Apache Hadoop Training Course

Project Based Immersive Learning Course

Apache Hadoop Training Course Overview

Sambodhi and Education Nest's Apache Hadoop training is a comprehensive Big Data course designed by industry experts to provide you with in-depth knowledge of Hadoop and Spark modules, considering current industry job requirements. This industry-recognized certification course combines Hadoop developer, administrator, testing, and analytics with Apache Spark, preparing you to clear the Cloudera CCA175 Big Data certification. The course covers in-depth knowledge of the Hadoop Ecosystem tools such as HDFS, YARN, Map Reduce, Hive, and Pig. This Big Data program is entirely hands-on, allowing you to work on real-world projects to gain practical experience. Upon successful completion of the course, you will possess the knowledge and skills necessary to pursue a career as a Hadoop developer, administrator, tester, or analyst.

Benefits of Apache Hadoop Training Course:

Sambodhi and Education Nest's Apache Hadoop training offers many benefits, including access to numerous job opportunities in the Big Data and Analytics field. Top companies worldwide are actively recruiting Apache Hadoop professionals, making up skilling in this area a smart career decision. According to recent data, the global Hadoop-as-a-Service (HaaS) market was worth approximately USD 7.35 billion in 2019 and is expected to grow at a CAGR of 39.3%, reaching

around USD 74.84 billion by 2026. By completing this course, you will be well-prepared to take advantage of these growing opportunities in the field of Big Data and Analytics.

Who should learn?

Big Data analytics market is expanding globally, offering great opportunities for IT professionals. Our Hadoop Certification course prepares both fresher and professionals for certified Data Analytics jobs. It is best suited for:

- Graduates looking to build a career in Big Data Field
- Programming Developers
- Software Architects
- Testing professionals
- Data Analysts & Business Intelligence Professionals
- Data Engineers
- System Administrators
- DBAs and DB professionals
- ETL and Data Warehousing Professionals
- Mainframe professionals
- Project Managers
- Senior IT Professionals

Why do you need Apache Hadoop Training Course?

Average Salary Growth:

According to Payscale, the average salary of a Hadoop Developer is around $92,000 per year in the United States. However, this figure can vary based on factors such as location, years of experience, and company size. According to Indeed.com salary data, the average salary of Big Data Hadoop Developers is $135,000 per year in the United States. With the increasing demand for Big Data professionals and the growth of the Hadoop ecosystem, there is potential for salary growth and career advancement in the field of Apache Hadoop.

Industries:

There is a significant demand for Hadoop Developers across various industries and MNCs. As more and more companies are adopting Big Data technologies, the demand for Hadoop Developers is increasing rapidly. Industries such as banking, healthcare, retail, telecommunications, and entertainment are actively seeking skilled Hadoop Developers. Top companies such as Amazon, Google, Microsoft, and IBM are actively recruiting Hadoop Developers to manage and analyze their vast amounts of data.

Position in Market:

According to various job research websites, the demand for Apache Hadoop professionals is projected to grow significantly in the coming years. The global Hadoop market size was valued at USD 17.19 billion in 2020 and is expected to grow at a compound annual growth rate (CAGR) of 13.5% from 2021 to 2028, according to Grand View Research. This growth is due to the increasing adoption of big data analytics, IoT, and cloud computing technologies.

Designations:

- Big Data Architect
- Big Data Engineer
- Big Data Developer
- Hadoop Administrator

Why Apache Hadoop Training Course from Education Nest

- **Free Demo on Request**
- **Live Interactive Learning**

- **Lifetime Access**
- **Flexible Schedules**
- **24x7 Support**
- **One on One Doubt Clearing**
- **Real Time Project-Based Learning**
- **Certificate Oriented Curriculum**

Key Skills Covered:

- Real-time Data Processing
- Functional Programming
- Spark Application
- Working with HDFS
- Batch Processing using Map Reduce
- Data Analysis using Pig
- Querying Big Data using Hive and HBase
- Workflow Scheduling using Oozie

Apache Hadoop Training Course Syllabus

Module 1: Introduction to Big Data and Hadoop

- Introduction to Big Data & Hadoop
- 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: Map Reduce Framework
- Different Hadoop Distribution

Module 2: Hadoop Architecture and Hadoop Distributed File System (HDFS)

- 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: Hadoop Map Reduce Framework

- Traditional way vs Map Reduce way
- Why Map Reduce
- YARN Components
- YARN Architecture
- YARN Map Reduce Application Execution Flow
- YARN Workflow
- Anatomy of Map Reduce Program
- Input Splits, Relation between Input Splits and HDFS Blocks

Module 4: Advanced Hadoop Map Reduce

- Counters
- Distributed Cache
- MR unit

- Reduce Join
- Custom Input Format
- Sequence Input Format

Module 5: Apache Pig

- Introduction to Apache Pig
- Map Reduce vs Pig
- Pig Components & Pig Execution
- Pig Data Types & Data Models in Pig

Module 6: Introduction to Apache Hive

- What is 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

Module 7: Advanced Apache Hive and HBase

- Hive QL: Joining Tables, Dynamic Partitioning
- Custom Map Reduce Scripts
- Hive Indexes and views
- Hive Query Optimizers
- Hive UDF
- 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: Advanced Apache HBase

- HBase Data Model

- HBase Shell

- Hive Data Loading Techniques

- Apache Zookeeper Introduction

- ZooKeeper Data Model

- Zookeeper Service

- HBase Filters

Module 9: Processing Distributed Data with Apache Spark

- Spark Overview

- Spark Ecosystem

- Spark Components

- What is Scala?

- Why Scala?

- SparkContext

- Spark RDD

Module 10: Oozie

- What is Oozie?

- Oozie Components

- Oozie Workflow

- Scheduling Jobs with Oozie Scheduler

- Oozie Coordinator and Bundles

- Oozie Web Console

- Oozie for Map Reduce
- Hive in Oozie

Career Support

Profile Building:

Experienced professionals are available to offer tailored assistance in crafting your CV and online profiles, taking into account your unique educational and experiential background.

Interview Preparation:

The upcoming interview preparation service will include personalized one-on-one sessions and the option for mock interviews if needed.

Job Referrals:

At Education Nest, we receive a variety of job requirements from diverse sources such as organizations, our clients, HR consultants, and a vast network of Education Nest currently employed in different companies. We strive to meet these varied requirements to the best of our abilities.

Continuous Support:

We offer continuous support for as much time as you need it, and a considerable number of our learners receive multiple interviews offers and promising employment opportunities as a result of the abilities they gain during the program.