

VIDYA JYOTHI INSTITUTE OF TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

IV Yr I SEM

BIG DATA ANALYTICS LAB

Week - 1 &2: Implement the following Data structures in Java

- a) Linked Lists
- b) Stacks
- c) Queues
- d) Set
- e) Map

Week - 3: Perform setting up and Installing Hadoop in Pseudo distributed mode

Week - 4: Implement the following file management tasks in Hadoop:

- i. Adding files and directories
- ii. Retrieving files
- iii. Deleting files

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.

Week - 5: Run a basic Word Count Map Reduce program to understand Map-Reduce Paradigm.

Week - 6: Write a Map Reduce program that mines weather data.

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 &8:

Install and Run Pig then write Pig Latin scripts to sort, group, join, project, and filter your Data.

Week - 9 &10:

- i. Install and Run Hive then use Hive to create, alter, and drop databases, tables, views, Functions and indexes
- ii. Performance techniques in Hive partitions, bucketing.