku Govern
- Suivi des modèles
- Déploiement par lots
- API en temps réel
- Quiz
- Récapitulatif

Examen Advanced Designer Certificate (Optionnel)

Examen ML Practitioner Certificate (Optionnel)

Examen Developer Certificate (Optionnel)

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.