UPSKILL
GLOBAL TECHNOLOGIES

BIG DATA & HADOOP COURSE CONTENT

Course Curriculum and Learning Outcomes

Module Description Learning Outcomes

Big Data and Hadoop Concepts

- Describe Big Data
- List out the limitations of the existing solution
- Define Hadoop and its components
- Describe how Hadoop solves the limitations of the existing solution

Hadoop Architecture

- Explain HDFS Architecture
- Describe the anatomy of File Read and Write
- Explain MapReduce Process Flow

Hadoop Environment Set Up

- Define a Cluster
- Describe different flavours of Hadoop
- Execute the word count example

Hadoop MapReduce Concepts

- Compare and contrast traditional approach with Map Reduce way
- Differentiate between Block and Split
- Describe Combiners
- Discuss Partitioners

Analytics Using Hive

- Discuss use cases for Hive
- Compare Hive and Pig, also Hive and RDBMS
- Describe Hive components
- Execute Hive queries
- Identify existing data challenges

Analytics Using Mysql

- Describe features of a likely solution
- Define Mysql
- Explain data model in Mysqlarchitecture

Hadoop 2.0 & Identify challenges with Hadoop 1.0

- Apache Fsqoop ,lume and Zookeeper
- Discuss solutions provided by Hadoop 2.0 in terms of YARN
- Explain Hadoop 2.0 Process Flow