

Branch Name:	Management (Business Analytics)
Program Code:	MS301
Course Name:	Big Data Analytics
Course Code:	2MS3010645T
Pre-requisite Course:	Basic of research, Computer application in Management

Course Objectives:

- 1) To optimize business decisions and create competitive advantage with Big Data analytics
- 2) To explore the fundamental concepts of big data analytics.
- 3) To learn to analyses the big data using intelligent techniques.
- 4) To understand the various search methods and visualization techniques.
- 5) To learn to use various techniques for mining data stream.
- 6) To understand the applications using Map Reduce Concepts.
- 7) To introduce programming tools PIG & HIVE in Hadoop echo system.

Teaching and Examination Scheme:

Teaching Scheme (Hours per week)				Evaluation Scheme (Marks)				
Lecture (L)	Tutorial (T)	Practical (P)	Credit	Theory (Marks)		Practical (Marks)		Total (Marks)
				University Assessment	Continuous Assessment	University Assessment	Continuous Assessment	
4	-	-	4	60	40	-	-	100

Subject Contents:

Sr No.	Topic	Total Hours	Weightage (%)
1	Introduction to big data: Introduction to Big Data Platform – Challenges of Conventional Systems - Intelligent Data Analysis – Nature of Data - Analytic Processes and Tools - Analysis vs Reporting.	10	20
2	Mining data streams: Introduction To Streams Concepts – Stream Data Model and Architecture -Stream Computing - Sampling Data in a Stream – Filtering Streams - Counting Distinct Elements in a Stream – Estimating Moments – Counting Oneness in a Window – Decaying Window - Real-time Analytics Platform (RTAP) Applications – Case Studies - Real-Time Sentiment Analysis- Stock Market Predictions	10	20
3	Hadoop: 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 - Hadoop environment.	10	20
4	Frameworks: Applications on Big Data Using Pig and Hive – Data processing operators in Pig – Hive services – HiveQL – Querying Data in Hive - fundamentals of HBase and ZooKeeper - IBM InfoSphere BigInsights and Streams.	10	20

5	Predictive Analytics- Simple linear regression- Multiple linear regression- Interpretation of regression coefficients. Visualizations - Visual data analysis techniques- interaction techniques - Systems and applications.	08	20
---	--	----	----

Text Books:

Sr. No.	Books
1	Michael Berthold, David J. Hand, “Intelligent Data Analysis”, Springer, 2007.
2	Tom White “Hadoop: The Definitive Guide” Third Edition, O’reilly Media, 2012.
3	Chris Eaton, Dirk DeRoos, Tom Deutsch, George Lapis, Paul Zikopoulos, “Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data”, McGrawHill Publishing, 2012.

Reference Books:

Sr. No.	Books
1	Anand Rajaraman and Jeffrey David Ullman, “Mining of Massive Datasets”, CUP, 2012.
2	Bill Franks, “Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics”, John Wiley & sons, 2012.
3	Glenn J. Myatt, “Making Sense of Data”, John Wiley & Sons, 2007.

Course Learning Outcomes (CLO): On completion of this course, the students will be able to:

CLO	Description	Bloom’s Taxonomy Level
CLO1	Understand the fundamentals of Big Data, its challenges, and the role of analytics in decision-making.	Understand
CLO2	Apply data stream mining techniques for real-time analytics and business intelligence applications.	Apply
CLO3	Analyze and implement Hadoop Distributed File System (HDFS) and MapReduce for big data processing.	Analyse
CLO4	Utilize big data frameworks such as Pig, Hive, HBase, and ZooKeeper for data storage and querying.	Apply
CLO5	Develop predictive analytics models using regression techniques for data-driven decision-making.	Create
CLO6	Interpret and visualize big data insights using advanced data visualization techniques and tools.	Evaluate

Mapping of CLOs with Pos & PSOs

Course Learning Outcomes	Program Outcomes (POs)								Program Specific Outcomes (PSOs)	
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PSO1	PSO2
CLO1	H	H					H		M	H
CLO2		M	H					H	M	
CLO3	M		M			H			H	
CLO4		H	H		M				M	
CLO5	H	M		H						H
CLO6			H	H				H		M

H: High, M: Medium, L: Low
