



Google Cloud Big Data and Machine Learning Fundamentals

Durée 1 jour(s) (BIGDATA-GOOGLE)

T-GCPBDML-B-04

Formation officielle



Description

Ce cours vous présentera les produits de Big Data et d'apprentissage automatique de Google Cloud. Vous commencerez par un aperçu rapide de Google Cloud, puis vous apprendrez ses capacités de traitement des données offertes par la plateforme.

Objectifs

- Identifier le cycle de vie des données vers l'IA sur Google Cloud et les principaux produits du Big Data et du Machine Learning.
- Concevoir des pipelines de streaming avec Dataflow et Pub/Sub.
- Analyser le Big Data à grande échelle avec BigQuery.
- Identifier différentes options pour créer des solutions de machine learning sur Google Cloud.
- Décrire un workflow de machine learning et les étapes clés avec Vertex AI.
- Créer un pipeline d'apprentissage automatique à l'aide d'AutoML.

Public

- Data analysts, data scientists, et business analysts débutants avec Google Cloud.
- Personnes chargées de la conception des pipelines et des architectures des traitements des données, de la création et de la maintenance des modèles de ML et des modèles statistiques, de l'interrogation des ensembles de données, de la visualisation des résultats des requêtes et de la création des rapports.
- Dirigeants et décideurs informatiques en charge de l'évaluation de l'utilisation de Google Cloud par les data scientists

Prérequis

Connaissance de base d'un ou plusieurs des éléments suivants :

- Langage de requête de base de données tel que SQL
- Flux de travail d'ingénierie des données depuis l'extraction, la transformation, le chargement jusqu'à l'analyse, la modélisation et le déploiement
- Modèles d'apprentissage automatique tels que les modèles supervisés et non supervisés

Répartition

À définir

Evaluations des acquis

L'évaluation des acquis de la formation se fera en séance au travers d'ateliers, d'exercices et/ou de travaux pratiques. Dans le cas d'une formation officielle éditeur, veuillez nous consulter afin que nous vous fassions part des modalités d'évaluation.

A l'issue de la formation, vous sera transmis une évaluation à chaud de l'action de formation qui vous permettra de nous faire part de vos retours quant à votre expérience apprenant avec Zenika.

Ressources pédagogiques

Les ressources pédagogiques proviennent de productions des équipes Zenika et/ou de la documentation éditeur dans le cas d'une formation "Officielle". Les documents sont en français ou en anglais.

RQTH et ma formation Zenika

Si vous êtes sujet à un handicap, prenez contact avec nos équipes pour que nous puissions définir ensemble comment nous pourrions aménager la session afin que vous puissiez vivre une expérience en formation inchangée.

Programme

Introduction

- Big Data et Machine Learning dans Google Cloud
- Introduction
- Compute
- Stockage
- Historique des produits Big Data et ML
- Catégories de produits Big Data et ML
- Exemple client : Gojek
- Pluralsight : Démarrer avec GCP et Qwiklabs
- TP : Explorer un Dataset publiques BigQuery

Ingénierie des données pour le streaming de données

- Introduction
- Les défis du Big Data
- Architecture orientée messages
- Conception de pipelines de streaming avec Apache Beam
- Implémentation de pipelines de streaming sur Cloud Dataflow
- Visualisation avec Looker
- TP: Création d'un pipeline de données en streaming pour un tableau de bord en temps réel avec Dataflow

Big Data avec BigQuery

- Introduction
- Stockage et analyses
- Démo BigQuery – Vélos en libre-service à San Francisco
- Introduction à BigQuery ML
- Utiliser BigQuery ML pour prédire "Customer lifetime value"
- Commandes clé de BigQuery ML
- TP: Prédire les achats des visiteurs à l'aide de BigQuery ML

Options de machine learning sur Google Cloud

- Introduction
- Options pour créer des modèles ML
- API prédéfinies
- ML automatique
- Custom training
- Vertex AI
- Solutions d'IA

Le workflow de machine learning avec Vertex AI

- Introduction
- Préparation des données
- Model training
- Model evaluation
- Déploiement et monitoring du modèle
- TP: Prédire le risque de prêt avec AutoML

Résumé