

Big data Hadoop Curriculum

Module 1

- Introduction to Big Data & Big Data Challenges
- Limitations & Solutions of Big Data Architecture
- Hadoop & its Features
- Hadoop Ecosystem
- Hadoop 2.x Core Components
- Hadoop Storage: HDFS (Hadoop Distributed File System)
- Hadoop Processing: MapReduce Framework
- Different Hadoop Distributions

Module 2

- Hadoop 2.x Cluster Architecture
- Federation and High Availability Architecture
- Typical Production Hadoop Cluster
- Hadoop Cluster Modes
- Common Hadoop Shell Commands
- Hadoop 2.x Configuration Files
- Single Node Cluster & Multi-Node Cluster set up
- Basic Hadoop Administration

Module 3

- YARN Workflow
- Anatomy of MapReduce Program
- Input Splits, Relation between Input Splits and HDFS Blocks
- MapReduce: Combiner & Partitioner
- Demo of Health Care Dataset
- Demo of Weather Dataset



+44 (020) 3950 8881



+1 332-600-8899



+91 9581512555



+46 76 441 48 79



trainings@slearn.academy

Module 4

- MapReduce vs Pig
- Pig Components & Pig Execution
- Pig Data Types & Data Models in Pig
- Pig Latin Programs
- Shell and Utility Commands
- Pig UDF & Pig Streaming
- Testing Pig scripts with Punit
- Aviation use-case in PIG
- Pig Demo of Healthcare Dataset



Module 5

- Introduction to Apache Hive
- Hive vs Pig
- Hive Architecture and Components
- Hive Metastore
- Limitations of Hive
- Comparison with Traditional Database
- Hive Data Types and Data Models
- Hive Partition
- Hive Bucketing
- Hive Tables (Managed Tables and External Tables)
- Importing Data
- Querying Data & Managing Outputs
- Hive Script & Hive UDF
- Retail use case in Hive
- Hive Demo on Healthcare Dataset

Module 6

- Hive QL: Joining Tables, Dynamic Partitioning
- Custom MapReduce Scripts
- Hive Indexes and views
- Hive Query Optimizers
- Hive Thrift Server



+44 (020) 3950 8881



+1 332-600-8899



+91 9581512555



+46 76 441 48 79



trainings@slearn.academy

- Hive UDF

Module 7

- Apache HBase: Introduction to NoSQL Databases and HBase
- HBase v/s RDBMS
- HBase Components
- HBase Architecture
- HBase Run Modes
- HBase Configuration
- HBase Cluster Deployment

Module 8

- Oozie
- Oozie Components
- Oozie Workflow
- Scheduling Jobs with Oozie Scheduler
- Demo of Oozie Workflow
- Oozie Coordinator
- Oozie Commands
- Oozie Web Console
- Oozie for MapReduce
- Combining flow of MapReduce Jobs

KEY FEATURES



Instructor-led Sessions



Learn from Experts



Assignments



24 x 7 Support



Real-time Case Studies



Get Certified

For More Info:

www.slearn.academy

www.linkedin.com/company/slearnacademy

www.facebook.com/slearnacademy

www.instagram.com/slearnacademy



+44 (020) 3950 8881



+1 332-600-8899



+91 9581512555



+46 76 441 48 79



trainings@slearn.academy