St. MIRA'S COLLEGE FOR GIRLS * POONA-411001 *

Computer Science Paper V
Big Data Analytics
[Elective Course]

Semester: III	Credits: 4	Subject Code: MSE32005	Lectures: 48

Course Outcomes:

At the end of this course, the learner will be able to:

- Describe and explain the Big Data Analytics, its methodology and its applications.
- Examine and assess the concepts of Hadoop Distributed File System
- Relate, Explain and examine the case studies related to real life situations using Big data Analytics
- Analyse the role of map-reduce and functional programming

Unit 1: Introduction to Big Data and Emerging Database Landscape	8
• **Chapter 1: Introduction to Big data** ○ Big Data: Definition & taxonomy ○ Sources of Big Data ○ Characteristics of Big Data-5V's of Big Data (Volume, Variety, Velocity, Veracity, Value) ○ The Structure of Big Data-Structured dada, Unstructured data, Semi-structured data ○ Why Big Data, Applications of Big Data ○ Risks of Big data ○ Big Data Implications for Industries - Big Data Analytics for Telecom, Banking, Insurance, Retail, HealthCare, IT Operations	6
• **Chapter 2: Emerging Database Landscape** ○ The Database Evolution ○ The Scale-Out Architecture, RDBMS Vs Non-Relational Database, Old SQL, NewSQL and the Emerging NoSQL ○ Database Workloads & Workload Characteristics ○ Implications of Big Data Scale on Data Processing	2

Unit 2: Application Architecture & Data Modeling for Big Data and The Hadoop Ecosystem	15
• **Chapter 3:Application Architecture & Data Modeling for Big Data And Analytics** ○ Big Data Warehouse & Analytics ○ Big data Warehouse System requirements & Hybrid Architectures ○ Enterprise Data Platform Ecosystem ○ Big Data and Master Data Management ○ Understanding data integration Pattern	6

Board of Studies	Member Name	Signature
Chairperson	Mrs. Ashwini Kulkarni	

○ Big Data Workload Design Approaches ○ Map-Reduce patterns, Algorithms and Use Cases • **Chapter 4: The Hadoop Ecosystem** ○ Introduction to Hadoop Ecosystem ○ History of Hadoop-Facebook, Dynamo, Yahoo, Google ○ Hadoop Architecture ○ Hadoop Ecosystem Components: HDFS, Map-reduce, YARN, Pig, Hive, HBase, Mahout	**9**

Unit 3: Extracting Value from Big Data and Big Data Analytics Methodology	**9**
• **Chapter 5:Extracting Value from Big Data** ○ Real Time Analytics ○ In-Memory Data Grid for real Time Analysis ○ Map reduce & Real Time Processing, Use Cases, document shingling, min hash function	**4**
• **Chapter 6:Big Data Analytics Methodology** ○ Big Data Analytics Methodology-Analyze & evaluate business cases ○ Develop Business Hypothesis –Analyze outcomes, Build & Prepare Data Sets, Select & Build Analytical Model, Design for Big Data scale. Build production ready system, setting up the Big Data Analytics system, Gathering data, measure & monitor	**5**

Unit 4: Big Data Analytics Practical	**16**
Assignments: • Assignment 1: Case studies • Assignment 2: Navigating in Hadoop environment • Understanding HDFS • Unix commands using Big Data • Development in Hadoop environment, using various Hadoop tools/utilities • Develop map-reduce programs	

***Contact hours=12 hours for Library work, practical or field work or research purposes**

Board of Studies	**Member Name**	**Signature**
Chairperson	Mrs. Ashwini Kulkarni	

Reference Books:

- Bernard Marr,,*Big Data Case Study,* Willey Publications.
- Bernard Marr, "*Big data in practice – How 45 successful companies used big data analytics to deliver extraordinary results*", Wiley, (2016).
- Cristian Molaro, Surekha Parekh, Terry Purcell, "*DB2 11: The Database for Big Data &Analytics*", MC Press, (2013)
- Frank J. Ohlhorst, "*Big Data Analytics: Turning Big Data into Big Money*", Wiley Publishers (2012)
- Madhu Jagdeesh, Soumendra Mohanty, Harsha Srivatsa, "*Big Data Imperatives: Enterprise Big Data Warehouse, BI Implementations and Analytics*",1stEdition, Apress, (2013)
- Tom White, "*Hadoop-The Definitive Guide, Storage and analysis at internet scale*", SPD, O'Really.

Board of Studies	Member Name	Signature(in white cell)	
Chairperson	Mrs. Ashwini Kulkarni	25/7/20	
Faculty	Mrs. Shubhangi Jagtap		25/07/20
Faculty	Mrs. Smita Borkar	25/7/20	
Faculty	Mrs. Swati pulate		25/7/20
Faculty	Mrs. Alka Kalhapure	25/07/20	
Subject Expert (Outside SPPU)	Dr. Manisha Divate		25/7/20
Subject Expert (Outside SPPU)	Prof Aniket Nagane	25/7/20	
VC Nominee	Dr. Manisha Bharambe		25/07/20
Industry Expert	Ms. SnehalBiyala	25/7/20	
Alumni	Ms. Mamta Choudhary		25/7/20

Board of Studies	Member Name	Signature
Chairperson	Mrs. Ashwini Kulkarni	