SparkWorkshop

Spark worshop

Spark worshop

SparkWorkshop - Version: 1

🕐 3 days Course

Description:

In the first training day, the course will introduce developers to basic Spark concepts (RDD, Transformations, Actions). In the next two training days, the course will cover Spark including: Spark Streaming, and Cluster deployment.

Intended audience:

Developers who wish to obtain the skills and knowledge to develop big data solutions with Spark.

Prerequisites:

Basic knowledge of programming in Python

Objectives:

Storing and analyzing data
Distributed programming with Spark

Topics:

Day 1 – Sparks Basics

• Spark's architecture.

- Python fundamentals
- Basic Spark concepts: RDD,Transformations,Actions
-

Day 2 – Data in Spark

- Loading and storing data.
- Spark SQL
- Spark Streaming

Day 3 – Clustering and Deployment

- Cluster deployment.
- Persist reusable results
- Controlling memory usage