FACULTY OF COMPUTER SCIENCE

Master of Computer Application (Sem-II)
In Effect from Academic Year 2023-24

Branch Name :	MCA
Program Code :	CS201
Course Name :	Big Data Analytics Practical
Course Code :	3CS2010208P
Pre-requisite Course :	Working knowledge of programming language and database concepts

Course Objectives:

1. To understand basics of Big Data and Big Data Tools (Hadoop, MapReduce)
2. To understand fundamental techniques used for Big data analytics
3. To help a student to perform a variety of "analytics" on different data sets and to arrive at positive conclusions.

Teaching Scheme (Hours per week)				Evaluation Scheme (Marks)				
				Theory		Practical		
Lecture	Tutorial	Practical	Credit	University Assessment	Continuous Assessment	University Assessment	Continuous Assessment	Total
-	-	3	3	-	-	25	25	50

LAB/ Practical:

Note: The practical list provided beneath is for reference only. The course teacher may Change/formulate it as per his/her methodology and requirement.

Sr.No	Practical Experiments
1.	HADOOP: How to run word count program on Hadoop and HDFS
2.	2.1 MongoDb Installation Steps 2.2 Run various commands like help(), auth(), createUser(),cloneDatabase() etc. 2.3 Run various database related command 2.4 Try createCollection() and collection.drop() methods
3.	1.1 Create a file into the local file system and export it to the HDFS systems 1.2 To work with Pig in both Hadoop local mode and HDFS / MapReduce mode 1.3 Define schema using Load command and validating it. 1.4 Try various Relational operators
4	**Algorithms Using MapReduce**: Matrix-Vector Multiplication by MapReduce, Relational operators (Selection, projection, Union, Intersection and difference), Computing Natural Join by MapReduce, Grouping and Aggregation by MapReduce, Matrix Multiplication, Matrix Multiplication with One MapReduce Step.

Text Books:

1. Radha Shankarmani, M Vijayalakshmi, Big Data Analytics, 2nd Edition, Wiley
2. Jure Leskovec, AnandRajaraman, Jeffrey David Ullman, "Mining of Massive Datasets", Cambridge University Press, Second Edition, 2014.

Reference Books:

1. Seema Acharya, Subhashini Chhellappan, BIG Data and Analytics ,Willey
2. Jiawei Han, Micheline Kamber, Jian Pei, "Data Mining Concepts and Techniques", Morgan Kaufman Publications, Third Edition, 2011.
3. VigneshPrajapati, "Big Data Analytics with R and Haoop", Packet Publishing 2013
4. Big Data Black Book, DreamTech

List of Open Source Software/learning website::

2. http://www.bigdatauniversity.com
3. http://www.mongodb.com
4. http://hadoop.apache.org

Course Learning Outcomes (CLO): On completion of this course, the students will be able to:

CLO	Description	Bloom's Taxonomy Level
CLO1	Understand the importance of Big Data and Big Data Tools for solving real world problems	1 Remembering 2 Understanding
CLO2	Understand the mechanism of fundamental techniques used for the Big Data Analytics.	1 Remembering 2 Understanding
CLO3	Analyze the Big Data framework like Hadoop and NOSQL to efficiently store and process Big Data to generate analytics.	4. Analyze 6. Creating
CLO4	Design of Algorithms to solve Data Intensive Problems using Map Reduce Paradigm	3. Applying 4. Analyze 6. Creating
CLO5	Access and Process Data on Distributed File System	1 Remembering 2 Understanding 3. Applying
CLO6	Understand the implementation of Data Stream Mining, Link Analysis, Recommendation System & Social Network Graphs	1 Remembering 2 Understanding 4. Analyze 5. Evaluate

Mapping of CLOs with POs & PSOs

Course Learning Outcomes	Program Outcomes (POs)												Program Specific Outcomes(PSOs)	
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CLO1	H	L	L				L				L		H	M
CLO2	H	M	L			L		M		H				M
CLO3		H			M	H		M				M	M	M
CLO4		H	H	H		M		M	M		M	L	H	M
CLO5		M	M	M	H	M	H	M	H				M	M
CLO6	H	M	L			L		M		H	L		M	M

H: High, M: Medium, L: Low