



# INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal - 500 043, Hyderabad, Telangana

## COURSE CONTENT

### BIG DATA AND ANALYTICS LABORATORY

VII Semester: CSE								
Course Code	Category	Hours / Week			Credits	Maximum Marks		
ACSC34	Core	L	T	P	C	CIA	SEE	Total
		0	0	3	1.5	30	70	100
<b>Contact Classes: Nil</b>		<b>Tutorials Classes: Nil</b>		<b>Practical Classes: 36</b>		<b>Total Classes: 36</b>		
<b>Prerequisite: There are no prerequisites to take this course</b>								
<b>I. COURSE OVERVIEW:</b>								
Fundamental principles of Big Data Analytics and its role in making better decisions and predictions in the organization. The course also covers the Technology, Infrastructure and Applications of BigData. Concepts of data identification, data cleansing and integration are also addressed. Software requirements of Big Data are addressed and case studies of Big Data Applications are discussed.								
<b>II. COURSE OBJECTIVES:</b>								
<b>The students will try to learn:</b>								
I. The implications and challenges of Big Data Analytics.								
II. The technologies used to store, manage, and analyze BigData in a Hadoop ecosystem.								
III. The hypothesis on the optimized business decisions in solving complex real-world problems.								
<b>III. COURSE SYLLABUS</b>								
<b>WEEK-1: INSTALLVMWARE</b>								
Installation of VMWare to set up the Hado openenvironment and its ecosystems								
<b>WEEK-2: HADOOP MODES</b>								
Perform setting up and Installing Hadoop in its three operating modes.								
a. Standalone.								
b. Pseudo distributed.								
c. Fully distributed.								
<b>WEEK-3: USING LINUX OPERATING SYSTEM</b>								
Implementing the basic commands of LINUX Operating System–File / Directory creation, deletion, update operations.								
a. Create a directory in HDFS at given path(s).								
b. List the contents of a directory.								
c. Upload and download a file in HDFS.								
d. See contents of a file								
e. Copy a file from source to destination								
f. Copy a file from / To Local file system to HDFS								
g. Move file from source to destination.								
h. Remove a file or directory in HDFS.								
<b>WEEK-4: FILE MANAGEMENT IN HADOOP</b>								
Implement the following file management tasks in Hadoop:								
a. Copy a file from/ To Local file system to HDFS								
b. Move file from source to destination.								
c. Remove a file or directory in HDFS.								
d. Display the aggregate length of a file.								
<b>WEEK-5: MAPREDUCE PROGRAM 1</b>								
Run a basic word count Map Reduce program to understand Map Reduce Paradigm.								

**WEEK-6: MAPREDUCE PROGRAM 2**

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.

**WEEK-7: MAPREDUCE PROGRAM 3**

Implement matrix multiplication with Hadoop Map Reduce

**WEEK-8: PIG LATIN LANGUAGE - PIG**

Installation of PIG.

**WEEK-9: PIG COMMANDS**

Write Pig Latin scripts sort, group, join, project, and filter your data.

**WEEK-10 PIG LATIN MODES, PROGRAMS**

- 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.

**WEEK-11: HIVE**

Installation of HIVE.

**WEEK-12: HIVE OPERATIONS**

Use Hive to create, alter, and drop databases, tables, views, functions, and indexes.

**IV. REFERENCEBOOKS:**

1. Jay Liebowitz, "Big Data and Business Analytics Laboratory", CRC Press.
2. DT Editorial Services, "Black Book, BigData", Dreamtech Publications, 1<sup>st</sup> Edition, 2017.
3. Tom White, "Hadoop: The Definitive Guide", O'Reilly, 3<sup>rd</sup> Edition, 2012.

**V. WEBREFERENCES:**

1. <https://www.iiitnr.ac.in/>
2. <https://ibse.iitm.ac.in/>