Reg. No. : ☐☐☐☐☐☐☐☐☐☐

Question Paper Code: 49809

B.E. / B.Tech. DEGREE EXAMINATION, AUGUST 2021

Elective

Information Technology

14UIT909- BIG DATA AND ANALYTICS

(Regulation 2014)

Duration: 1:45 hour   Maximum: 50 Marks

PART A - (10 x 2 = 20 Marks)

(Answer any ten of the following questions)

1. State the significance of resampling.

2. State the challenges of conventional systems in handling big data.

3. Differentiate between examples and instances from rule induction.

4. What is Language bias?

5. State the definition of data stream.

6. State the different filtering schemes used in streams.

7. What is the purpose of Apriori Algorithm?

8. Compare the Agglomerative and Divisive Hierarchical Clustering.

9. State the general form of Map and Reduce function in Hadoop Map Reduce.

10. What is the use of Hive Metastore?

11. Draw the Credit Risk Framework.

12. What is NoSQL?

13. Why Key Type need be both Writable and Comparable in Map-Reduce Programs.**?**

14. What is Hadoop streaming?

15. What is ZOO keeper?

(Answer any three of the following questions)

16.   What is Inter-Firewall and Trans-Firewall Analytics? Explain the Value Chain for Inter-Firewall and Trans-Firewall Analytics.   (10)

17.   Write a short notes on following terminology in NoSQL with suitable application

(i) Aggregates.   (10)

(ii) Application Side Joins8

18.   Detail the Read and Write operations in Hadoop Distributed File System (HDFS) and detail the Java API required for that.   (10)

19.   How does map reduce on YARN technology work? Details the failures in YARN over map reduce.   (10)

.20.   Write short notes on primary and complex data types in Hive data types.   (10)