



# INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal, Hyderabad -500 043

## COMPUTER SCIENCE AND ENGINEERING

### COURSE DESCRIPTOR

Course Title	<b>BIG DATA AND BUSINESS ANALYTICS LABORATORY</b>				
Course Code	<b>ACS111</b>				
Programme	<b>B.Tech</b>				
Semester	<b>VII</b>	<b>CSE/IT</b>			
Course Type	<b>Core</b>				
Regulation	<b>IARE - R16</b>				
Course Structure	<b>Theory</b>			<b>Practical</b>	
	<b>Lectures</b>	<b>Tutorials</b>	<b>Credits</b>	<b>Laboratory</b>	<b>Credits</b>
	-	-	-	3	2
Chief Coordinator	<b>Ms. S Swarajya Laxmi, Assistant professor</b>				
Course Faculty	<b>Ms. E Uma Shankari, Assistant Professor Ms. G Sulakshana, Assistant professor Ms. G Srilekha, Assistant Professor</b>				

#### I. COURSE OVERVIEW:

Big data and Business Analytics Laboratory demonstrates distributed computing environment .It includes hands on experience installation process of VMWare, Pig, Hive, program setting in three operating modes like standalone, Pseudo code distributed, fully distributed, implementing basic LINUX commands, HDFS file management, MapReduce functions, Pig commands and Hive operations.

#### II. COURSE PRE-REQUISITES:

Level	Course Code	Semester	Prerequisites	Credits
UG	ACS104	III	Database Management Systems Lab	2

#### III. MARKS DISTRIBUTION:

Subject	SEE Examination	CIA Examination	Total Marks
Big Data and Business Analytics Laboratory	70 Marks	30 Marks	100

#### IV. DELIVERY / INSTRUCTIONAL METHODOLOGIES:

✗	Chalk & Talk	✗	Quiz	✗	Assignments	✗	MOOCs
✓	LCD / PPT	✗	Seminars	✓	Mini Project	✓	Videos
✓	Open Ended Experiments						

#### V. EVALUATION METHODOLOGY:

Each laboratory will be evaluated for a total of 100 marks consisting of 30 marks for internal assessment and 70 marks for semester end lab examination. Out of 30 marks of internal assessment, continuous lab assessment will be done for 20 marks for the day to day performance and 10 marks for the final internal lab assessment.

**Semester End Examination (SEE):** The semester end lab examination for 70 marks shall be conducted by two examiners, one of them being Internal Examiner and the other being External Examiner, both nominated by the Principal from the panel of experts recommended by Chairman, BOS.

The emphasis on the experiments is broadly based on the following criteria:

20 %	To test the preparedness for the experiment.
20 %	To test the performance in the laboratory.
20 %	To test the calculations and graphs related to the concern experiment.
20 %	To test the results and the error analysis of the experiment.
20 %	To test the subject knowledge through viva – voice.

#### Continuous Internal Assessment (CIA):

CIA is conducted for a total of 30 marks (Table 1), with 20 marks for continuous lab assessment during day to day performance, 10 marks for final internal lab assessment.

Table 1: Assessment pattern for CIA

Component	Laboratory		Total Marks
	Day to day performance	Final internal lab assessment	
CIA Marks	20	10	30

#### Continuous Internal Examination (CIE):

One CIE exams shall be conducted at the end of the 15<sup>th</sup> week of the semester. The CIE exam is conducted for 10 marks of 3 hours duration.

Preparation	Performance	Algorithm	Results and Error Analysis	Viva	Total
2	2	2	2	2	10

#### VI. HOW PROGRAM OUTCOMES ARE ASSESSED:

Program Outcomes (POs)		Strength	Proficiency assessed by
PO 1	<b>Engineering knowledge:</b> Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.	1	Experiments and Viva voice

Program Outcomes (POs)		Strength	Proficiency assessed by
PO 2	<b>Problem analysis:</b> Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.	2	Experiments and Viva voice
PO 3	<b>Design/development of solutions:</b> Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.	3	Mini Project
PO 5	<b>Modern tool usage:</b> Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.	2	Mini Project

3 = High; 2 = Medium; 1 = Low

#### VII. HOW PROGRAM SPECIFIC OUTCOMES ARE ASSESSED:

Program Specific Outcomes (PSOs)		Strength	Proficiency assessed by
PSO 1	<b>Professional Skills:</b> The ability to understand, analyze and develop computer programs in the areas related to algorithms, system software, multimedia, web design, big data analytics, and networking for efficient analysis and design of computer - based systems of varying complexity.	2	Videos
PSO 2	<b>Problem solving skills:</b> The ability to apply standard practices and strategies in software project development using open-ended programming environments to deliver a quality product for business success.	1	Mini Project
PSO 3	<b>Successful career and Entrepreneurship:</b> The ability to employ modern computer languages, environments, and platforms in creating innovative career paths to be an entrepreneur, and a zest for higher studies.	-	-

3 = High; 2 = Medium; 1 = Low

#### VIII. COURSE OBJECTIVES (COs):

The course should enable the students to:	
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.

