

Machine Learning on Google Cloud

Bien démarrer vos projets machine learning avec Google Cloud

5 jour(s) / 35h

Objectifs pédagogiques

- Créer, entraîner et déployer un modèle de Machine Learning sans écrire une seule ligne de code à l'aide de Vertex AI AutoML.
- Comprendre quand utiliser AutoML et Big Query ML.
- Créer des ensembles de données gérés par Vertex AI.
- Effectuer une analyse exploratoire des données.
- Décrire comment améliorer la qualité des données.
- Créer des ensembles de données d'entraînement, d'évaluation et de test reproductibles et évolutifs.
- Créer et utiliser un notebook Vertex AI Workbench
- Construire et entraîner des modèles d'apprentissage supervisé.
- Mettre en œuvre des modèles de ML à l'aide de TensorFlow/Keras.
- Créer une tâche d'entraînement personnalisée, et la déployer à l'aide d'un conteneur Docker.
- Optimiser et évaluer les modèles à l'aide des fonctions de perte et des indicateurs de performance.
- Comprendre les avantages de l'utilisation du feature engineering.
- Décrire comment représenter et transformer des entités.
- Ajouter des features à un Feature Store.
- Décrire le réglage des hyperparamètres à l'aide de Vertex Vizier et comment il peut être utilisé pour améliorer les performances du modèle.
- Expliquer Vertex AI Pipelines
- Décrire Analytics Hub, Dataplex et Data Catalog.
- Décrire les prédictions par lots et en ligne et la surveillance des modèles.

Public cible

- Aspirants machine learning data analysts, data scientists et data engineers
- Les participants qui souhaitent être exposés au ML à l'aide de Vertex AI AutoML, BQML, Feature Store, Workbench, Dataflow, Vizier pour le réglage des hyperparamètres et TensorFlow/Keras

Prérequis

Pour tirer le meilleur parti de ce cours, les participants doivent avoir:

- Une certaine familiarité avec les concepts de base du machine learning
- Des compétences de base avec un langage de script, de préférence Python

Programme

Module 1: Comment Google fait du Machine Learning

Sujets

- Quelles sont les bonnes pratiques pour mettre en œuvre le machine learning sur Google Cloud ?
- Qu'est-ce que Vertex AI et comment pouvez-vous utiliser la plate-forme pour créer, entraîner et déployer rapidement des modèles de machine learning AutoML sans écrire une seule ligne de code ?
- Qu'est-ce que le Machine Learning et quels types de problèmes peut-il résoudre ?

Google envisage le Machine Learning de manière légèrement différente : il s'agit de fournir une plate-forme unifiée pour les ensembles de données gérés, un features store, un moyen de créer, d'entraîner et de déployer des modèles de Machine Learning sans écrire une seule ligne de code, offrant la possibilité d'étiqueter les données, de créer notebook Workbench utilisant des frameworks tels que TensorFlow, SciKit Learn, Pytorch, R et autres.

Notre plate-forme Vertex AI inclut également la possibilité d'entraîner des modèles personnalisés, de créer des pipelines de composants et d'effectuer des prédictions en ligne et par lots.

Nous discutons également des cinq phases de la conversion d'un cas d'utilisation candidat pour qu'il soit piloté par le Machine Learning, et examinons pourquoi il est important de ne pas sauter les phases. Nous terminons par une reconnaissance des biais que le Machine Learning peut amplifier et comment les reconnaître.

Objectifs

- Décrire Vertex AI Platform et son utilisation pour créer, entraîner et déployer rapidement des modèles de Machine Learning AutoML sans écrire une seule ligne de code.
- Décrire les bonnes pratiques pour mettre en œuvre le machine learning sur Google Cloud.
- Développer une stratégie de données autour du Machine Learning.
- Examiner les cas d'utilisation qui sont ensuite repensés à travers une lentille ML.
- Tirer parti des outils et de l'environnement Google Cloud Platform pour faire du ML.
- Tirer parti de l'expérience de Google pour éviter les pièges courants.
- Effectuer des tâches de science des données dans des notebooks collaboratifs en ligne.

Activités

- Labs pratiques
- Quiz des modules
- Lectures

Module 2: Se lancer dans le Machine Learning

Sujets

- Le cours commence par une discussion sur les données : comment améliorer la qualité des données et effectuer une analyse exploratoire des données.
- Nous décrivons Vertex AI AutoML et comment créer, former et déployer un modèle ML sans écrire une seule ligne de code. Vous comprendrez les avantages de BigQuery ML.
- Nous discutons ensuite de la manière d'optimiser un modèle de Machine Learning (ML) et de la manière dont la généralisation et l'échantillonnage peuvent aider à évaluer la qualité des modèles ML pour un entraînement personnalisé.

Objectifs

- Décrire Vertex AI AutoML et comment créer, entraîner et déployer un modèle ML sans écrire une seule ligne de code.

