Nirma University

Institute of Technology

Computer Engineering Department

Practical List

2CS702 **Big Data Analytics** **[2-0-2-3]**

Laboratory details: (List of Experiments, Schedule, assessment policy)

Sr. No.	Problem definition	Hours	CLO
1.	Study and explore various applications of big data in different domains. Prepare a study report for any selected application which includes, data sources, analytics and visualization algorithms, etc. The domain may be selected from the followings, • Big Data in Retail • Big Data in Healthcare • Big Data in Education • Big Data in E-commerce • Big Data in Media and Entertainment • Big Data in Finance • Big Data in Travel Industry • Big Data in Telecom • Big Data in Automobile	02	3
2	Identify the data sources for big data. Find the technological limitations of conventional data analysis algorithms to perform analytics on big data. Justify your answer with any one of the applications.	02	3
3.	Install and configure single node Hadoop cluster. Perform HDFS commands on singlenode Hadoop Cluster.	04	3
4.	Design and implement MapReduce program to find phrase frequency from the given dataset.	04	3
5	Design and implement the MapReduce program to find the sum of given values.	02	3

6	Configure the number of mappers and reducers to execute practical 4 and 5. Evaluate the results and prepare a report on performance.	02	3
7	Design and implement a k-means clustering algorithm using a map-reduce programming model.	04	3
8	Install and configure MongoDB. Perform CRUD operations on the given dataset.	02	3
9	Install and configure Cassandra. Perform CRUD operations on the given dataset.	02	3
10	Install and configure Spark. Analyze the performance of Spark in comparison with Hadoop using a given example scenario.	04	3