deerwalk training center

DTI-DM-001

Hadoop for Big Data Analysts

deerwalktrainingcenter.com

Program Information

Nature of the Course
Theory + Practical

Total Hours per Day
2 hours

Course Duration
80 Hours

Course Summary

This training is designed to make you a certified Big Data practitioner by providing you rich hands-on training on Hadoop Ecosystem. This Big Data $ Hadoop developer certification training is a stepping stone to your Big Data journey and you will get the opportunity to work on various Big data projects. Big Data Hadoop architects are among the highest paid professionals in the IT industry.

Course Details

Module 1 - Introduction To Big Data

- What Is Big Data And Where It Is Produced
- Rise Of Big Data
- Attributes Of Big Data
- Types Of Data
- Other Technologies Vs Big Data

Module 2 – Hadoop Architecture And Hdfs

- What Is Hadoop? Hadoop History
- Hdfs Architecture
- Daemon Types
- Learn How HDFS Works

Module 3 – Hadoop Administration

- Learn To Install Hadoop
- Installation Mode Of Hadoop
- Learn To Create And Configure Hadoop Cluster
- Troubleshooting

Module 4 – Hadoop Mapreduce Framework
Part - 1

- Overview Of Mapreduce Framework
- Installation Mode Of Hadoop
- Mapreduce Architecture
- Use Cases Of Mapreduce
- Anatomy Of Mapreduce Program
- Mapreduce API Concepts
- Writing Mapreduce Driver, Mappers, And Reducers

Module 5 – Hadoop Mapreduce Framework
Part - 2

- Learn About Combiners
- Working With Partitioners
- Multiple Output Class
- Joins With Mapreduce
- Yarn Components
- Yarn Architecture
- Working With Real Data Set

Module 6 – Apache SQOOP

- Why And What SQOOP is?
- Importing Data Using Sqoop
- Exporting Data Using SQOOP

Module 7 – Hive And Hiveql

- What Is Hive?
- Hive DDL and DML
- Partitions and Bucketing In Hive
- Managed Vs Unmanaged Tables
- Hive Serdes
- Learn To Write Own Hive Udfs
- Demo With Airlines Data Sets

Module 8 – Pig And Pig Latin

- What Is Pig And Pig Latin
- Pig Architecture
- Shell And Utility Components
- Pig Udf
- Pig Vs Mapreduce Vs Hive
- Pig Latin Demo With Movies DataSets

Module 9 – Hbase

- Hbase Architecture
- Hbase Components
- Hbase Vs RDBMS

Module 10 – Mongodb

Part - 1

- Introduction To Nosql Database
- What Is Mongodb
- Brief History Of Mongodb
- How To Download & Install Mongodb
- Starting And Stopping Mongodb Servers
- Crud Operations
- Data Modeling In Mongodb

Module 11 – Mongodb
Part - 2

- Indexing In Mongodb
- Replications
- Learn About Sharding
- Importing And Exporting Data
- Mongodb Vs Sql

Module 12 – Apache Spark

- What is Spark
- Spark Ecosystem
- Spark Components
- Downloading and installing Spark standalone
- RDD and DataFrames
- Understand how to create parallelized collections and external datasets
- Rdd Transformation and actions
- Working with spark sql
- Understanding spark streaming

Module 13 – Hadoop Security

- Ring of Defence
- Security overview of hadoop

Module 14 – Hadoop on VMs

- Learn about pre configured vms of Hadoop

---

deerwalk
training center

Sifal, Kathmandu, Nepal
Phone: +977 - 01 - 5913021 | 4567153
Mobile: +977 - 9765355167 | 9860422021
Email: training@deerwalkcompware.com
Website: deerwalktrainingcenter.com