

JAMIA HAMDARD
Department of Computer Science & Engineering
School of Engineering Sciences & Technology
New Delhi-110062

Examination: Even Semester Examination 2023

Paper Name: Big Data Analysis for Bioinformatics Data

Paper Code: MCSB 203

Time: 2.5Hours

Maximum Marks: 60

Note: This question paper has two sections : Section A and Section B. Assume missing data, if any.

Attempt (i) Section-A: All questions are compulsory.

(ii) Section-B: Attempt any FOUR QUESTIONS.

Section –A (All questions are compulsory) 4×5=20 Marks

1. Describe Hive integration and workflow steps?	CLO5	1
2. What are the features of Hadoop HDFS? Explain the functions of Name Node and Data Node?	CLO2	3
3. Explain structures, semi-structured and unstructured Data?	CLO1	4
4. Differentiate between Map-reduce, Hive and Pig?	CLO5	2

Section –B (Attempt any Four questions) 4×10=40 Marks

1. Explain the Map and Reduce components in word count example?	CLO3	4
2. Explore various Execution Models of PIG?	CLO4	3
3. Explain all the steps in execution of a problem using MapReduce. Draw diagram?	CLO3	2
4. Differentiate “Scale up and Scale out” Explain with an example How Hadoop uses Scale out feature to improve the Performance.	CLO2	3
5. What are the components of YARN? It is said YARN components work on master-slave relationship. Explain with the help of the diagram.	CLO2	4
6. What are the different types of the tables available in Hive? Where does the data of the Hive table get stored?	CLO5	4