

BIG DATA AND BUSINESS ANALYTICS

VII Semester: CSE/IT								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
ACS012	Core	L	T	P	C	CIA	SEE	Total
		3	1	-	4	30	70	100
Contact Classes: 45	Tutorial Classes: 15	Practical Classes: Nil			Total Classes: 60			
<p>COURSE OBJECTIVES: Students will try to learn:</p> <ol style="list-style-type: none"> I. The implications and challenges of Big Data Analytics. II. The technologies used to store, manage, and analyze big data in a Hadoop ecosystem. III. The techniques and principles in big data analytics with scalability and streaming capability. IV. The hypothesis on the optimized business decisions in solving complex real-world problems. <p>COURSE OUTCOMES (COs): After successful completion of this course, students will be able to:</p> <ol style="list-style-type: none"> 1. Describe the concepts and challenges of Big Data comparing with traditional business intelligence. 2. Recognize the importance and challenges involved in big data analytics in view of technology landscape. 3. Illustrate data processing methods with the help of MapReduce abstraction techniques. 4. Explain the processing and querying techniques included in big data technologies such as Hadoop, MapReduce, Pig and Hive. 5. Make use of appropriate components for processing, scheduling and knowledge extraction from large volumes in Hadoop Ecosystem. 6. Translate the data from traditional file system to HDFS for analyzing big data in Hadoop ecosystem. 7. Develop a Map Reduce application for optimizing the jobs. 8. Develop applications for handling huge volume of data using Pig Latin. 9. Demonstrate the use of data analytics using Hadoop component HIVE with its built-in functions and services like DDL. 10. Develop business models with the help of big data analytics. 11. Make use of appropriate Hadoop components for developing real time big data analytics in various applications like recommender systems, social media applications etc. 								
UNIT-I	INTRODUCTION TO BIG DATA						Classes: 09	
<p>Introduction to Big data: Characteristics of Data, Evolution of Big Data, Definition of Big Data, Challenges with Big Data, Traditional Business Intelligence (BI) versus Big Data. Big data analytics: Classification of Analytics, Importance and challenges facing big data, Terminologies Used in Big Data Environments, The Big Data Technology Landscape.</p>								

UNIT -II	INTRODUCTION TO HADOOP	Classes: 09
Introducing Hadoop,RDBMS versus Hadoop,Distributed Computing Challenges, History and overview of Hadoop, Use Case of Hadoop,Hadoop Distributors,Processing Data with Hadoop, Interacting with Hadoop Ecosystem		
UNIT-III	THE HADOOP DISTRIBUTED FILESYSTEM	Classes: 09
Hadoop Distributed File System(HDFS):The Design of HDFS, HDFS Concepts, Basic Filesystem Operations, Hadoop Filesystems. The Java Interface- Reading Data from a Hadoop URL, Reading Data Using the Filesystem API, Writing Data. Data Flow- Anatomy of a File Read, Anatomy of a File Write, Limitations.		
UNIT-IV	UNDERSTANDING MAP REDUCE FUNDAMENTALS	Classes: 09
Map Reduce Framework: Exploring the features of Map Reduce, Working of Map Reduce, Exploring Map and Reduce Functions, Techniques to optimize Map Reduce jobs, Uses of Map Reduce. Controlling MapReduce Execution with InputFormat, Reading Data with custom RecordReader,-Reader, Writer, Combiner, Partitioners, Map Reduce Phases,Developing simple MapReduce Application.		
UNIT-V	INTRODUCTION TO PIG and HIVE	Classes: 09
Introducing Pig: Pig architecture, Benefits, Installing Pig, Properties of Pig, Running Pig, Getting started with Pig Latin, Working with operators in Pig, Working with functions in Pig. Introducing Hive: Getting started with Hive, Hive Services, Data types in Hive, Built-in functions in Hive, Hive DDL.		
Text Books:		
<ol style="list-style-type: none"> 1. Seema Acharya, Subhashini Chellappan, -Big Data and Analytics, Wiley Publications, 2nd Edition, 2014 2. Tom White, -Hadoop: The Definitive Guide, O'Reilly, 3rd Edition, 2012. 3. Big Data Black Book, dreamtech publications, 1st Edition, 2017. 		
Reference Books:		
<ol style="list-style-type: none"> 1. Michael Minelli, Michele Chambers, Ambiga Dhiraj, -Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business, Wiley CIO Series, 1st Edition, 2013. 2. Rajiv Sabherwal, Irma Becerra- Fernandez, -Business Intelligence -Practice, Technologies and Management, John Wiley, 1st Edition, 2011. 3. Arvind Sathi, -Big Data Analytics: Disruptive Technologies for Changing the Game, IBM Corporation, 1st Edition, 2012. 		
Web References:		
<ol style="list-style-type: none"> 1. https://www.sas.com/en_us/insights/analytics/big-data-analytics.html 2. https://www.searchbusinessanalytics.techtarget.com/definition/big-data-analytics 3. https://www.webopedia.com 		
E-Text Books:		
<ol style="list-style-type: none"> 1. https://www.books.google.co.in/books?id=rkWPojgfeM8C&printsec=frontcover&dq=HIGH+PERFORMANCE+COMPUTING. 2. http://www.datameer.com/pdf/big-data-analytics-ebook.pdf?mkt_tok. 		