| CS423: Big Data Analytics | L | T | P | Mathematics, aptitude. |
|---|---|---|---|---|
| | 3 | 0 | 2 | |

**Course Objective:** Be Exposed With the Basic Rudiments of Business Intelligence System. Understand The Modeling Aspects Behind Business Intelligence. Understand Of the Business Intelligence Life Cycle and the Techniques Used In It.

| S. No. | Course Outcomes (CO) |
|---|---|
| CO1 | Understand big data characteristics, storage, and high-performance architectures like HDFS and MapReduce. |
| CO2 | Apply clustering (K-means) and classification methods (decision trees, Naïve Bayes) and evaluate their effectiveness. |
| CO3 | Implement association rules and recommendation systems using algorithms like Apriori and various recommendation approaches. |
| CO4 | Analyze stream data, real-time analytics, and apply graph analytics. |
| CO5 | Utilize NoSQL databases and visualization tools for big data management and analysis. |

| S. No | Contents | Contact Hours |
|---|---|---|
| UNIT 1 | Introduction to Big Data: Evolution of Big data, Best Practices for Big data Analytics – Big data characteristics, Validating, The Promotion of the Value of Big Data, Big Data Use Cases- Characteristics of Big Data Applications, Perception and Quantification of Value, Understanding Big Data Storage, A General Overview of High-Performance Architecture, HDFS – MapReduce and YARN, Map Reduce Programming Model. | 8 |
| UNIT 2 | Clustering and Classification: Advanced Analytical Theory and Methods: Overview of Clustering ,K-means, Use Cases – Overview of the Method, Determining the Number of Clusters, Diagnostics, Reasons to Choose and Cautions, Classification: Decision Trees Overview of a Decision Tree, The General Algorithm, Decision Tree Algorithms Evaluating a Decision Tree, Decision Trees in R, Naïve Bayes, Bayes' Theorem, Naïve Bayes Classifier. | 6 |
| UNIT 3 | Association and Recommendation System: Advanced Analytical Theory and Methods-Association Rules, Overview – Apriori Algorithm, Evaluation of Candidate Rules, Applications of Association Rules, Finding Association& finding similarity, Recommendation System: Collaborative Recommendation- Content Based Recommendation, Knowledge Based Recommendation, Hybrid Recommendation Approaches. | 8 |
| UNIT 4 | Classification :Classification: Decision Trees, Overview of a Decision Tree, The General Algorithm, Decision Tree Algorithms, Evaluating a Decision Tree, Decision Trees in R, Naïve Bayes, Bayes' Theorem, Naïve Bayes Classifier. | 6 |
| UNIT 5 | Stream Memory: Introduction to Streams Concepts, Stream Data Model and Architecture, Stream Computing, Sampling Data in a Stream, Filtering Streams, Counting Distinct Elements in a Stream, Estimating moments, Counting oneness in a Window, Decaying Window, Real time Analytics Platform(RTAP) applications, Case Studies, Real Time Sentiment Analysis, Stock Market Predictions. Using Graph Analytics for Big Data: Graph Analytics. | 8 |
| UNIT 6 | NoSQL Data Management For Big Data And Visualization: NoSQL Databases: Schema-less Models: Increasing Flexibility for Data Manipulation- Key Value Stores- Document Stores, Tabular Stores, Object Data Stores, Graph Databases Hive, Sharding, Hbase, Analyzing big data with twitter, Big data for E-Commerce Big data for blogs, Review of Basic Data Analytic Methods using R. | 6 |
| | Total | 42 |