# GITAM (Deemed to be University) [19EAI433]

# **B.Tech. Degree Examination** (Computer Science & Engineering)

### VII SEMESTER

#### **BIG DATA ANALYTICS**

(Effective from the admitted batch 2019-20)

Time: 3 Hours Max.Marks: 60

**Instructions:** All parts of the unit must be answered in one place only.

-----

#### **SECTION-A**

## 1. Answer all Questions:

(10x2=20)

- a) Differentiate between structured and Unstructured Data
- b) List out Big Data Applications in Healthcare
- c) Summarize the features of Distribution Models
- d) What is database sharding?
- e) Differentiate between Hadoop streaming and pipes.
- f) What are the functions of Name Node?
- g) What are the functions of YARN resource manager in map reduce?
- h) How to handle Task Failure during map reduce process?
- i) What are the basic attributes of a Keyspace in Cassandra?
- j) Write short notes on Cassandra clients.

# **SECTION-B**

# **Answer the following:**

(5x8=40)

#### **UNIT-I**

2. Analyze how big data and marketing is helpful for increase the profit in business.

#### OR

3. Illustrate the Role of Big Data in Algorithmic Trading.

#### **UNIT-II**

4. Scheme less databases plays major role in big data – Justify.

## OR

5. Explain the process of composing map-reduce calculation with examples.

# **UNIT-III**

6. Illustrate the process of data Flow in map reduce.

### OR

7. Analyze the need of data integrity in HDFS system.

## **UNIT-IV**

8. Explain about Hadoop YARN Architecture with diagram.

## OR

9. How can you determine the number of map tasks and reduce tasks based on requirements?

# **UNIT-V**

10. Explain working methods with HBase Client API.

## OR

11. How to integrate Hadoop with other languages such as R Programming, Python, or Ruby? Explain with example.

[46/VII S/123]