C. U. SHAH UNIVERSITY – WADHWAN CITY

FACULTY OF TECHNOLOGY AND ENGINEERING
DEPARTMENT OF COMPUTER ENGINEERING
M. TECH. SEMESTER: - II

SUBJECT NAME: Big Data and Analytics (BDA)
SUBJECT CODE: 5TE02BDA1

Teaching & Evaluation Scheme: -

Subject Code	Subject Name	Teaching Scheme (Hours)				Credits	Evaluation Scheme							Total
							Theory				Practical (Marks)			
							Sessional Exam		University Exam		Internal		University	
		Th	Tu	Pr	Total		Marks	Hours	Marks	Hours	Pr/Viva	TW	Pr	
5TE02BDA1	Big Data and Analytics	4	0	2	6	5	30	1.5	70	3.0	-	20	30	150

Objectives:
- To explore the fundamental concepts of big data analytics
- To learn to analyze the big data using intelligent techniques.
- To understand the various search methods and visualization techniques.
- To understand the applications using Hadoop and Map Reduce Concepts.
- To develop algorithms to analyze and model data.

Prerequisites:

This Course needs the thorough understanding of Big Data and its Tools & Techniques.

Course outline:

Sr. No.	Course Contents
1	**Introduction of Big Data Analytics** What Big Data Is, How Data Is Changing, Shift in Processing Due to Big Data. Changing Focus with Big Data. The Role of the Data Analyst, Implementing Big Data Analytics within an Organization, Blending Data from Multiple Sources.
2	**Analyzing Big Data in Context and Predictive Analysis** Focus on Context, Combining Big Data with Spatial Data, Leveraging External Data Provider Resources. Analytics and Big Data, Why Do Predictive Analytics on Big Data? Moving Predictive Analytics to the Front Lines, Gaining Real Business Value from Predictive Analysis.
3	**Humanizing Big Data Analytics.** Humanizing Data Design Principles. Humanizing Big Data Analytics Workflow, Considering Consumerization of Big Data Analytics, Publishing Data and Analytics to Cloud Service, Focusing on Consuming Applications, Platform for Strategic Analytics.

4	**Hadoop and Map Reduce** History of Hadoop ,The Hadoop Distributed File System, Components of Hadoop Analyzing the Data with Hadoop ,Scaling Out ,Hadoop Streaming , Design of HDFS ,Java interfaces to HDFS , Basics Developing a Map Reduce Application ,How Map Reduce Works ,Anatomy of a Map Reduce Job run ,Failures ,Job Scheduling ,Shuffle and Sort ,Task execution , Map Reduce Types and Formats, Map Reduce Features
5	**Framework and Visualization** Hive, MapR, Hive, Pig, S3, Visualizations, Visual Data Analysis Techniques, Interaction Techniques.

Learning Outcomes:

At the end of this course the student will be able to
- Work with big data platform
- Analyze the big data analytic techniques for useful business applications.
- Design efficient algorithms for mining the data from large volumes.
- Explore on Big Data applications Using Pig and Hive
- Understand the fundamentals of various big data analysis techniques.

Books Recommended:

1. Big Data Analytics for Dummies, **Michael Wessler**, Wiley Publication(2013).
2. Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data, **Zikopoulos, Paul, Chris Eaton**, Tata McGraw Hill Publications, 2011.
3. Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business, **Michael Minelli, Michehe Chambers**, 1st Edition, Wiley, 2013.

Reference Books:

1. Hadoop: The Definitive Guide, **Tom White**, 3rd Edition, O'reilly Media, 2012.
2. Data-Intensive Text Processing with MapReduce, **Jimmy Lin and Chris Dyer** , Morgan & laypool Publishers, 2010.