A)Scala:

- Why Scala?
- What is Scala?
- Introducing Scala
- Installing Scala
- Scala Basics
- Scala Basic Types
- Defining Functions
- If statements
- Scala For Comprehensions
- While Loops
- Do-While Loops
- Conditional Operators
- Base Object Oriented Programming in Scala
- Case Objects and Classes

B)Spark

- Introduction to Big Data
- Challenges with Big Data
- Batch Vs. Real Time Big Data Analytics
- Batch Analytics – Hadoop Ecosystem Overview
- Spark Opportunity and Solution
- In Memory Data – Spark

- What is Spark?
- Modes of Spark
- Spark Installation
- Spark Standalone Cluster
- Capabilities and Ecosystem
- Spark Components vs Hadoop
- Loading a File in spark Shell
- Performing Some Basic Operations on Files in Spark Shell

C)RDD Fundamentals

- Purpose and Structure of RDDs
- Transformations, Actions, and DAG
- RDD programming API

D)Spark SQL / Dataframes

- Spark SQL and DataFrame Uses
- DataFrame / SQL APIs
- Catalyst Query Optimization

E)Spark Job Execution

- Jobs, Stages, and Tasks
- Partitions and Shuffles
- Data Locality
- Job Performance

F)Spark Streaming

•Streaming Sources and Tasks

• DStream APIs and Stateful Streams

• Reliability and Fault Recovery

G)Spark Mlib

• Classification Algorithm

• Clustering Algorithm

• Sequence Mining Algorithm

• Collbrative filtering

H)Spark GraphX

•Graph analysis with Spark

•GraphX for graphs

•Graph-parallel computation