

## SECTION-D

**Note:** Long answer type questions. Attempt any two questions out of three questions. (2x8=16)

Q.23 Describe the Architecture of Hadoop, including its main components and their functions.

Q.24 What is Hive? Briefly explain the working with Hive. Also explain the benefits of Apache Hive.

Q.25 Write a short note on:

- i) Job Tracker
- ii) Taxonomy of Big Data
- iii) Task Tracker

No. of Printed Pages : 4

223861

Roll No. ....

**6th Sem / Artificial Intelligence & Machine Learning**

**Subject:- Big Data Analytics**

Time : 3Hrs.

M.M. : 60

## SECTION-A

**Note:** Multiple choice questions. All questions are compulsory (6x1=6)

Q.1 What are the main components of big data?

- a) HDFS
- b) Map Reduce
- c) YARN
- d) All of the Above

Q.2 What is a Data Node in HDFS?

- a) A node that stores actual data blocks
- b) A node that manages metadata
- c) A node responsible for job tracking
- d) A node responsible for resource management

Q.3 \_\_\_\_\_ function is responsible for consolidating the results produced by each of the Map() functions/tasks.

- a) Reduce
- b) Map
- c) Reducer
- d) All of the mentioned.

- Q.4 What is the primary function of the Map Reducer programming model in Hadoop?
- a) Real time data processing
  - b) Batch processing of large datasets
  - c) Querying Structured Data
  - d) High Speed data ingestion.
- Q.5 Pig operates in mainly how many Hadoop?
- a) Two                                      b) Three
  - c) Four                                      d) Five
- Q.6 What is Hive used for in Hadoop?
- a) Data storage                              b) Data Processing
  - c) Data querying                              d) Data Ingestion

### SECTION-B

**Note:** Objective/Completion type questions. All questions are compulsory. (6x1=6)

- Q.7 What is big data?
- Q.8 Expand the term YARN?
- Q.9 Hadoop has \_\_\_\_\_ type of file structure.

- Q.10 Which language is used in Hadoop
- Q.11 What is Hive in big data?
- Q.12 Define Map Reduce.

### SECTION-C

**Note:** Short answer type questions. Attempt any eight questions out of ten questions. (8x4=32)

- Q.13 Explain the distributed file system?
- Q.14 What is Hadoop and what are its advantages?
- Q.15 Differentiate between RDBMS & Hadoop
- Q.16 What are the main characteristics of Big Data? Discuss at least three
- Q.17 What do you mean by Word Count and Word Count Problem?
- Q.18 Define the different stages of MapReduce.
- Q.19 List different sources of Big Data.
- Q.20 Differentiate between Apache Pig & MapReduce.
- Q.21 Briefly describe the various components of Pig?
- Q.22 Briefly explain the working with Pig.