Computer Science Paper V
Big Data Analytics
[Elective Course]

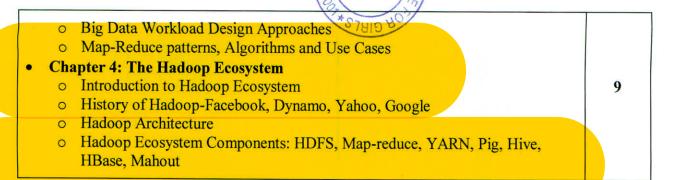| Semester: III	Credits: 4	Subject Code: MSE32005	Lectures: 48

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**
  o Big Data: Definition & taxonomy
  o Sources of Big Data
  o Characteristics of Big Data-5V's of Big Data (Volume, Variety, Velocity, Veracity, Value)
  o The Structure of Big Data-Structured dada, Unstructured data, Semi-structured data
  o Why Big Data, Applications of Big Data
  o Risks of Big data
  o Big Data Implications for Industries - Big Data Analytics for Telecom, Banking, Insurance, Retail, HealthCare, IT Operations
• **Chapter 2: Emerging Database Landscape**
  o The Database Evolution
  o The Scale-Out Architecture, RDBMS Vs Non-Relational Database, Old SQL, NewSQL and the Emerging NoSQL
  o Database Workloads & Workload Characteristics
  o Implications of Big Data Scale on Data Processing
	6
	
	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**
  o Big Data Warehouse & Analytics
  o Big data Warehouse System requirements & Hybrid Architectures
  o Enterprise Data Platform Ecosystem
  o Big Data and Master Data Management
  o Understanding data integration Pattern
	6

Board of Studies	Member Name	Signature
Chairperson	Mrs. Ashwini Kulkarni	

o Big Data Workload Design Approaches
    o Map-Reduce patterns, Algorithms and Use Cases
• **Chapter 4: The Hadoop Ecosystem**
    o Introduction to Hadoop Ecosystem
    o History of Hadoop-Facebook, Dynamo, Yahoo, Google
    o Hadoop Architecture
    o 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**
    o Real Time Analytics
    o In-Memory Data Grid for real Time Analysis
    o Map reduce & Real Time Processing, Use Cases, document shingling, min hash function
	4

• **Chapter 6:Big Data Analytics Methodology**
    o Big Data Analytics Methodology-Analyze & evaluate business cases
    o 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		*Shubhangi 25/07/20*
Faculty	Mrs. Smita Borkar	*25/7/20*	
Faculty	Mrs. Swati pulate		*25/7/20*
Faculty	Mrs. Alka Kalhapure	*Alka 25/07/20*	
Subject Expert (Outside SPPU)	Dr. Manisha Divate		*Divate 25/7/20*
Subject Expert (Outside SPPU)	Prof Aniket Nagane	*25/7/20*	
VC Nominee	Dr. Manisha Bharambe		*MBharambe 25/07/20*
Industry Expert	Ms. SnehalBiyala	*Kulkarni 25/7/20*	
Alumni	Ms. Mamta Choudhary		*25/7/20*

Board of Studies	Member Name	Signature
Chairperson	Mrs. Ashwini Kulkarni	