

GANPAT UNIVERSITY

FACULTY OF ENGINEERING & TECHNOLOGY	
-------------------------------------	--

Programme	Bachelor of Technology					Branch/Spec.	Computer Engineering/ Information Technology/ Computer Engineering (Artificial Intelligence)		
Semester	VII					Version	2.0.0.2		
Effective from Academic Year			2025-26			Effective for the Batch admitted in		July 2022	
Course Code		2CEIT702		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	-	1	-	4	Theory	40	60	100
Hours	3	-	2	-	5	Practical	30	20	50
Pre-requisites									
Programming concepts, Basic Mathematical concepts and statistics									
Course Outcomes									
On successful completion of the course, the students will be able to:									
CO1	Understand fundamental concepts of big data, big data applications and analytics in the real world.								
CO2	Apply different big data tools and technologies to effectively handle the big data.								
CO3	Analyze and visualize the real-world big data to extract meaningful insights for decision making.								
CO4	Design and implement the big data analytics algorithms to solve the data intensive problems.								
Theory Syllabus									
Unit	Content								Hrs.
1	Introduction: Big Data Definition, Types of Big Data, Application of Big Data, Big Data Drivers, Big Data Characteristics, Big Data Challenges, Concept of Big data analytics, Types of Big data analytics , Lifecycle of big data analytics , Concept of Distributed file systems								05
2	Big Data Enabling Technologies: Hadoop stack for Big Data: Hadoop distributed file system, Map reduce, YARN, Concept of Zookeeper								08
3	Big Data Storage Platforms: Concept of Large-Scale Data Storage, CAP Theorem, ACID and BASE Properties, Concept of Cassandra, Concept of HBase								08
4	Apache Spark: Concept of Data Streaming, Big data pipelines architecture, Concept of Apache Spark: Introduction, Philosophy of Spark, Spark APIs (RDD, Dataframes), Real time data processing using spark.								08
5	Apache Kafka: Introduction to Kafka, Kafka streaming ecosystem and applications								08
6	Advanced Topics in Big Data Analytics: Big data Machine learning, Graph processing using Spark GraphX, Introduction to big data algorithms like Bloom Filters, Count-min Sketch, Property Testing Model, Markov Chains and Random Walk								08
Practical Content									
Practicals, 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	Big Data Analytics: Introduction to Hadoop, Spark, and Machine-Learning by Raj Kamal and Preeti Saxena, McGrawhill								
2	Big Data, Black Book: Covers Hadoop 2, MapReduce, Hive, YARN, Pig, R and Data Visualization by DT Editorial Services, Wileyindia Publication								
3	Hadoop: The Definitive Guide by Tom White, O'Reilly Media Publication								
3	Taming The Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics by Bill Franks, Wiley								

4	Big Data Analytics by S Chandramouli, Asha A George, CR Rene Robin, D Doreen Hephzibah Miriam, J Jasmine Christina Magdalene, Orient BlackSwan Publication
5	Big Data Analytics with Hadoop 3 by Sridhar Alla, O'Reilly Publication
ICT/MOOCs Reference	
1	https://nptel.ac.in/courses/106104189
2	https://archive.nptel.ac.in/courses/106/106/106106142/

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	1	1	1	2	1	0	0	1	0	2	3	3	0
CO2	3	3	2	2	3	1	0	0	2	1	1	2	2	3	1
CO3	2	3	3	3	3	2	1	1	2	2	2	2	2	3	1
CO4	3	3	3	3	3	1	0	1	2	2	2	2	2	2	0