

# University of Mumbai

Program: **Electronics and Telecommunications**

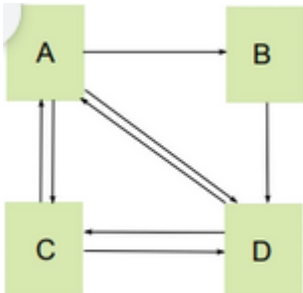
Curriculum Scheme: Rev 2016

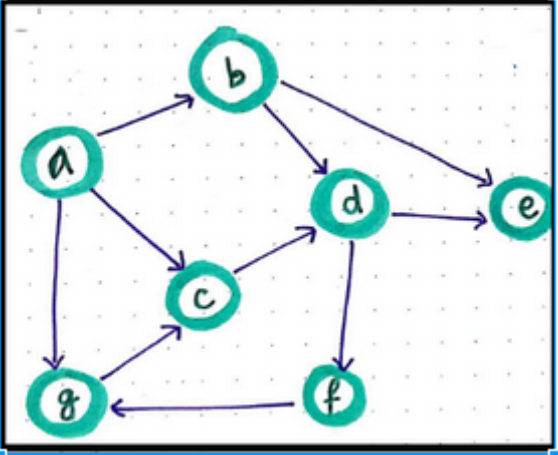
Examination: BE Semester:VII

Course Code:ECCDL7032 and Course Name: Big Data Analytics

Time: 2 hour 30 minutes

Max. Marks: 80

| Q1.       | Choose the correct option for following questions. All the Questions are compulsory and carry equal marks   |
|-----------|---|
| 1.        | Which of the following is not application of Nearest Neighbor Similarity search?  |
| Option A: | Optical character Recognition   |
| Option B: | Content based image retrieval   |
| Option C: | Document Similarity   |
| Option D: | Transaction processing system   |
| 2.        | Pages with no inlinks or outlinks to or from the SCC are called as  |
| Option A: | In-component  |
| Option B: | Tendrils  |
| Option C: | Tubes   |
| Option D: | Isolated components   |
| 3.        | Which node will have the highest page rank value after the first iteration ?<br> |
| Option A: | B   |
| Option B: | A   |
| Option C: | D   |
| Option D: | C   |
| 4.        | Which NoSQL database is mostly used for storing product catalog data in ecommerce sites?  |
| Option A: | Key-value   |
| Option B: | Columnar  |
| Option C: | Graph   |
| Option D: | Document  |
| 5.        | Arrange the following phases of a MapReduce program in the order that they execute? i. Partitioner ii. Mapper iii. Combiner iv. Shuffle/Sort                        |
| Option A: | i, iii , ii , iv  |

|           |  |
|-----------|--|
| Option B: | i , ii , iii , v   |
| Option C: | iii, ii, i, iv   |
| Option D: | ii, iii , i , iv   |
| 6.        | Cold Start Problems are associated with _____type of recommendation system.  |
| Option A: | Collaborative filtering  |
| Option B: | Content based  |
| Option C: | Hybrid   |
| Option D: | Item based filtering   |
| 7.        | How many dead ends exist in the following Web graph?<br> |
| Option A: | 2  |
| Option B: | 1  |
| Option C: | 0  |
| Option D: | 3  |
| 8.        | Which of the following is a core concept of Hadoop?  |
| Option A: | Applications depend on the network infrastructure.   |
| Option B: | All data are distributed in advance.   |
| Option C: | If a node fails performance will dip.  |
| Option D: | Nodes talk to each other as much as possible.  |
| 9.        | What is the problem faced by the Page Rank algorithm?  |
| Option A: | Dead Ends  |
| Option B: | Orphan Pages   |
| Option C: | Tendrils   |
| Option D: | Tubes  |
| 10.       | Which of the following is not the characteristic of stream data?   |
| Option A: | Continuous   |
| Option B: | Ordered  |
| Option C: | Persistent   |
| Option D: | Huge   |

|                                |  |                      |
|--------------------------------|--|----------------------|
| <b>Q2.<br/>(20 Marks Each)</b> | <b>Solve any Two Questions out of Three</b>  | <b>10 marks each</b> |
| A                              | Write the Map Reduce Pseudo code for any two of the following relational algebra operations<br>i) Projection ii) Selection iii) Natural Join iv) Matrix-Vector Multiplication v) Grouping & Aggregation. |                      |
| B                              | Explain any one clustering algorithm for mining Social Network Graph.  |                      |
| C                              | Explain Column family store and Graph Store NoSQL Architectural pattern with examples.   |                      |

|                                |  |                      |
|--------------------------------|--|----------------------|
| <b>Q3.<br/>(20 Marks Each)</b> | <b>Solve any Two Questions out of Three</b>  | <b>10 marks each</b> |
| A                              | Compare Content based Recommendation system with Collaborative filtering based Recommendation system |                      |
| B                              | What is MapReduce? Explain how map and reduce works?   |                      |
| C                              | Explain the recommendation system model along with the applications of it.                           |                      |

|                                |  |                      |
|--------------------------------|--|----------------------|
| <b>Q4.<br/>(20 Marks Each)</b> | <b>Solve any Two Questions out of Three</b>                              | <b>10 marks each</b> |
| A                              | Explain with diagrams the Park Chen Yu (PCY) Algorithm.                  |                      |
| B                              | What is NoSQL Data Architectural Pattern? What are its different types?  |                      |
| C                              | What is search engine? Discuss in detail Algorithm-Based Ranking System. |                      |