



U20IT701 BIG DATA ANALYTICS

IV YEAR IT - REGULATION 2023

PART B QUESTION BANK

UNIT-I INTRODUCTION TO BIG DATA

1. Briefly explain the Evolution of Bigdata.
2. Describe in detail about application suites to Big Data Analytics .
3. Write note on
 - a. Big data use case.
 - b. Characteristics of Big Data Applications
4. Write in detail note on HDFS.
5. Explain in detail about the following
 - i) Map Reduce and YARN
 - ii) Map Reduce Programming Model
6. Explain the Characteristics of Big Data
7. Describe in detail about the practices for Big Data Analytics.
8. Explain Big Data Validating.

UNIT-2 CLUSTERING AND CLASSIFICATION

1. Explain in detail about the Decision Trees.
2. Describe in detail about the reasons to choose and cautions of k-means.
3. Explain the General Algorithm.
4. Explain in detail about the Decision Tree algorithms.
5. Explain the Decision Tree in R.
6. Describe in brief on the Bayes' Theorem.
7. Explain in brief about the Naïve Bayes Classifier.

UNIT-3 ASSOCIATION AND RECOMMENDATION SYSTEM

1. Explain in detail about the Apriori Algorithm and Evaluation of Candidate Rules.
2. Write the Application of Association Rules and finding similarities and association.
3. Briefly explain the Collaborative Recommendation.
4. Briefly explain the Content Based Recommendation
5. Briefly explain the Knowledge based Recommendation

6. Briefly explain the Hybrid Recommendation
7. Explain the Demographic Filtering and evaluation of Recommender Systems.

UNIT-4 STREAM THEORY

1. Explain the Stream Data Model and Architecture of the streams.
2. Briefly describe the Stream Computing.
3. Explain Sampling data in a stream.
4. Describe in detail about the Big Data using the Graph Analytics.
5. Write a note on the following
 - i) Decaying Window
 - ii) RTAP
6. Describe the counting oneness in a window for stream.
7. Estimating and explain the computer moments in stream.
8. Briefly on the following.
 - a) counting Distinct Element in a stream
 - b) filtering Streams
9. How the stream memory used in real time application as case studies?

UNIT-5 NOSQL DATA MANAGEMENT FOR BIG DATA AND VISUALIZATION

1. Explain briefly on the R Programming language.
2. Briefly describe detail about the Exploratory Method in R.
3. Briefly Analysis the big data for blog.
4. Give the brief analysis of big data for the real time application like E-Commerce.
5. List out the sharding along with the brief explanation of its Architecture.
6. Describe the Architecture of HBase and its mechanism.
7. Explain the brief analysis of big data in social networking.
8. Write a short note on the following.
 - a) NoSQL Database
 - b) Hive
9. Explain in detail about the schema-less Models.