INFOSOFT IT SOLUTIONS

Training | Projects | Placements

Revathi Apartments, Ameerpet, 1st Floor, Opposite Annapurna Block, Infosoft It solutions, Software Training & Development Institute, +91-9059683947|+91-9182540872

Big Data Analytics Content

Introduction to Big Data Analytics

- Sources of Big Data
- Structured and Unstructured Data
- Data Storage and Processing Technologies
- Data Analytics Frameworks and Tools
- Introduction to Hadoop Ecosystem
- Hadoop Distributed File System (HDFS)
- MapReduce Programming Model
- Apache Spark,Apache Hive, Apache PigIntroduction
- Data Warehousing Concepts
- Data Mining Techniques

Data Mining Techniques

- Introduction to data mining
- Data preprocessing techniques
- Data transformation techniques
- Association rule mining

- Clustering techniques
- Classification techniques
- Anomaly detection
- Text mining
- Time series analysis
- Evaluation and validation of data mining model

Data Warehousing and ETL

- Introduction to Data Warehousing
- Data Warehousing Architecture
- OLAP and OLTP Concepts
- ETL (Extract, Transform, Load) Process and its importance
- Data Modeling for Data Warehousing
- Dimensional Modeling
- Fact Tables and Dimension Tables
- ETL Tools and Techniques
- Data Cleansing and Data Profiling
- Performance Tuning and Optimization for Data Warehousing

Big Data Analytics Tools

- Hadoop ecosystem tools: This includes HDFS, MapReduce, Pig, Hive, Sqoop, Flume, etc.
- NoSQL databases: MongoDB, Cassandra, HBase, etc.
- Tools for data visualisation include Tableau, QlikView, Power BI, and others.
- Machine learning tools: Python, R, Mahout, etc.

- Cloud-based Big Data Analytics tools: AWS, Azure, Google Cloud Platform, etc.
- Apache Spark and related tools: Spark SQL, Spark Streaming, Spark MLlib, etc.
- Graph databases: Neo4j, OrientDB, etc.

Data Visualization

- An overview of data visualisation and its significance
- Types of data visualization techniques and tools
- Best practices for designing effective visualizations
- Principles of visual perception and how they apply to data visualization
- Exploratory data analysis using visualizations
- Interactive visualizations and dashboards
- Storytelling with data and creating compelling narratives
- Designing for different audiences and purposes
- Common data visualization pitfalls to avoid
- Advanced data visualization techniques and tools

Algorithms for Machine Learning in Big Data Analytics

- Overview of different types of machine learning algorithms
- Preprocessing and cleaning of big data for machine learning
- Regression analysis for big data using linear and logistic regression
- Clustering algorithms for unsupervised learning
- Dimensionality reduction techniques
- Classification algorithms
- Deep learning algorithms

- Evaluating and selecting appropriate machine learning models for Big Data Analytics
- Deploying and scaling machine learning models in distributed computing
- environments such as Hadoop and Spark

Real-time Big Data Analytics

- Understanding the characteristics of real-time data
- Overview of real-time big Data Analytics architecture
- Real-time data ingestion techniques and tools
- frameworks like Apache Kafka, Apache Storm
- Real-time data processing and analysis techniques
- Real-time machine learning techniques
- Real-time data visualization techniques and tools
- Use cases and applications of real-time Big Data Analytics

Case Studies and Projects

- Customer Segmentation
- Fraud Detection
- Sentiment Analysis
- Predictive Maintenance
- Supply Chain Optimization
- Healthcare Analytics
- Financial Analytics
- IoT Analytics
- Social Network Analysis
- Image and Video Analytics