

GANPAT UNIVERSITY									
FACULTY OF MANAGEMENT STUDIES									
Programme	Bachelor of Business Administration				Branch/Spec.	Business Analytics			
Semester	VII				Version	1.0.0.0			
Effective from Academic Year		2026-27			Effective for the Batch admitted in			July 2023	
Course Code	7A01IBD	Course Name			Introduction to Big Data Analytics				
Teaching Scheme					Examination Scheme (Marks)				
(Per week)	Lecture (DT)		Practical (Lab.)		Total		CE	SEE	Total
	L	TU	P	TW					
Credit	04	00	00	00	04	Theory	40	60	100
Hours	04	00	00	00	04	Practical	00	00	00
Pre-requisites									
Database Management System, JAVA/Python Programming Language									
Course Objective:									
Provide students with a foundational understanding of Big Data concepts and distributed computing.									
Course Outcomes									
On successful completion of the course, the students will be able to:									
CO1	Understand distributed file system concepts and algorithms								
CO2	Understand big data phenomena and its applications								
CO3	Implement distributed storage and big data processing concepts using open-source tools like Hadoop and Spark								
CO4	Apply Hadoop and Apache Spark ecosystem tools to implement big data related case studies.								
Theory Syllabus									
Unit	Content								Hrs.
1	Distributed file system fundamentals: Introduction to distributed file system, Difference between normal file System and distributed file system, Benefits and requirement of distributed file system.								15
2	Introduction to Big data: What exactly is big data? Comparison of small data large data and big data, Big data types, Source of big data, Big data examples, Big data challenges, Big data analytics, Data Science vs. Big Data vs. Data Analytics, Types of Big Data Analytics, Lifecycle of Big Data Analytics.								15
3	Hadoop Architecture (HDFS, MapReduce): Introduction to Hadoop Distributed, File System (HDFS), HDFS commands, HDFS Architecture, HDFS read and write operations, MapReduce Framework, Map Reduce phases.								15
4	Apache Spark: Apache Spark Architecture, Concept of Data Lake, Concept of Streaming data, Challenges Building Data, Streaming Application, Use Cases for Streaming data Application.								15
Exam: Theory 100%, Numerical 0%									
Practical Content									
NA									
Text Books									
1	Cloud Computing: Principles and Paradigms by Rajkumar Buyya, James Broberg, Andrzej M Goscinski								
Reference Books									
1	Hadoop: The Definitive Guide, By Tom White								
2	Big Data and Analytics, by Subhashini Chellappan Seema Acharya								
3	Big Data Analytics with Hadoop 3 by Sridhar Alla								
ICT/MOOCs Reference									

1	https://nptel.ac.in/courses/106/104/106104189/														
Mapping of CO with PO and PSO:															
	Course Outcome (CO) No.	PO-CO Mapping								PSO-CO Mapping					
		PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6
	CO1	3	1	1	1	1	1	1	1	2	1	1	2	1	1
	CO2	3	1	1	1	1	1	2	2	3	2	1	2	1	1
	CO3	3	2	1	2	1	1	3	3	3	2	1	3	1	1
	CO4	3	2	2	2	1	1	3	3	3	3	2	3	2	1