Szkolenie: Microsoft

DP-3011: Implement a Data Analytics Solution with Azure Databricks

Microsoft Partner

Cel szkolenia:

Ten kurs bada, jak używać środowiska Databricks i Apache Spark na platformie Azure, aby przeprowadzić projekty danych od etapu eksploracji do produkcji. Dowiesz się, jak pozyskiwać, przekształcać i analizować zestawy danych o dużej skali za pomocą Spark DataFrames, Spark SQL i PySpark, jednocześnie budując pewność siebie w zarządzaniu rozproszonym przetwarzaniem danych. Zagłębisz się również w praktyki inżynierii danych, w tym projektowanie potoków ETL, obsługę ewolucji schematów i egzekwowanie jakości danych. Następnie kurs przechodzi do orkiestracji, pokazując, jak automatyzować i zarządzać obciążeniami za pomocą Lakeflow Jobs i potoków. Na koniec poznasz możliwości zarządzania i bezpieczeństwa, takie jak Unity Catalog i integracja z usługą Purview.

Profil odbiorców:
Przed rozpoczęciem tego kursu uczestnicy powinni już swobodnie posługiwać się podstawami języków Python i SQL. Ważna jest znajomość witryny Azure portal i podstawowych usług, takich jak Azure Storage, wraz z ogólną świadomością koncepcji danych, takich jak przetwarzanie wsadowe w porównaniu do przetwarzania strumieniowego oraz dane ustrukturyzowane w porównaniu do nieustrukturyzowanych.

Plan szkolenia:

- Wdrażanie rozwiązań analizy danych przy użyciu usługi Azure Databricks
  - Moduł 1: Eksploracja usługi Azure Databricks
  - Moduł 2: Przeprowadzanie analizy danych w usłudze Azure Databricks
  - Moduł 3: Używanie platformy Apache Spark w usłudze Azure Databricks
  - Moduł 4: Zarządzanie danymi przy użyciu usługi Delta Lake
  - Moduł 5: Tworzenie deklaratywnych potoków Lakeflow
  - Moduł 6: Wdrażanie obciążeń za pomocą Lakeflow Jobs

Wymagania:

Biegłość w podstawach języków Python i SQL. Umiejętność pisania prostych skryptów w języku

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001

Python, pracy z popularnymi strukturami danych oraz pisania zapytań SQL w celu filtrowania, łączenia i agregowania danych. Podstawowe zrozumienie popularnych formatów plików (CSV, JSON, Parquet). Znajomość witryny Azure portal i podstawowych usług, takich jak Azure Storage. Świadomość koncepcji danych (przetwarzanie wsadowe vs strumieniowe, dane ustrukturyzowane vs nieustrukturyzowane).

Poziom trudności

Certyfikaty:

Uczestnicy otrzymają certyfikaty podpisane przez Microsoft (ukończenie kursu).

Prowadzący:

Microsoft Certified Trainer (MCT)

Compendium Centrum Edukacyjne Spółka z o.o.
ul. Tatarska 5, 30-103 Kraków, tel.: (12) 298 47 77
ul. Postępu 18B, 02-676 Warszawa, tel.: (22) 417 41 70
BNP Paribas Bank Polska S.A.
Oddział w Warszawie
Nr konta: PL 20 1600 1462 1853 6451 3000 0001