Big Data

Duration – 40 Hours

Program Description

This program provides learners with a strong foundation in Big Data technologies, frameworks, and applications. It equips participants to process, analyze, and manage large-scale data sets efficiently, enabling data-driven decision-making across industries. The program covers modern architectures, distributed computing frameworks, data pipelines, and integration with cloud platforms. By the end of the course, learners will be prepared to design, implement, and optimize Big Data solutions in real-world scenarios.

Learning Goals

- Understand the fundamentals of Big Data and its ecosystem
- Gain proficiency in distributed storage and processing frameworks
- Learn to design and manage scalable data pipelines
- Apply data ingestion, transformation, and integration techniques
- Explore cloud-based Big Data solutions for scalability and cost efficiency
- Apply analytics and visualization for business insights
- Develop problem-solving and solution-design skills for enterprise use cases

Course Topics

- Introduction to Big Data and Ecosystem
- Hadoop & Distributed Computing
- Spark for Big Data Processing
- Data Ingestion & Pipeline Design
- NoSQL & Big Data Storage Systems
- Big Data on Cloud (AWS, Azure, GCP)
- Big Data Analytics & Visualization
- Security, Governance & Compliance
- Capstone Project

Modules can be customized to suit client's specific needs and duration accordingly