

7. (i) What do you mean by serialization and deserialization and its importance in Hadoop working environment ?
- (ii) Elaborate the HDFS architecture and list down different daemons working in HDFS cluster.

Unit-IV

16 each

8. (i) Detail out the working of all three major components Code Driver, Mapper and Reducer of Map Reduce framework.
- (ii) List various functions performed by Pig Execution Engine to successfully run the Pig queries.
9. (i) Why partitions and buckets are created in Hive ? Explain with the help of an example, adding, renaming and dropping a partition in Hive.
- (ii) Elaborate how a request from a client is propagated in Hbase architecture that is handled by its different components.

Roll No. :

Total No. of Questions : 9] [Total No. of Pages : 4

67101-N

MCA 3rd Semester (Regular) Examination,
February-2022
(MCA 2 Years Programme)
(w.e.f. 2021-22)

Paper-21MCA23C1

DATA MINING AND BIG DATA ANALYSIS

Time : Three Hours]

[Maximum Marks : 80

Before answering the questions, candidates should ensure that they have been supplied the correct and complete question paper. No complaint in this regard, will be entertained after examination.

Note :- Attempt five questions in all, selecting one question from each Unit. Q. No. 1 is compulsory. All questions carry equal marks.

- Compulsory Question** 2×8=16
1. (i) Pictorially present the Knowledge Discovery from Data Process.
- (ii) Give an illustration to show vertical partitioning performed in data mining tasks.

