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



**UNIVERSITY**

(Karunya Institute of Technology & Sciences)

(Declared as Deemed-to-be University under Sec.3 of the UGC Act, 1956)

**End Semester Examination –April/May– 2017**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| **Code :** | **15CS3007** | **Duration :** | **3hrs** |
| **Sub. Name :** | **BIG DATA PLATFORMS** | **Max. marks :** | **100** |

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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Q. No. | | Sub Div. | | Questions | Course  Outcome | Marks |
| 1. | a. | | Classify the digital data. Describe its sources with example. Give a comparativestudy on the same. | | CO1 | 10 |
| b. | | Define big data. Explain its characteristics with example. Discuss its sources and applications. | | CO1 | 10 |
| (OR) | | | | | | |
| 2. | a. | | In your university, you are getting streaming data from weather monitoring sensors. Also you have weather data for the past 3 years and spatial images of your region. You are requested to do an analysis by combining all these data and to make predictive analysis. In this scenario what type of database would you prefer? Discuss in detail with its types and advantages. | | CO1 | 12 |
| b. | | Outline the different phases of big data analytics and identify its challenges. | | CO1 | 8 |
| 3. | a. | | Distinguish SQL and NoSQL databases. | | CO1 | 10 |
| b. | | State the limitations of NoSQL database. | | CO1 | 5 |
| c. | | Illustrate the Brewers’ theorem with suitable examples. | | CO1 | 5 |
| (OR) | | | | | | |
| 4. | a . | | Write the difference between, “insert” and “save” in MongoDB. Give examples | | CO1 | 8 |
| b. | | Categorize the the following NoSQL Databases:  MongoDB, Cassandra, Neo4j, Riak, Hbase, Membase, Redis, Hyper table, RavenDB, Allegro graph, couchDB, Infinite graph.  Mention the corresponding vendors and clients | | CO1 | 6 |
| c. | | List the various data types in MongoDB. Also demonstrate “UPSERT” | | CO1 | 6 |
| 5. | a. | | Discuss the key aspects of Hadoop. | | CO2 | 5 |
| b. | | Explain the core components of Hadoopframework with a neat diagram. | | CO2 | 15 |
| (OR) | | | | | | |
| 6. | a. | | Compare Pig, Hive and Impala | | CO2 | 5 |
| b. | | List the components of the Hadoop Eco System. Describe each component in detail. | | CO2 | 15 |
| 7. | a. | | Consider the following schema and write the Cassandra command for the given queries.  Book(book\_id,book\_name,title,price,author)   1. Create a keyspace called “library” with replication factor- 2 and simple strategy and create a “book” table with the above mentioned fields, insert the necessary data 2. Insert a row into the book table with new fields quantity and year-of-publication. 3. Update the title of book\_id=1, to ‘Big Data' 4. Alter the book table by adding a set called “category” and update the values for “category” 5. Alter the book table by adding a map “subject” and update the values for “subject” 6. Display the details of the book which is published in the year 2007. 7. Display the details of the book which is written by the author”Seema”. | | CO1 | 14 |
| b. | | Mention the main features of Cassandra. | | CO1 | 6 |
| (OR) | | | | | | |
| 8. | a. | | Discuss the types of failures in MapReduce and its causes. | | CO2 | 4 |
| b. | | How status updates are propagated through the MapReduce system? Demonstrate with a neat diagram, necessary classes and methods. | | CO2 | 8 |
|  | c. | | Describe the process of shuffle and sort in MapReduce with a neat diagram. | | CO2 | 8 |
|  | | | **Compulsory**: | |  |  |
| 9. | a. | | Explain the functions of Master Server and Region Server in Hbase with a neat diagram. Also express the role of Zookeeper Hbase. | | CO1 | 15 |
| b. | | Decribe compaction in Hbase. | | CO1 | 5 |

ALL THE BEST