- Décrire BigQuery ML et ses avantages.
- Décrire comment améliorer la qualité des données.
- Effectuer une analyse exploratoire des données.
- Construire et entraîner des modèles d'apprentissage supervisé.
- Optimiser et évaluer les modèles à l'aide de fonctions de perte et de mesures de performance.
- Atténuer les problèmes courants qui surviennent dans le Machine Learning.
- Créer des ensembles de données de formation, d'évaluation et de test reproductibles et évolutifs

Activités

- Labs pratiques
- Quiz des modules
- Lectures

Module 3: TensorFlow dans Google Cloud

Sujets

Les modules couvrent:

- la conception et la création d'un pipeline de données d'entrée TensorFlow
- la création de modèles ML avec TensorFlow et Keras
- l'amélioration de la précision des modèles ML, l'écriture de modèles ML pour une utilisation à l'échelle et l'écriture de modèles ML spécialisés.

Objectifs

- Créer des modèles de Machine Learning TensorFlow et Keras.
- Décrire les composants clés de TensorFlow.
- Utiliser la bibliothèque tf.data pour manipuler des données et de grands ensembles de données.
- Créer un modèle ML à l'aide des couches de prétraitement tf.keras.
- Utiliser les API séquentielles et fonctionnelles Keras pour la création de modèles simples et avancés.
- Comprendre comment le sous-classement de modèles peut être utilisé pour des modèles plus personnalisés.
- Utiliser les utilitaires tf.keras.preprocessing pour travailler avec des données d'image, des données de texte et des données de séquence.
- Entraîner, déployer et produire des modèles de ML à grande échelle avec Cloud AI Platform

Activités

- Labs pratiques
- Quiz des modules
- Lectures

Module 4: Feature Engineering

Sujets

- Vous voulez en savoir plus sur Vertex AI Feature Store ?
- Vous voulez savoir comment vous pouvez améliorer la précision de vos modèles de ML ?
- Qu'en est-il de la recherche des colonnes de données qui présentent les fonctionnalités les plus utiles ?

Bienvenue dans Feature Engineering, où nous discutons des bonnes et des mauvaises features et comment vous pouvez les prétraiter et les transformer pour une utilisation optimale dans vos modèles.

Ce cours comprend du contenu et des ateliers sur le Feature Engineering à l'aide de BigQuery ML, Keras et TensorFlow

Objectifs

- Décrire le feature store de Vertex AI.
- Comparer les principaux aspects requis d'une bonne feature.
- Combiner et créer de nouvelles combinaisons de fonctionnalités grâce à des feature crosses.
- Effectuer le feature engineering à l'aide de BigQuery ML, Keras et TensorFlow.
- Comprendre comment prétraiter et explorer les features avec Dataflow et Dataprep by Trifacta.
- Comprendre et appliquer comment TensorFlow transforme les features.

Activités

- Labs pratiques
- Quiz des modules
- Lectures

Module 5: Machine Learning en Entreprise

Sujets

Ce cours englobe une approche pratique du monde réel du flux de travail ML : une approche d'étude de cas qui présente une équipe ML confrontée à plusieurs exigences commerciales et cas d'utilisation ML. Cette équipe doit comprendre les outils requis pour la gestion et la gouvernance des données et envisager la meilleure approche pour le prétraitement des données : de la présentation d'un aperçu de Dataflow et de Dataprep à l'utilisation de BigQuery pour les tâches de prétraitement.

L'équipe se voit présenter trois options pour créer des modèles de Machine Learning pour deux cas d'utilisation spécifiques. Ce cours explique pourquoi l'équipe utiliserait AutoML, BigQuery ML ou un entraînement personnalisé pour atteindre ses objectifs.

Une plongée plus profonde dans l'entraînement personnalisé est présentée dans ce cours. Nous décrivons les exigences liées à l'entraînement personnalisé, de la structure du code d'entraînement, du stockage et du chargement de grands ensembles de données à l'exportation d'un modèle entraîné.

Vous allez créer un modèle de machine learning personnalisé, qui vous permet de créer une image de conteneur avec peu de connaissances sur Docker.

L'équipe de l'étude de cas examine le réglage des hyperparamètres à l'aide de Vertex Vizier et comment il peut être utilisé pour améliorer les performances du modèle. Pour mieux comprendre l'amélioration des modèles, nous plongeons dans un peu de théorie : nous discutons de la régularisation, de la gestion de la parcimonie et de nombreux autres concepts et principes essentiels.

Nous terminons par un aperçu de la prédiction et de la surveillance des modèles et de la manière dont Vertex AI peut être utilisé pour gérer les modèles ML.

Objectifs

- Comprendre les outils requis pour la gestion et la gouvernance des données.
- Décrire la meilleure approche pour le prétraitement des données : de la présentation de Dataflow et de Dataprep à l'utilisation de SQL pour les tâches de prétraitement.
- Expliquer en quoi AutoML, BigQuery ML et l'entraînement personnalisé diffèrent et quand utiliser un framework particulier.
- Décrire le réglage des hyperparamètres à l'aide de Vertex Vizier et comment il peut être utilisé pour améliorer les performances du modèle.
- Expliquer la prédiction et la surveillance des modèles et comment Vertex AI peut être utilisé pour gérer les modèles ML.
- Décrire les avantages de Vertex AI Pipelines.

Activités

- Labs pratiques
- Quiz des modules
- Lectures

Les concepts suivants ne sont pas couverts dans ce cours:

- Concepts et fondamentaux du cloud
- Mise en réseau
- Sécurité