

SHRI G.S. INSTITUTE OF TECHNOLOGY & SCIENCE, INDORE
DEPARTMENT OF INFORMATION TECHNOLOGY
Lecture Plan -2024-25

Subject Nomenclature: Big data Analytics
Subject Code: IT 48601
Year/semester: IV yr 8th sem
Instructor Name Dr Sunita Varma/Ms. Neha Agrawal

Course Outcome : At the end of this course student will be able to:

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

Prerequisite: knowledge of cloud,Data Base Management System and java programming.

Lecture Content:

S No.	Date	Content Covered	Interaction mode lecture/live coding/method
1	08-Jan-2025	Introduction to Big Data: Definition, Importance, Examples	Online mode - Slide,Notes
2	13-Jan-2025	Evolution of Big Data and Data Lifecycle	Online mode - Slide,Notes
3	14-Jan-2025	Characteristics of Big Data: The 5 Vs	Online mode - Slide,Notes
4	15-Jan-2025	Challenges with Big Data: Storage, Processing, Governance	Online mode - Slide,Notes
5	20-Jan-2025	Technologies for Big Data: Hadoop, Spark, NoSQL Overview	Online mode - Slide,Notes
6	21-Jan-2025	Infrastructure for Big Data: Distributed Computing & Cloud	Online mode - Slide,Notes

7	22-Jan-2025	Use of Data Analytics: Types and Tools	Online mode - Slide,Notes
8	03-Feb-2025	Desired Properties of Big Data Systems: Scalability, Fault Tolerance	Online mode - Slide,Notes
9	04-Feb-2025	Introduction to Hadoop: History, Components	Online mode - Slide,Notes
10	05-Feb-2025	Hadoop Core Components: HDFS & MapReduce Overview	Online mode - Slide,Notes
11	10-Feb-2025	HDFS in Detail: Architecture, Namenode, Datanode	Online mode - Slide,Notes
12	11-Feb-2025	MapReduce in Detail: Mapper, Reducer, Workflow	Online mode - Slide,Notes
13	12-Feb-2025	Hadoop Limitations: Batch Processing, Real- Time Needs	Online mode - Slide,Notes
14	17-Feb-2025	RDBMS vs Hadoop: Comparison and Use Cases	Online mode - Slide,Notes
15	18-Feb-2025	Hadoop Distributed File System: Read/Write Operations	Online mode - Slide,Notes
16	24-Feb-2025	YARN Architecture: Resource Manager, Node Manager	Online mode - Slide,Notes
17	25-Feb-2025	Managing Applications with YARN	Online mode - Slide,Notes
18	03-Mar-2025	Hadoop Ecosystem Tools Overview: Hive, Pig, Sqoop, etc.	Online mode - Slide,Notes
19	04-Mar-2025	Introduction to Hive and Hive Architecture	Online mode - Slide,Notes
20	05-Mar-2025	Hive Data Types and Table Creation	Online mode - Slide,Notes
21	10-Mar-2025	Hive Query Language (HQL): Basics and Examples	Online mode - Slide,Notes
22	11-Mar-2025	Hive Architecture in Depth: Metastore,	Online mode -

		Compiler	Slide,Notes
23	12-Mar-2025	Introduction to Pig and Its Use Cases	Online mode - Slide,Notes
24	17-Mar-2025	Pig Architecture and Execution Modes	Online mode - Slide,Notes
25	18-Mar-2025	Pig Latin Basics: Syntax and Operators	Online mode - Slide,Notes
26	19-Mar-2025	ETL Processing with Pig	Online mode - Slide,Notes
27	24-Mar-2025	Pig Data Types and Eval Functions	Online mode - Slide,Notes
28	25-Mar-2025	Execution Model of Pig and Comparison with Hive	Online mode - Slide,Notes
29	26-Mar-2025	Introduction to NoSQL: Need and Concepts	Online mode - Slide,Notes
30	01-Apr-2025	Business Drivers for NoSQL	Online mode - Slide,Notes
31	02-Apr-2025	NoSQL Data Architectural Patterns	Online mode - Slide,Notes
32	07-Apr-2025	Variations in NoSQL Architectures: Key-Value, Column, Graph, Document	Online mode - Slide,Notes
33	08-Apr-2025	Comparison of NoSQL Types	Online mode - Slide,Notes
34	09-Apr-2025	Managing Big Data using NoSQL	Online mode - Slide,Notes
35	14-Apr-2025	Introduction to Social Network Mining and Applications	Online mode - Slide,Notes
36	15-Apr-2025	Social Networks as Graphs: Nodes and Edges	Online mode - Slide,Notes