BIG DATA ANALYTICS

Functional Programming With Scala – Overview

Scala Immutable and Mutuable Variable Declaration

Values (concept of val and var)

Scala Functions

Create normal functions or Anonymous Functions , Applying Functions
Partially Curried Functions , Var-args Support in Functions

Scala Classes, Traits

Classes in Scala , Constructors , Expressions , Functions vs Methods
Inheritance, Overloading Methods, Abstract Class, Traits, Inheritance in Traits
When to use Trait instead of Abstract Class
Using Generics/Types with Traits

Exception Handling

try-catch-finally. Expression-oriented programming

Case Classes & Collections

Case classes, objects, packages, apply, update, Functions are Objects (uniform access principle), pattern matching
Collections : Lists, Maps, functional combinators (map, foreach, filter, zip, folds)
Pattern matching & functional composition

BigData Overview and Apache Hadoop

What is Big Data & Hadoop ?
Hadoop Architecture : Yarn , MapReduce and HDFS Overview
Use Cases of Hadoop

Apache Spark Overview
Spark Core

Introduction to Apache Spark and its Core Architecture
What are Spark RDD and how they internally works?
Exploring various functions of Spark RDD API's (map, flatMap, filter, reduce , reduceByKey etc)

Spark SQL, DataFrames and DataSets
The DataFrames/Datasets API , Data Aggregation , Column Operations,
The Functions API: date/time, string manipulation, aggregation
Caching and caching storage levels

Structured Spark Streaming
Need of Real-time Stream Processing?
Sources and sinks
Structured Streaming APIs
Windowing & Aggregation

Machine Learning with Apache Spark MLlib
Build machine learning pipelines for both supervised and unsupervised learning
Transformer/Estimator/Pipeline API

Graph Processing with Apache Spark
Basic Graph Analysis
GraphFrames API
GraphFrames ETL

Project : Real-time Sentiment Analysis on Real-time Twitter Data