

Lecture Plan

Name of Course	Big Data
AKTU Course Code	BCS-061
Branch	Computer Science and Engineering
Semester	6
Section	CS
Total Number of Students	72
Name of Faculty	Ram Prakash
Number of Lecture Proposed	45

S.No.	Unit No	Topic	CO	No of Lectures Required	Actual Date of Completion	Suggested Reference
1	1	Introduction to Big Data: Type of digital data, history of Big Data innovation, Introduction to Big Data Platform	1	01	02/02/2026	4,1
2		Drivers for Big Data, Big Data architecture and characteristics, 5Vs of Big Data		01	04/02/2026	4,1
3		Big data technology components, Big Data importance		01	05/02/2026	4
4		Big Data features		01	06/02/2026	2
5		Big Data Analytics, Challenges of conventional systems,		01	07/02/2026	4,1
6		Intelligent data analysis		01	09/02/2026	3
7		Nature of Data, analysis vs reporting		01	11/02/2026	3
No of Lectures Required to complete Unit 1				7		
8	2	Hadoop: History of Hadoop, Apache Hadoop	2	01	12/02/2026	2,4
9		The Hadoop Distributed File System, components of Hadoop		01	13/02/2026	3
10		Data format, Hadoop pipes, Hadoop Echo System		01	14/02/2026	3
11		MapReduce: MapReduce framework and basic, How MapReduce works		01	16/02/2026	3
12		Unit tests with MR unit, test data and local tests		01	18/02/2026	3
13		Anatomy of a MapReduce job run, failures, job scheduling, shuffle and sort, task execution		01	19/02/2026	2
14		MapReduce types, input formats, output formats		01	20/02/2026	2
15		MapReduce features, real-world MapReduce		01	20/02/2026	2
No of Lectures Required to complete Unit 2				8		
16	3	HDFS (Hadoop Distributed File System): Design of HDFS, Benefits and challenges	3	1	23/02/2026	3,2
17		File sizes, block sizes and block abstraction inn HDFS		1	25/02/2026	3,2
18		Data replication, how does HDFS store, read, and write files		1	09/03/2026	3,2
19		Java interfaces to HDFS, command line interface, Hadoop file system interfaces		1	26/02/2026	3,2
20		Data flow data ingest with Flume and Scoop, Hadoop archives		1	27/02/2026	3,1
21		Hadoop I/O		1	28/02/2026	3,1
22		Hadoop Environment: Hadoop cluster, cluster specification, HDFS monitoring & maintenance		1	06/03/2026	3,1
23		Hadoop configuration, security in Hadoop, Hadoop in the cloud		1	07/03/2026	3,1
No of Lectures Required to complete Unit 3				8		

24	4	Hadoop Eco System and YARN: Hadoop ecosystem components, schedulers, fair and capacity	4	1	09/03/2026	2
25		Hadoop 2.0 New Features- NameNode high availability		1	11/03/2026	2
26		HDFS federation MRv2, Running MR1 in YARN		1	16/03/2026	2,4
27		NoSQL Databases: Introduction to NoSQL		1	18/03/2026	2,3
28		MongoDB: Introduction, data types		1	19/03/2026	2,3
29		Updating and deleting documents, querying		1	20/03/2026	2,3
30		Spark: Installing spark, spark applications, jobs		1	23/03/2026	4
31		Stages and tasks, Resilient Distributed Databases, Spark on YARN		1	25/03/2026	1,2
32		SCALA: Introduction, classes and objects, basis types and operators		1	27/03/2026	1
33		Basic types and operators, Functions and closures		1	28/03/2026	2,3
No of Lectures Required to complete Unit 4				10		
34	5	Hadoop Eco System Frameworks: Applications on Big Data using Pig, Hive, HBase	5	1	30/03/2026	1,4
35		Pig- Introduction to PIG, Execution Modes of Pig		1	02/04/2026	1,4
36		Grunt, Pig Lattin, UserDefined Functions		1	03/04/2026	1,3
37		Hive- Apache Hive architecture and installation, Hive shell, Hive services		2	06/04/2026	1,4
38		HiveQL, tables, querying data and user defined functions. MapReduce scripts, join & subqueries		1	08/04/2026	1,4
39		HBase- HBase concepts, clients, example		1	11/04/2026	1
40		HBase vs RDBMS, scheme design.		2	13/04/2026	1
41		Zookeeper- How it helps in monitoring a cluster, how to build applications with Zookeeper		2	16/04/2026	1
42		BigInsights and Big Sheets, introduction to Big SQL		1	20/04/2026	1
No of Lectures Required to complete Unit 5				12		

Text Books & References	
1	Thomas Erl, Wazid Khattak, Paul Buhler, Big Data Fundamentals: Concepts, Drivers and Techniques”, Prentice Hall.
2	ArshdeepBahga, Vijay Madiseti, “Big Data Science & Analytics: A HandsOn Approach”, VPT.
3	DT Editorial Services, Big-Data Black Book, Wiley.
4	Radha Shankarmani, “Big Data Analytics”, Wiley.

Course Outcomes (COs)		
At the end of this course students will demonstrate the ability to:		
CO 1	Demonstrate knowledge of Big Data Analytics concepts and its applications in business.	
CO 2	Demonstrate functions and components of MapReduce Framework and HDFS.	
CO 3	Discuss Data Management concepts in NoSQL environment.	
CO 4	Explain process of developing MapReduce based distributed processing applications.	
CO 5	Explain process of developing applications using HBASE, Hive, Pig etc.	

Name of Faculty with Signature: Ram Prakash

Lecture Plan In-charge:

Head of Department:

Dean Academics: