M.C.A DEGREE EXAMINATION, AUGUST 2023

Second Semester

MASTER OF COMPUTER APPLICATION

CORE - MCACT204 - DATA SCIENCE & BIG DATA ANALYSIS

2020 Admission Onwards

A7C54F86

Time: 3 Hours   Maximum: 75 Marks

Part A

*Answer any **ten** questions*

*Each question carries **3** marks*

1. Explain the steps of knowledge discovercy from data

2. Write down any three Datamining Task primitives

3. How do we handle mising values in data preprocessing?

4. What is apriori property?

5. Compare supervised learning and unsupervised learning process.

6. What is cluster?

7. Which are the facets of data science?

8. How is Big Data useful in advertisement?

9. Explain Regions in HBase.

10. What are the features of MapReduce ?

11. Define RDBMS

12. Discuss about the data types in Hive

(10×3=30 marks)

Part B

*Answer **all** questions*

*Each question carries **9** marks*

13. a) Explain KDD(Knowledge Discovercy from Data) process with suitable diagram

    OR

    b) Diffrence Between OLAP and OLTP

14. a) Discuss FP Growth Algorithm with an example.

    OR

    b) Explain the general approach for building a Classification model in detail.

15. a) Which are the facets of data science?

    OR

    b) Describe the steps involved in data preparation in Data Science process.

16. a) Write a short note on different technologies for handling big data

    OR

    b) Describe elaborately on Hadoop architecture with HDFS and MapRdeuce

17. a) Explain the features of non-relational database.

    OR

    b) Explain a) architecture of Hive b) Any 5 built-in functions in Hive

(5×9=45 marks)