

Databricks

CODICE	DT0286
DURATA	2 gg
PREZZO	1.190,00 €
EXAM	

DESCRIZIONE

Il corso è strutturato in modo tale da fornire una breve panoramica del cloud di azure in modo tale da poter effettuare le corrette configurazioni propedeutiche all'utilizzo, l'installazione e la configurazione di Databricks, una piattaforma che offre strumenti che consentono di connettere le origini dati a un'unica piattaforma per elaborare, archiviare, condividere, analizzare, modellare e monetizzare i set di dati con soluzioni dalla BI all'intelligenza artificiale generativa.

TARGET

Professionisti IT, data scientist, data engineer e analisti

PREREQUISITI

Conoscenza di base di programmazione, fondamentali di cloud computing

CONTENUTI

Giorno 1

- Il Cloud Computing ed Azure
- Architettura di Azure ed i servizi principali
- Azure ed i Data Services principali
- Gestione dei gruppi di risorse in Azure
- Gli Storage Accounts in Azure
- Il Database Sql Server sul Cloud di Microsoft
- Introduzione a DataBricks
- Architettura DataBricks
- Configurazione dell'ambiente
- Installare e configurare Databricks
- La UI di DataBricks e come navigare
- Cosa sono i Workspace di DataBricks

- Come creare un Workspace DataBricks
- Data Management in DataBricks
- Apache Spark al centro della piattaforma Azure DataBricks
- Spark per alimentare i Cluster di calcolo e Sql Warehouse
- Qual è la relazione tra Apache Spark e Azure DataBricks ?
- Come funziona Apache Spark in Azure DataBricks
- Perché utilizzare Apache Spark in Azure DataBricks

Giorno 2

- Configurare le proprietà di Spark in DataBricks SQL
- Collegare una fonte di dati Apache Spark
- Comprensione delle colonne di base e dei tipi di dati
- Gestione del File System nei Notebook
- Gestione dei Jobs e dei Cluster
- Creazione e configurazione di Cluster
- Creazione di Job utilizzando il Notebook
- Come eseguire un Job
- Visualizzazione dei Job e dei dettagli
- Introduzione a Delta Lake
- Delta Lake, una evoluzione del Data Lake tra Big Data e Spark
- Utilizzo di Delta Lake in Databricks
- Caricamento dei dati in Delta Lake
- Gestione dei dati in Delta Lake
- Panoramica sulla sicurezza in Databricks
- Gestione della sicurezza di Databricks
- Gestione del Backup e del Recovery
- Risoluzione dei problemi
- Troubleshooting