

New Scheme Based On AICTE Flexible Curricula

CSE-Artificial Intelligence and Machine Learning/ Artificial Intelligence and Machine Learning, VIII-Semester

AL802 (C) Big Data Analytics

Course Outcomes:

1. Students should be able to understand the concept and challenges of Big data.
2. Students should be able to demonstrate knowledge of big data analytics.
3. Students should be able to develop Big Data Solutions using Hadoop Eco System
4. Students should be able to gain hands-on experience on large-scale analytics tools.
5. Students should be able to analyse the social network graphs.

Course Content

Unit1: Introduction to Big data, Big data characteristics, Types of big data, Traditional Versus Big data, Evolution of Big data, challenges with Big Data, Technologies available for Big Data, Infrastructure for Big data, Use of Data Analytics, Desired properties of Big Data system.

Unit2: Introduction to Hadoop, Core Hadoop components, Hadoop Eco system, Hive Physical Architecture, Hadoop limitations, RDBMS Versus Hadoop, Hadoop Distributed File system, Processing Data with Hadoop, Managing Resources and Application with Hadoop YARN, Map Reduce programming.

Unit3: Introduction to Hive Hive Architecture, Hive Data types, Hive Query Language, Introduction to Pig, Anatomy of Pig, Pig on Hadoop, Use Case for Pig, ETL Processing, Data types in Pig running Pig, Execution model of Pig, Operators, functions, Data types of Pig.

Unit4: Introduction to NoSQL, NoSQL Business Drivers, NoSQL Data architectural patterns, Variations of NOSQL architectural patterns using NoSQL to Manage Big Data, Introduction to Mango DB.

Unit5: Mining social Network Graphs: Introduction Applications of social Network mining, Social Networks as a Graph, Types of social Networks, Clustering of social Graphs Direct Discovery of communities in a social graph, Introduction to recommender system.

Text Books:

1. RadhaShankarmani, M. Vijaylakshmi, " Big Data Analytics", Wiley, Second edition
2. Seema Acharya, SubhashiniChellappan, " Big Data and Analytics", Wiley, First edition

Reference Books:

1. I.KaiHwang, Geoffrey C., Fox. Jack, J. Dongarra, "Distributed and Cloud Computing", Elsevier, First edition
2. Michael Minelli, Michele Chambers, AmbigaDhiraj, "Big Data Big Analytics", Wiley for old question papers visit <http://www.rgpvonline>