#### IX. COURSE LEARNING OUTCOMES (CLOs):

CLO Code	CLO's	At the end of the course, the student will have the ability to:	PO's Mapped	Strength of Mapping
ACS111.01	CLO 1	Understand the installation of VMWare	PO1, PO3	2
ACS111.02	CLO 2	Understand and apply the Perform setting up and Installing Hadoop in its three operating modes.	PO1, PO3	3
ACS111.03	CLO 3	Implementing the basic commands of LINUX Operating System	PO2	2
ACS111.04	CLO 4	Implement the file management tasks in Hadoop.	PO2, PO3, PO5	3
ACS111.05	CLO 5	Understand Map Reduce Paradigm.	PO2	2

CLO Code	CLO's	At the end of the course, the student will have the ability to:	PO's Mapped	Strength of Mapping
ACS111.06	CLO 6	Apply Map Reduce program that mines weather data.	PO3, PO5	3
ACS111.07	CLO 7	Implement matrix multiplication with Hadoop Map Reduce	PO2, PO5	2
ACS111.08	CLO 8	Apply Map Reduce program that makes the dataset to be compressed.	PO1, PO2	2
ACS111.09	CLO 9	Understand the installation of PIG.	PO1	2
ACS111.10	CLO 10	Understand Pig Latin scripts sort, group, join, project, and filter your data.	PO1, PO2	2
ACS111.11	CLO 11	Implement the Pig Latin scripts in two different modes	PO2	2
ACS111.12	CLO 12	Understand the installation of HIVE	PO1, PO5	2
ACS111.13	CLO 13	Apply Hive to create, alter, and drop databases, tables, views, functions, and indexes.	PO5	3

**3 = High; 2 = Medium; 1 = Low**

**X. MAPPING COURSE LEARNING OUTCOMES LEADING TO THE ACHIEVEMENT OF PROGRAM OUTCOMES AND PROGRAM SPECIFIC OUTCOMES:**

Course Learning Outcomes (CLOs)	Program Outcomes (POs)												Program Specific Outcomes (PSOs)		
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CLO 1	1		3										2		
CLO 2	2		3										2	1	
CLO 3		2													
CLO 4		2	2		2								2	2	
CLO 5		2													
CLO 6			2		3								1	1	
CLO 7		2			1								1	2	
CLO 8	2	1											2	1	
CLO 9	2														
CLO 10	1	3											1		
CLO 11		2													
CLO 12	1				3								1	1	
CLO 13					3								2	1	

**3 = High; 2 = Medium; 1 = Low**

### XI. ASSESSMENT METHODOLOGIES – DIRECT

CIE Exams	PO1, PO2 PO3, PO5	SEE Exams	PO1, PO2 PO3, PO5	Assignments	-	Seminars	-
Laboratory Practices	PO1, PO2 PO3, PO5	Student Viva	PO1,PO2	Mini Project	PO3,PO5	Certification	-

### XII. ASSESSMENT METHODOLOGIES - INDIRECT

✓	Early Semester Feedback	✓	End Semester OBE Feedback
✗	Assessment of Mini Projects by Experts		

### XIII. SYLLABUS

<b>LIST OF EXPERIMENTS</b>	
<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. <ul style="list-style-type: none"> <li>i. Standalone.</li> <li>ii. Pseudo distributed.</li> <li>iii. Fully distributed.</li> </ul> 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: <ul style="list-style-type: none"> <li>i. Adding files and directories</li> <li>ii. Retrieving files</li> <li>iii. Deleting files</li> </ul> Hint: 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>
a. Run the Pig Latin Scripts to find Word Count. b. Run the Pig Latin Scripts to find a max temp for each and every year.	

<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>Text Books:</b>	
<ol style="list-style-type: none"> <li>1. Michael Minelli, Michele Chambers, Ambiga Dhiraj, “Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today’s Business”, Wiley CIO Series, 1<sup>st</sup> Edition, 2013.</li> <li>2. Tom White, “Hadoop: The Definitive Guide”, O’Reilly, 3<sup>rd</sup> Edition, 2012.</li> <li>3. Rajiv Sabherwal, Irma Becerra- Fernandez, “Business Intelligence –Practice, Technologies and Management”, John Wiley, 1<sup>st</sup> Edition, 2011.</li> </ol>	
<b>Reference Books:</b>	
<ol style="list-style-type: none"> <li>1. Jay Liebowitz, “Big Data and Business Analytics Laboratory”, CRC Press.</li> </ol>	

#### XIV. COURSE PLAN:

The course plan is meant as a guideline. Probably there may be changes.

Week No.	Topics to be covered	Course Learning Outcomes (CLOs)	Reference
1	Install VMware	CLO9	T2:1
2	Hadoop Modes	CLO1, CLO2	T1:2.6-2.9
3	Using Linux Operating System	CLO1, CLO2	T1:4.1-4.2
4	File Management In Hadoop	CLO1, CLO2	T1:2.1-2.4 T1:2.6-2.9
5	Mapreduce Program 1	CLO2, CLO4	T1:4.2-4.3
6	Mapreduce Program 2	CLO1, CLO2,CLO4	T1:4.4-4.7
7	Mapreduce Program 3	CLO2, CLO4,CLO5	T1:5.6- 5.10
8	Mapreduce Program 4	CLO2, CLO4	T1:5.6- 5.10
9	Pig Latin Language - Pig	CLO2, CLO4	T1:5.6- 5.10
10	Pig Commands	CLO10,CLO11	T2:1, 2
11	Pig Latin Modes, Pig Program	CLO10,CLO11	T2:2
12	Hive	CLO11,CLO12,CLO13,CLO14	T2:10
13	Hive Operations	CLO11,CLO12,CLO13,CLO14	T2:10

#### XV. GAPS IN THE SYLLABUS - TO MEET INDUSTRY / PROFESSION REQUIREMENTS:

S No	Description	Proposed actions	Relevance with POs	Relevance with PSOs
1	Implementation of application that stores big data in MongoDB	Seminars	PO1,PO2,PO5	PSO 1
2	Implementation of application that stores big data in R language	Seminars / NPTEL	PO2,PO5	PSO 3

**Prepared by:**

Ms. S Swarajya Laxmi, Assistant Professor

**HOD, CSE**