Databricks

CODICE	**DT0286**
DURATA	**2 gg**
PREZZO	**890,00 €**
EXAM	

DESCRIZIONE

Il corso Databricks fornisce una visione completa e operativa della piattaforma Azure Databricks, oggi uno degli standard di riferimento per la gestione e l'analisi dei dati su larga scala.

Partendo dai fondamenti del cloud Azure, il percorso guida i partecipanti nell'utilizzo della Lakehouse Platform, che consente di unificare data engineering, analytics e machine learning in un unico ambiente.

Attraverso un approccio pratico, i partecipanti imparano a:

- configurare e utilizzare ambienti Databricks
- integrare fonti dati eterogenee
- progettare pipeline di elaborazione scalabili con Apache Spark
- gestire e ottimizzare dataset tramite Delta Lake
- orchestrare job e workload in ambienti distribuiti

Il corso consente di comprendere come trasformare dati grezzi in informazioni strategiche e utilizzabili, supportando dall'analisi descrittiva fino a scenari avanzati di AI e data-driven decision making.

Al termine, i partecipanti saranno in grado di operare in autonomia su ambienti Databricks e contribuire allo sviluppo di soluzioni dati moderne, scalabili e pronte per contesti enterprise.

TARGET

Professionisti IT, data scientist, data engineer e analisti

PREREQUISTI

Conoscenza di base di programmazione, fondamenti 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