

Spark for Developers Eğitimi

Açıklama

Apache Spark dağıtık bilgi işlem motoru, büyük ölçekli veri kümelerinin işlenmesi ve analizi konusunda hızlı bir şekilde birincil araç haline geliyor. Hadoop gibi mevcut motorlara göre 10 ila 100 kat daha hızlı çalışma süresi hızları ve daha basit programlama modeli dahil olmak üzere birçok avantaja sahiptir. Bu eğitimi tamamladıktan sonra, Spark ile daha bilinçli ve üretken bir şekilde çalışmaya hazır olacaksınız.

Bu Eğitimde Neler Öğreneceksiniz ?

- Katılımcılar, Spark'ın Büyük Veri ekosistemine nasıl uyum sağladığını ve veri analizi için Spark'ın nasıl kullanılacağını öğrenirler. Eğitim, Spark'la etkileşimli veri analizi, Spark iç işlevleri, Spark API'leri, Spark SQL, Spark Streaming, makine öğrenimi gibi konu başlıklarını içerir.

Eğitim İçeriği

Spark Basics

- Spark and Hadoop
- Spark Concepts and Architecture
- Spark eco System (core, spark sql, ml, streaming)
- Spark SQL
- RDD

Spark API programming

- Introduction to Spark API / RDD API
- Submitting the First Program to Spark
- Debugging / Logging
- Configuration Properties

Machine Learning Fundamentals

- ML and DL fundamentals
- Regression
- Classification
- Clustering

Machine Learning on bigdata with SparkML

- Feature operations
- Preparing data for ML
- Onehotencoding, scaling..
- Training Models
- classification, regression
- Hyper parameter tuning
- Cross validation, Train Validation Split
- Basic Sentiment analysis on text data

Introduction to Structured Streaming

- Apache Spark Streaming Overview
- Creating Streaming DataFrames
- Transforming DataFrames
- Executing Streaming Queries

Structured Streaming with Apache Kafka

- Receiving Kafka Messages
- Sending Kafka Messages

Aggregating and Joining Streaming DataFrames

- Streaming Aggregation
- Joining Streaming DataFrames

Spark Distributed Processing

- Apache Spark on a Cluster
- RDD Partitions
- Stages and Tasks
- Job Execution Planning

Ön Koşullar

Temel programlama bilgisine sahip olmak