

GANPAT UNIVERSITY									
FACULTY OF ENGINEERING & TECHNOLOGY									
Programme	Bachelor of Technology				Branch/Spec.	Electronics and Communication Engineering			
Semester	VII				Version	1.0.0.0			
Effective from Academic Year	2026-27				Effective for the Batch admitted in	July 2023			
Course Code	2EC71PE06	Course Name			Big Data Analytics				
Teaching Scheme					Examination Scheme (Marks)				
(Per week)	Lecture (DT)		Practical (Lab.)		Total	CE	SEE	Total	
	L	TU	P	TW					
Credit	3	0	1	0	4	Theory	40	60	100
Hours	3	0	2	0	5	Practical	30	20	50
Pre-requisites									
Database Management Systems (DBMS)									
Course Outcomes									
On successful completion of the course, the students will be able to:									
CO1	Students will to build and maintain reliable, scalable, distributed systems with Apache Hadoop.								
CO2	Students will be able to write Map-Reduce based Applications								
CO3	Students will be able to design and build MongoDB based Big data Applications and learn MongoDB query language								
CO4	Students will learn difference between conventional SQL query language and NoSQL basic concepts								
CO5	Students will learn tips and tricks for Big Data use cases and solutions.								
Theory Syllabus									
Unit	Content								Hrs.
1	Introduction To Big Data: Introduction– distributed file system–Big Data and its importance, Four Vs, Drivers for Big data, Big data analytics, Big data applications. Algorithms using map reduce								5
2	Introduction To Hadoop And Hadoop Architecture: Big Data – Apache Hadoop & Hadoop EcoSystem, Moving Data in and out of Hadoop – Understanding inputs and outputs of MapReduce -, Data Serialization.								10
3	HDFS, HIVE and HIVEQL, HBASE HDFS-Overview, Installation and Shell, Java API; Hive Architecture and Installation, Comparison with Traditional Database, HiveQL Querying Data, Sorting And Aggregating, Map Reduce Scripts, Joins & Sub queries, HBase concepts, Advanced Usage, Schema Design, Advance Indexing, PIG, Zookeeper , how it helps in monitoring a cluster, HBase uses Zookeeper and how to Build Applications with Zookeeper								8
4	SPARK: Introduction to Data Analysis with Spark, Downloading Spark and Getting Started, Programming with RDDs, MachineLearning with MLlib.								10
5	NoSQL: What is it?, Where It is Used Types of NoSQL databases, Why NoSQL?, Advantages of NoSQL, Use of NoSQL in Industry, SQL vs NoSQL, NewSQL								5
6	Data Base for the Modern Web: Introduction to MongoDB key features, Core Server tools, MongoDB through the JavaScript’s Shell, Creating and Querying through Indexes, Document-Oriented, principles of schema design, Constructing queries on Databases, collections and Documents , MongoDB Query Language.								7
Practical Content									
Practical assignments and tutorials are based on the above syllabus.									
Text Books									
1	Big Data and Analytics by Seema Acharya, Subhashini Chellappan, Wileyindia Publication								

Reference Books	
1	MongoDB in Action, Kyle Banker, Piter Bakkum , Shaun Verch, Dream tech Press
2	Chris Eaton, Dirk deRooset al. , “Understanding Big data ”, McGraw Hill, 2012.
3	Tom White, “HADOOP: The definitive Guide”, O Reilly 2012.
ICT/MOOCs Reference	
1	https://www.coursera.org/learn/introduction-to-big-data-with-spark-hadoop
2	https://elearn.nptel.ac.in/shop/nptel/big-data-computing/
3	https://www.coursera.org/specializations/big-data-processing-using-hadoop

Mapping of CO with PO and PSO:															
	P O 1	P O 2	P O 3	P O 4	P O 5	P O 6	P O 7	P O 8	P O 9	P O 10	P O 11	P O 12	P S O 1	P S O 2	P S O 3
CO1	3	2	3	2	3	0	0	0	1	0	0	2	2	3	2
CO2	3	3	3	2	3	0	0	0	1	0	0	2	2	3	2
CO3	2	2	3	2	3	0	0	0	1	0	0	2	2	3	3
CO4	2	3	2	2	2	0	0	0	0	0	0	2	2	2	2
CO5	2	3	3	3	3	2	1	1	2	1	2	3	3	2	3