

Dataiku : prise en main de DS Studio

L'objectif de cette formation est de familiariser les participants avec Dataiku et son interface DS Studio. Les stagiaires apprendront à manipuler des données, à créer des modèles de machine learning et à visualiser les résultats. À l'issue de la formation, ils seront en mesure d'utiliser Dataiku pour des projets d'analyse de données.

Pré-requis : Bases en analyse de données, connaître les concepts de machine learning, compétences en informatique.



Modalité :

- Distanciel en classe virtuelle
- E-learning : à venir
- Présentiel

Communauté:

community.reconvert.net

Durée totale :

21 H (3 jours)

PLAN DETAILLE

Introduction à Dataiku

Présentation de Dataiku
 Comparaison avec d'autres outils d'analyse de données
 Installation et configuration de Dataiku
 Navigation dans l'interface de DS Studio
 Création d'un projet dans DS Studio
 Gestion des utilisateurs et des permissions
 Bonnes pratiques pour démarrer un projet
Travaux pratiques : création d'un projet simple dans DS Studio

Importation et préparation des données

Types de sources de données compatibles
 Connexion à une base de données
 Importation de fichiers CSV et Excel
 Nettoyage et transformation des données
 Utilisation des recettes de préparation
 Gestion des valeurs manquantes et des doublons
 Exploration des données avec des visualisations
Travaux pratiques : importer et nettoyer un jeu de données

Exploration et visualisation des données

Introduction aux graphiques et visualisations
 Création de visualisations interactives
 Analyse descriptive des données
 Utilisation de tableaux de bord dans Dataiku
 Personnalisation des graphiques
 Exportation des visualisations
 Partage des résultats avec l'équipe
Travaux pratiques : création d'un tableau de bord interactif

Optimisation des modèles

Importance de l'optimisation des hyperparamètres
 Techniques pour améliorer la performance du modèle
 Utilisation de la validation croisée
 Comparaison de plusieurs modèles
 Interprétation des métriques de performance
 Stratégies pour éviter le surapprentissage
 Déploiement des modèles optimisés
Travaux pratiques : optimisation d'un modèle existant

Introduction au machine learning

Concepts fondamentaux du machine learning
 Types de modèles de machine learning disponibles
 Sélection des variables pertinentes
 Création d'un modèle prédictif
 Validation du modèle et évaluation des performances
 Interprétation des résultats du modèle
 Introduction aux pipelines de machine learning
Travaux pratiques : créer et évaluer un modèle simple

Dataiku : prise en main de DS Studio

L'objectif de cette formation est de familiariser les participants avec Dataiku et son interface DS Studio. Les stagiaires apprendront à manipuler des données, à créer des modèles de machine learning et à visualiser les résultats. À l'issue de la formation, ils seront en mesure d'utiliser Dataiku pour des projets d'analyse de données.

Pré-requis : Bases en analyse de données, connaître les concepts de machine learning, compétences en informatique.



- Modalité :**
- Distanciel en classe virtuelle
 - E-learning : à venir
 - Présentiel

Communauté: community.reconvert.net

Durée totale : 21 H (3 jours)

PLAN DETAILLE

Automatisation des workflows

Introduction aux workflows dans Dataiku
Création d'un workflow automatisé
Planification des tâches récurrentes
Surveillance des exécutions de workflows
Gestion des erreurs et des alertes
Utilisation de l'API Dataiku
Intégration avec d'autres outils et systèmes
Travaux pratiques : automatiser un workflow simple

Collaboration et partage des résultats

Meilleures pratiques pour le travail en équipe
Partage de projets et de recettes
Utilisation des commentaires et des annotations
Exportation des résultats vers des formats variés
Gestion des versions des projets
Introduction à la documentation dans Dataiku
Cas d'utilisation et témoignages
Travaux pratiques : collaboration