

## BIG DATA AND BUSINESS ANALYTICS LABORATORY

<b>VII Semester: CSE / IT</b>								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
ACS111	Core	L	T	P	C	CIA	SEE	Total
		-	-	3	2	30	70	100
<b>Contact Classes: Nil</b>		<b>Tutorial Classes: Nil</b>		<b>Practical Classes: 36</b>			<b>Total Classes: 36</b>	
<p><b>COURSE OBJECTIVES:</b>  <b>The course should enable the students to:</b>                      I. Optimize business decisions and create competitive advantage with Big data analytics                      II. Practice java concepts required for developing map reduce programs.                      III. Impart the architectural concepts of Hadoop and introducing map reduce paradigm.                      IV. Practice programming tools PIG and HIVE in Hadoop eco system.                      V. Implement best practices for Hadoop development.</p> <p><b>COURSE LEARNING OUTCOMES (CLOs):</b>                      1. Understand the installation of VMWare.                      2. Understand and apply the Perform setting up and Installing Hadoop in its three operating modes.                      3. Implementing the basic commands of LINUX Operating System.                      4. Implement the file management tasks in Hadoop.                      5. Understand Map Reduce Paradigm.                      6. Apply Map Reduce program that mines weather data.                      7. Implement matrix multiplication with Hadoop MapReduce                      8. Apply Map Reduce program that makes the dataset to be compressed.                      9. Understand the installation of PIG.                      10. Understand Pig Latin scripts sort, group, join, project, and filter your data.                      11. Implement the Pig Latin scripts in two different modes                      12. Understand the installation of HIVE                      13. Apply Hive to create, alter, and drop databases, tables, views, functions, and indexes.</p>								
<b>Week-1</b>	<b>INSTALL VMWARE</b>							
Installation of VMWare to setup the Hadoop environment and its ecosystems.								
<b>Week-2</b>	<b>HADOOP MODES</b>							
a. Perform setting up and Installing Hadoop in its three operating modes. i. Standalone. ii. Pseudo distributed. iii. Fully distributed. b. Use web based tools to monitor your Hadoop setup.								
<b>Week-3</b>	<b>USING LINUX OPERATING SYSTEM</b>							
Implementing the basic commands of LINUX Operating System – File/Directory creation, deletion, update operations.								
<b>Week-4</b>	<b>FILE MANAGEMENT IN HADOOP</b>							
a. Implement the following file management tasks in Hadoop: i. Adding files and directories ii. Retrieving files iii. Deleting files  <b>Hint:</b> A typical Hadoop workflow creates data files (such as log files) elsewhere and copies them into HDFS using one of the above command line utilities.								
<b>Week-5</b>	<b>MAPREDUCE PROGRAM 1</b>							
Run a basic word count Map Reduce program to understand Map Reduce Paradigm.								

<b>Week-6</b>	<b>MAPREDUCE PROGRAM 2</b>
Write a Map Reduce program that mines weather data. Hint: Weather sensors collecting data every hour at many locations across the globe gather a large volume of log data, which is a good candidate for analysis with Map Reduce, since it is semi structured and record-oriented.	
<b>Week-7</b>	<b>MAPREDUCE PROGRAM 3</b>
Implement matrix multiplication with Hadoop Map Reduce.	
<b>Week-8</b>	<b>PIG LATIN LANGUAGE - PIG</b>
Installation of PIG.	
<b>Week-9</b>	<b>PIG COMMANDS</b>
Write Pig Latin scripts sort, group, join, project, and filter your data.	
<b>Week-10</b>	<b>PIG LATIN MODES, PROGRAMS</b>
<ul style="list-style-type: none"> <li>a. Run the Pig Latin Scripts to find Word Count</li> <li>b. Run the Pig Latin Scripts to find a max temp for each and every year.</li> </ul>	
<b>Week-11</b>	<b>HIVE</b>
Installation of HIVE.	
<b>Week-12</b>	<b>HIVE OPERATIONS</b>
Use Hive to create, alter, and drop databases, tables, views, functions, and indexes.	
<b>Reference Books:</b>	
1. Jay Liebowitz, —Big Data And Business Analytics Laboratory, CRC Press.	