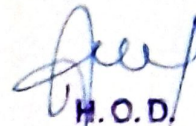


<b>BIG DATA ANALYTICS</b> <b>[As per Choice Based Credit System (CBCS) scheme]</b> <b>(Effective from the academic year 2016 -2017)</b> <b>SEMESTER – VIII</b>			
Subject Code	15CS82	IA Marks	20
Number of Lecture Hours/Week	4	Exam Marks	80
Total Number of Lecture Hours	50	Exam Hours	03
<b>CREDITS – 04</b>			
<b>Course objectives:</b> This course will enable students to			
<ul style="list-style-type: none"> <li>• Understand Hadoop Distributed File system and examine MapReduce Programming</li> <li>• Explore Hadoop tools and manage Hadoop with Ambari</li> <li>• Appraise the role of Business intelligence and its applications across industries</li> <li>• Assess core data mining techniques for data analytics</li> <li>• Identify various Text Mining techniques</li> </ul>			
<b>Module – 1</b>			<b>Teaching Hours</b>
Hadoop Distributed File System Basics, Running Example Programs and Benchmarks, Hadoop MapReduce Framework, MapReduce Programming			<b>10 Hours</b>
<b>Module – 2</b>			
Essential Hadoop Tools, Hadoop YARN Applications, Managing Hadoop with Apache Ambari, Basic Hadoop Administration Procedures			<b>10 Hours</b>
<b>Module – 3</b>			
Business Intelligence Concepts and Application, Data Warehousing, Data Mining, Data Visualization			<b>10 Hours</b>
<b>Module – 4</b>			
Decision Trees, Regression, Artificial Neural Networks, Cluster Analysis, Association Rule Mining			<b>10 Hours</b>
<b>Module – 5</b>			
Text Mining, Naïve-Bayes Analysis, Support Vector Machines, Web Mining, Social Network Analysis			<b>10 Hours</b>
<b>Course outcomes:</b> The students should be able to:			
<ul style="list-style-type: none"> <li>• Master the concepts of HDFS and MapReduce framework</li> <li>• Investigate Hadoop related tools for Big Data Analytics and perform basic Hadoop Administration</li> <li>• Recognize the role of Business Intelligence, Data warehousing and Visualization in decision making</li> <li>• Infer the importance of core data mining techniques for data analytics</li> <li>• Compare and contrast different Text Mining Techniques</li> </ul>			
<b>Question paper pattern:</b> The question paper will have ten questions. There will be 2 questions from each module. Each question will have questions covering all the topics under a module. The students will have to answer 5 full questions, selecting one full question from each module.			
<b>Text Books:</b> 1. Douglas Eadline, "Hadoop 2 Quick-Start Guide: Learn the Essentials of Big Data Computing in the Apache Hadoop 2 Ecosystem", 1 <sup>st</sup> Edition, Pearson Education, 2016. ISBN-13: 978-9332570351			

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

**Reference Books:**

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



Dept. Of Computer Science & Engineering  
Alva's Institute of Engg. & Technology  
Mijar, MOODBIDRI - 574 225