**POORNIMA UNIVERSITY, JAIPUR**

**END SEMESTER EXAMINATION, April 2023**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **3BT6101** | Roll No. | Total Printed Pages: 2 |
| **3BT6101** |  |
| BCA I Year II-Semester (Main/Back) End Semester Examination, April 2023  **(All Spl.)** | |
| **BDS06101 : Big Data Analytics** | | | |

# Time: **3** Hours. Total Marks: **60**

Min. Passing Marks: **21**

Attempt **five** questions selecting one question from each Unit. There is internal choice from Unit I to Unit V. Marks of each question or its parts are indicated against each question / parts. Draw neat sketches wherever necessary to illustrate the answer. Assume missing data suitably (if any) and clearly indicate the same in the answer.

Use of following supporting material is permitted during examination for this subject.

# **1.--------------------------Nil--------------------** **2.------------------Nil-----------------------**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **UNIT-I (CO1)** | **Marks** |
| **Q.1** | **(a)** | Differentiate between Structured, Semi structured and unstructured data. | **(6)** |
|  |  |  |  |
|  | **(b)** | Discuss different sources of data generation. Also explain how the rate of data generation varies from one source to another. | **(6)** |
|  |  |  |  |
|  |  | **OR** |  |
|  |  |  |  |
| **Q.2** | **(a)** | Explain the V’s associated with big data. What is their significance with respect to big data analysis? | **(6)** |
|  |  |  |  |
|  | **(b)** | Why is RDBMS unable to store Big data? Justify. | **(6)** |
|  |  |  |  |
|  |  | **UNIT-II (CO2)** |  |
|  |  |  |  |
| **Q.3** | **(a)** | Explain the Apache Hadoop ecosystem in detail. | **(6)** |
|  |  |  |  |
|  | **(b)** | With the help of suitable diagram describe the Hadoop architecture. | **(6)** |
|  |  |  |  |
|  |  | **OR** |  |
|  |  |  |  |
| **Q.4** | **(a)** | What is the Master slave architecture? Discuss its advantages and disadvantages. | **(6)** |
|  |  |  |  |
|  | **(b)** | With reference to Hadoop explain the concepts of cluster, commodity hardware and rack awareness. | **(6)** |
|  |  | **UNIT-III (CO3)** |  |
|  |  |  |  |
| **Q.5** | **(a)** | Discuss in detail, Hadoop distributed file system. | **(6)** |
|  |  |  |  |
|  | **(b)** | What is the role of name node and secondary name node in Hadoop? | **(6)** |
|  |  |  |  |
|  |  | **OR** |  |
|  |  |  |  |
| **Q.6** | **(a)** | What is job tracker in Hadoop? How is it different from the task tracker? | **(6)** |
|  |  |  |  |
|  | **(b)** | Explain in detail the process of block allocation in Hadoop cluster. | **(6)** |
|  |  |  |  |
|  |  | **UNIT-IV (CO4)** |  |
|  |  |  |  |
| **Q.7** | **(a)** | Discuss how MapReduce works. | **(6)** |
|  |  |  |  |
|  | **(b)** | What is Serialization? Explain. | **(6)** |
|  |  |  |  |
|  |  | **OR** |  |
|  |  |  |  |
| **Q.8** | **(a)** | With the help of example, explain different types of joins in Hadoop. | **(6)** |
|  |  |  |  |
|  | **(b)** | Write a MapReduce application of your choice. | **(6)** |
|  |  |  |  |
|  |  | **UNIT V (CO5)** |  |
|  |  |  |  |
| **Q.9** | **(a)** | Describe YARN architecture. | **(6)** |
|  |  |  |  |
|  | **(b)** | What are the components of YARN? | **(6)** |
|  |  |  |  |
|  |  | **OR** |  |
|  |  |  |  |
| **Q.10** | **(a)** | How is Hadoop different from relational databases? | **(6)** |
|  |  |  |  |
|  | **(b)** | Explain the process of cluster setup and installation. | **(6)** |