

Three-days workshop on "Big Data Analytics using Hadoop"

Dates: 6th Sep, 7th Sep and 8th Sep

Venue: GIMT, Guwahati

Join our World Class Workshop on "Big Data Analytics using Hadoop" being conducted by leadingindia.ai

Objective of Workshop is to:

- Learn more about the trends in Big data and how they are impacting the business world (Risk, Marketing, Healthcare, Financial Services, etc.)
- Explain this new technology and how companies can use them effectively to gather the data that they need and glean critical insights
- Explores relevant topics such as data privacy, data visualization, unstructured data, crowd sourcing data scientists, cloud computing for Big data, and much more.
- Learn about data management using NoSQL
- Explain how NoSQL databases work and the ways that they may be a superior alternative to a traditional RDBMS
- learn about file systems that manage storage across a network of machines
- Install, configure, and run Hadoop and HDFS
- Perform map-reduce analytics using Hadoop
- learn about the key features such as Pig Latin scripting language, Grunt shell, and learn how to use Hive's SQL dialect—HiveQL—to summarize, query, and analyze large datasets stored in Hadoop's distributed filesystem

For Registration [CLICK HERE](#)

Last Date of Registration (Extended to): **30 August, 2018**

Registration Fee: **1000/- Only.**

About the Institute [Click Here](#)

Contents:

I. Big Data analytics: An overview

- Introduction to Big Data Analytics
- SMAC model
- Structured and Unstructured data
- Use of hadoop and other nosql techniques
- Application domains of Hadoop

II. Hadoop Architecture

- Linux architecture and file system
- Hadoop Architecture
- Hadoop ecosystem components
- Hadoop Storage: HDFS
- Hadoop Processing: MapReduce Framework

III. Hadoop Cluster Configuration

- Hadoop Cluster Architecture
- Single-Node Hadoop Configuration
- A Typical Production Hadoop Cluster
- MapReduce Job execution
- Common Hadoop Shell commands

IV. Hadoop MapReduce framework

- Basics of Java programming
- Object oriented programming concepts
- Working with eclipse IDE
- Exception handling
- Arrays and strings in Java
- Collection in java
- Architecture of MapReduce
- Map and Reduce tasks (mapper, reducer & combiner)
- MapReduce Execution Framework
- Partitioners and Combiners
- MapReduce Programming with case study

V. Pig and Pig Latin

- Grunt
- Pig's Data Model
- Writing Evaluation
- Filter
- Load & Store Functions

VI. Hive / Jaql

- Comparison with Traditional Database
- HiveQL: Data Types
- Operators and Functions
- Introduction to Jaql
- Working with Jaql

VII. Hadoop Project:

Analysis of social sites (eg. Twitter)

Merging multiple language dictionaries in a single master dictionary.

Analysis of business models based on big data

Product review system design and many more...

Software Requirements:

1. [VMWare workstation 14](#)
2. [Cloudera VM 4.2](#)
3. VTX Should be enabled in BIOS
4. 64 bit Windows 7 or above

Hardware Requirements:

8GB Ram (4GB Ram will work but it will be very slow)

Contact Person for any query regarding the workshop at GIMT, Guwahati:

Sandip Bordoloi

Email: sandip_aei@gimt-guwahati.ac.in

Phone Number: +91 9864504431

Workshop Coordinator: Adarsh Pradhan

Email: adarsh@gimt-guwahati.ac.in Phone: +91 9864364161

Other Details

***For Participants who need accomodation in our hostel, extra accomodation charges will be applicable.**

***Only Snacks and lunch will be free during the duration of workshop.**

Nearby Hotels: 1. Treebo Rudraksh near Guwahati Airport

2. Oyo Aashiyana near Guwahati Airport

3. Eagle Nest near Guwahati Airport