(Big) Data Engineering In Depth

Moustafa Mahmoud
Data Solution Architect

Chapter: Apache Spark

Course Objectives

- Understand the fundamentals of Apache Spark.

- Build data processing applications using Apache Spark.

- Understand Apache Spark APIs (RDD, DataFrames, and Datasets).

- Optimize and tune Apache Spark applications.

- Build streaming applications using Apache Spark.

- Build scalable machine learning applications using Apache Spark MLlib.

- Deploy Apache Spark in production environments.

Section: Course References

Course References

- [Learning Spark, 2nd Edition](#) By Jules S. Damji, Brooke Wenig, Tathagata Das, Denny Lee.

- [Spark: The Definitive Guide](#) By Bill Chambers, Matei Zaharia.

- Random blog posts.

Section: Course Prerequisites

Course Prerequisites

Before beginning this course, participants should have:

- Experience in programming using Python.

- Basic programming skills using shell.

- An understanding of MapReduce foundations and Hive. [Garage Education YouTube Playlist](#)

Thank you for watching!

See you in the next video ☺