

TIB ACADEMY

TRAINING IN BANGALORE

DELIVERING THE
BETTER TRAINING



Flexible, Affordable, Accessible
service.



CALL US

+91 9513332301



TIB ACADEMY'S SERVICE FEATURE



Affordable Fees

We provide quality training with lowest price. This opportunity is available only at TIB Academy.



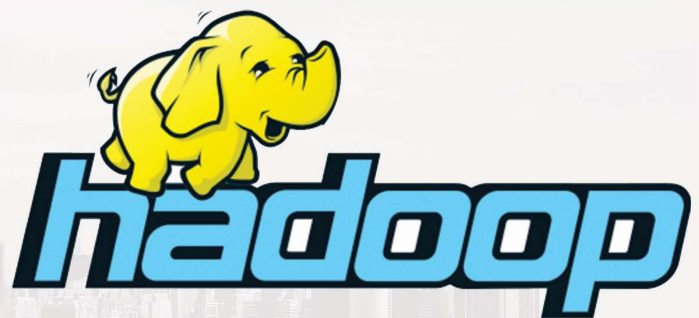
Experienced Trainers

Learn technology with a experienced professional who have expertise in their particular technology.



Flexible Timings

We believe that everyone should get the opportunity to learn their desired course. So we provide flexibility timings.



Quick Contact



Second Floor and Third Floor, 5/3 BEML
Layout, Varathur Main Road, Kundalahalli
Gate, Bengaluru, Karnataka 560066

www.traininginbangalore.com



info@tibacademy.com



+91 9513332306



Course - Curriculum

Overview of Big Data

Big Data analytics project lifecycle – brief

Hadoop Overview

Environment setup

- Setting up Single node cluster
- Setting up Multi node cluster

Hadoop Distributed File System – HDFS

- HDFS Overview
- HDFS Architecture, Operations
- Data Storage in Hadoop – Replication and Block storage
- Hadoop Daemons and its functionality
- Java based approach (Access HDFS through Java)

Map Reduce

- History of Map Reduce
- Map Reduce architecture – Programming paradigm
- How Map Reduce model works?
- Mapper and Reducer concepts
- Data types in Map Reduce
- Input and output formats of Map Reduce



Course - Curriculum

- Different phases of Map Reduce
- Partitioner and Combiner
- Data Localization in Map Reduce
- Examples of Map Reduce program – **Log Analysis**
- MR Job Configurations

YARN Framework

- Yarn Architecture
- Advantages of YARN over MR engine V1
- MR Job submission and execution procedure

Sqoop

- Overview of Sqoop
- Sqoop Installation and configuration
- Importing data from RDBMS to HDFS
- Incremental import of data
- Conditional based import
- Sqoop – to hive import
- Sqoop – Export to RDBMS
- Sqoop – Meta store
- Sqoop – Jobs configure and execution



Course - Curriculum

Hive

- Introduction to Hive
- Hive Architecture
- RDBMS vs Hive
- Map Reduce Vs Apache Pig Vs Hive
- Hive shell and HiveQL(HQL)
- Hive server & Beeline client
- Tables Management (External Vs Managed tables)
- Schemas and data types
- Partitions (Static & Dynamic partitioning)
- Bucketing
- Different file formats – rc, orc, parquet, Json, text formats
- Hive – Built in functions
- Hive – User defined functions (UDF)
- HQL - Query execution plan & performance tuning

Introduction to NoSQL

- What is NoSQL Databases?
- Fixed schema(RDBMS) Vs Flexible Schema(NoSQL)
- Json (JavaScript Object Notation) format
- Ideas Behind MongoDB / Cassandra



Course - Curriculum

HBase

- HBase Introduction
- HBase architecture
- HBase vs RDBMS (fixed Vs flexible schema)
- Master and Region servers
- HBase commands

Introduction to Apache Spark

- Introduction to Spark
- Apache Hadoop Vs Apache Spark
- Need for Spark
- Sample Programs to understand the advantages of Spark

OTHER HADOOP ECOSYSTEM COMPONENTS

Overview of Apache Pig

Overview of Flume

Overview of Oozie

Overview of Apache Kafka

HADOOP DISTRIBUTIONS

Overview & need for Cloudera (CDH) distribution

Overview of Hortonworks (HDP)



Other Activities

Assessments

- **Objective Assessments**

- Syntactical based
- Scenario based

Note: At least 2 objective based assessments in each module

- **Hands On – Practical Assessments**

- Scenario based

Note: At least 2 Hands on assessments in each module

Assignments

- **Hands On – Practical Assignments**

Note: At least 4 Hands on assignments in each module

Tasks – Home Work

- **Regular tasks on each module**

Note: Tasks are focused more to improve self learning

Resume Support & Interview Preparation

- **Work on one or two mini projects**
- **Discuss and convert the current working project into Hadoop project to add in resume & to explain Hadoop experience in interviews**