Reg. No. \_\_\_\_\_\_\_\_\_\_\_\_



**End Semester Examination – Nov / Dec – 2019**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| **CODE** | **: 18CA2001** | **Duration** | **: 3HRS** |
| **SUB. NAME** | **: BIG DATA ANALYTICS** | **Max. Marks** | **: 100** |

**ANSWER ALL QUESTIONS (5 x 20 = 100 Marks)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Q. No.** | **Sub Div.** | **Questions** | | **Course**  **Outcome** | | **Marks** | |
| 1 | a. | Discuss about Unstructured data. State the issues and steps to deal with unstructured data. | | CO2 | | 14 | |
|  | b. | Following words are to be placed in the relevant basket:   |  |  |  | | --- | --- | --- | | Structured | Unstructured | Semi-structured |  1. EMail ii) MS Access iii) Facebook iv) Images v) Database   vi) MS Excel vii) XML viii) Videos ix) Images x) Tables | | CO1 | | 6 | |
| **(OR)** | | | | | | | |
| 2 | a. | Give the evolution of Big data. | | CO2 | | 6 | |
| b. | Write short notes on   1. Typical data warehouse environment. 2. Coexistence of Big data and Data warehouse. | | CO2 | | 14 | |
|  | | | | | | | |
| 3 | a. | State the difference between predictive and descriptive analytics. | | CO3 | | 4 | |
| b. | Briefly discuss about data science and responsibilities of data scientist. | | CO2 | | 16 | |
| **(OR)** | | | | | | | |
| 4 | a. | With neat diagram explain the components of Hadoop ecosystem. | | CO3 | | 10 | |
| b. | List the differences between Hadoop and SQL. | | CO3 | | 10 | |
|  | | | | | | | |
| 5. | a. | List the four modules that make up the Apache Hadoop framework. | | CO4 | | 4 | |
| b. | With neat diagram, explain processing of data with Hadoop. | | CO3 | | 16 | |
| **(OR)** | | | | | | | |
| 6 | a. | List the terms used in RDBMS and MongoDB. | | CO4 | | 4 | |
| b. | With suitable examples list various operations in MongoDB (CRUD). | | CO4 | | 16 | |
|  | | | | | | | |
| 7. | a. | Briefly describe about various map task. | | CO5 | | 5 | |
| b. | Write short notes on:   1. REDUCER iii) SEARCHING 2. PARTITIONER iv) SORTING | | CO3 | | 15 | |
| **(OR)** | | | | | | | |
| 8. | a. | List the Hive Collection Data Types. | | CO5 | | 4 | |
| b. | What is RCFILE? | | CO6 | | 2 | |
| c. | 1. With suitable example write about Hive Query Language (HQL). | | CO4 | | 7 | |
| 1. Explain the Hive architecture. | | CO5 | | 7 | |
| **Compulsory**: | | | | | | |
| 9. | a. | Write notes on decision tree. | CO5 | | 4 | |
| b. | With suitable example explain unsupervised machine learning algorithm. | CO6 | | 16 | |