BI	G DATA ANA	LYTICS	
[As per Choice Based Credit Syst	em (CBCS) sc	heme] (Effective from t	he academic year
2016	-2017) SEMES	STER – VIII	20
Subject Code	150.582	IA Marks	20
Number of Lecture Hours/Week	4	Exam Marks	80
Total Number of Lecture Hours	50	Exam Hours	03
	CREDITS –	04	
Course objectives: This course will	l enable student	is to	
Understand Hadoop Distribution	ited File system	and examine MapReduc	e Programming
• Explore Hadoop tools and m	anage Hadoop	with Ambari	
• Appraise the role of Busines	s intelligence a	nd its applications across	industries
Assess core data mining tech	iniques for data	analytics	
 Identify various Text Mining 	g techniques		
Module – 1			Teaching Hours
Hadoop Distributed File System I	Basics, Runnin	g Example Programs a	nd 10 Hours
Benchmarks, Hadoop MapReduce F	ramework, Ma	pReduce Programming	
Module – 2			
Essential Hadoop Tools, Hadoop YA	ARN Applicatio	ons, Managing Hadoop w	vith 10 Hours
Apache Ambari, Basic Hadoop Adn	ninistration Pro-	cedures	
Module – 3			
Business Intelligence Concepts an	d Application,	Data Warehousing, Da	ita 10 Hours
Mining, Data Visualization			
Module – 4	• 1 > 7 1 > 7		·
Decision Trees, Regression, Artifi	cial Neural Ne	etworks, Cluster Analys	18, I0 Hours
Association Rule Mining			
Module – 5 Tract Mining Nation Descent Anglania	Constant Marth		. 10 II
Social Network Analysis	, Support Vecto	or Machines, web Minin	g, I Hours
Course outcomes: The students sho	ould be able to:		
 Master the concepts of HDF 	S and MapRed	uce framework	
 Investigate Hadoon related to 	ools for Big Da	ta Analytics and perform	basic Hadoop
Administration		······································	ourie manoch
• Recognize the role of Busine	ess Intelligence,	Data warehousing and V	Visualization in
decision making			
• Infer the importance of core	data mining tec	hniques for data analytic	S
Compare and contrast difference	ent Text Mining	g Techniques	
Question paper pattern:			
The question paper will have ten qu	estions.		
There will be 2 questions from each	module.		
Each question will have questions co	overing all the t	opics under a module.	
The students will have to answer 5 f	ull questions, s	electing one full question	1 from
each module.			
Text Books:			diala of Di D
1. Douglas Eadline, "Hadoop 2	Quick-Start	Guide: Learn the Esser	mais of Big Data
Computing in the Apache	Hadoop 2 Ec	osystem", 1 ^{°°} Edition, P	earson Education,

2016. ISBN-13: 978-9332570351

2. Anil Maheshwari, "Data Analytics", 1st Edition, McGraw Hill Education, 2017. ISBN-13: 978-9352604180

Reference Books:

- 1) Tom White, "Hadoop: The Definitive Guide", 4th Edition, O'Reilly Media, 2015.ISBN-13: 978-9352130672
- 2) Boris Lublinsky, Kevin T.Smith, Alexey Yakubovich,"Professional Hadoop Solutions", 1stEdition, Wrox Press, 2014ISBN-13: 978-8126551071 3) Eric Sammer, "Hadoop Operations: A Guide for Developers and
- Administrators", 1stEdition, O'Reilly Media, 2012.ISBN-13: 978-9350239261