

BIG DATA ANALYTICS

Paper Code: ETIT-406

Paper: Big Data Analytics

L	T/P	C
3	0	3

INSTRUCTIONS TO PAPER SETTERS:

MAXIMUM MARKS: 75

1. Question No. 1 should be compulsory and cover the entire syllabus. This question should have objective or short answer type questions. It should be of 25 marks.
2. Apart from Question No. 1, rest of the paper shall consist of four units as per the syllabus. Every unit should have two questions. However, student may be asked to attempt only 1 question from each unit. Each question should be of 12.5 marks.

Objective: To introduce the students about knowledge of Data Management, Big Data stacks and Data analysis.

UNIT-I

Big Data Introduction: The Evolution of Data Management, Defining Big Data, Traditional and advanced analytics. Distributed Computing, need of distributed computing for big data, economics of computing, latency problem.

Examining Big Data Types, Structured Data, sources of big structured data, role of relational databases in big data, Unstructured Data, sources of unstructured data, role of a CMS in big data management.

[T1][R1][No. of Hrs. 11]

UNIT-II

Big Data Stack: Redundant Physical Infrastructure, Security Infrastructure, Operational Databases. Organizing Data Services and Tools, Analytical Data Warehouses, Big Data Analytics, Big Data Applications.

Virtualization and big data: Server virtualization, Application virtualization, Network virtualization, Processor and memory virtualization, Data and storage virtualization, Managing Virtualization with the Hypervisor.

[T1][No. of Hrs. 10]

UNIT-III

MapReduce Fundamentals, Putting map and reduce Together, Optimizing MapReduce Tasks. Hadoop, Hadoop Distributed File System (HDFS), Name Nodes, Data nodes, Hadoop MapReduce.

[T1][T2][R1][No. of Hrs. 11]

UNIT-IV

Big Data Analytics: Basic analytics, Advanced analytics, Operationalized analytics, Monetizing analytics, Text Analytics and Big Data, Social media analytics, Text Analytics Tools for Big Data, Attensity, Clarabridge, OpenText.

Integrating Data Sources: Dealing with Real-time Data Streams and Complex Event Processing, Operationalizing Big Data, Applying Big Data within Your Organization, Security and Governance for Big Data Environments.

[T1][No. of Hrs. 11]

Text Books:

[T1] Judith S. Hurwitz, Alan F. Nugent, Fern Halper, Marcia A. Kaufman, "Big Data For Dummies", John Wiley & Sons, Inc.(2013)

[T2] Robert D. Schneider, "Hadoop For Dummies", John Wiley & Sons, Inc. (2012)

Reference Books:

[R1] Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data, by Paul Zikopoulos, McGraw Hill 2012.