



Course Syllabus
Gyanmanjari Institute of Management Studies
Semester-4 (MBA)

Subject: Big Data Analytics – MBABA14514

Type of course: Major (Core)

Prerequisite:

Students should have knowledge of Database Management Systems, Data Warehousing, and Data Mining.

Rationale:

This course introduces students to the emerging field of Big Data Analytics, a crucial area for businesses seeking competitive advantage in the digital era. Students will learn the fundamentals of Big Data, including data generation, storage, analysis, and applications. The course emphasizes frameworks like Hadoop, data analytic methods, and practical applications in enterprises, IoT, healthcare, and social networks.

Teaching and Examination Scheme:

Teaching Scheme			Credits	Examination Marks					Total Marks
CI	T	P		C	Theory Marks		Practical Marks		
			ESE		MS E	V	P	ALA	
03	00	02	04	60	30	10	20	30	150

Legends: CI-Classroom Instructions; T – Tutorial; P - Practical; C – Credit; ESE – End Semester Examination; MSE- Mid Semester Examination; V – Viva; CA-Continuous Assessment; ALA- Active Learning Activities.

3 Credits * 25 Marks = 75 Marks (each credit carries 25 Marks) Theory

1 Credits * 25 Marks = 25 Marks (each credit carries 25 Marks) Practical

SEE 100 Marks will be converted in to 50 Marks

CCE 100 Marks will be converted in to 50 Marks

It is compulsory to pass in each individual component.



Course Content:

Sr. No	Course content	Hrs.	% Weightage
1	Introduction to Big Data, Generation and Acquisition Dawn of the Big Data Era, Definition and Features of Big Data, Big Data Value, Development of Big Data, Challenges of Big Data. Related Technologies: Cloud Computing, IoT, Data Center, Hadoop. Enterprise Data, IoT Data, Internet Data, Biomedical Data. Data Collection, Data Transportation, Data Pre-processing.	15	25
2	Big Data Storage and Hadoop Framework Storage Systems for Massive Data, Distributed Storage, Storage Mechanisms for Big Data (Database Technology, Design Factors, Programming Models). Hadoop & MapReduce: History, Hadoop Ecosystem, Data Analysis with Unix Tools, Java MapReduce, Scaling Out, Hadoop Streaming, Hadoop Pipes.	15	25
3	Big Data Analysis Traditional Data Analysis vs Big Data Analysis, Big Data Analytic Methods, Architecture for Big Data Analysis (Real-Time vs Offline, Different Levels, Different Complexity), Tools for Big Data Mining and Analysis.	15	25
4	Big Data Applications Structured, Text, Web, Multimedia, Network, and Mobile Traffic Analysis. Applications in Enterprises, IoT, Social Networks, Healthcare, Collective Intelligence, Smart Grid.	15	25

Sr. No	Practical's	Unit no	App hours
1	Study and analysis of Big Data characteristics (5Vs) using real-life datasets (Enterprise, IoT, social media, Biomedical).	1	2
2	Identify and classify different sources of Big Data (Enterprise Data, IoT Data, Internet Data, Biomedical)	1	2



	Data) and analyze data generation challenges.		
3	Perform data collection and preprocessing on a sample dataset: Data cleaning, Handling missing values, Data transformation	1	4
4	Study of Distributed Storage Systems and comparison of traditional databases vs Big Data storage systems (HDFS, NoSQL – conceptual).	2	4
5	Installation and configuration of Hadoop Single Node Cluster (or demonstration using cloud/virtual lab).	2	2
6	Hands-on practice with HDFS commands for file operations (upload, download, delete, permissions).	2	2
7	Implementation of a simple MapReduce program (Word Count) using Hadoop / MapReduce simulator.	2	4
8	Perform Big Data analysis using Unix/Linux tools (grep, awk, sort, wc) on large text datasets.	3	2
9	Comparative analysis of Traditional Data Analytics vs Big Data Analytics using a case study (batch vs real-time processing).	3	4
10	Case study analysis of Big Data applications in any one domain: Enterprise Analytics, IoT, social media, Healthcare, Smart Grid (Students prepare and present findings).	4	4
			30

Continuous Assessment:

Sr. No	Active Learning Activities	Marks
1	Hadoop Mini Project: Students will implement a small project using Hadoop/MapReduce ecosystem tools, document dataset, process, outputs, and insights, and upload a PDF project report on GMIU Web Portal.	10
2	Diagram Preparation: Students will create a diagram showing Big Data architecture (data flow from generation to analysis) and upload PDF report on GMIU web portal.	10
3	Report on Applications: Students will prepare a detailed report on the applications of Big Data Analytics in enterprises, IoT, healthcare, or other sectors (any two of their choice) covering benefits, challenges, and future scope, and upload a PDF report on GMIU Web Portal.	10



Suggested Specification table with Marks (Theory): 60

Distribution of Theory Marks (Revised Bloom's Taxonomy)						
Level	Remembrance (R)	Understanding (U)	Application (A)	Analyze (N)	Evaluate (E)	Create (C)
Weightage	30%	30%	10%	10%	10%	10%

Note: This specification table shall be treated as a general guideline for students and teachers. The actual distribution of marks in the question paper may vary slightly from above table.

Course Outcome:

After learning the course, the students should be able to:	
CO1	Understand the concept, features, and challenges of Big Data.
CO2	Apply Hadoop, MapReduce, and distributed storage for Big Data processing.
CO3	Analyze data using Big Data analytic methods and tools.
CO4	Evaluate applications of Big Data in enterprises, IoT, healthcare, and smart grids.

Instructional Method:

The course delivery method will depend upon the requirement of content and the needs of students. The teacher, in addition to conventional teaching methods by black board, may also use any tools such as demonstration, role play, Quiz, brainstorming, MOOCs etc.

From the content 10% topics are suggested for flipped mode instruction. Students will use supplementary resources such as online videos, NPTEL/SWAYAM videos, e-courses. The internal evaluation will be done on the basis of Active Learning Assignment.

Practical/Viva examination will be conducted at the end of semester for evaluation of performance of students in the laboratory.

Reference Books:

- [1.] Min Chen, Shiwen Mao, Yin Zhang, Victor C.M. Leung, Big Data: Related Technologies, Challenges and Future Prospects, Springer, 2014.
- [2.] Tom White, Hadoop: The Definitive Guide, O'Reilly, 2nd Edition.
- [3.] Vignesh Prajapati, Big Data Analytics with R and Hadoop, PACKT Publishing, 2013.

