**18CSD12**

**Hall Ticket Number:**

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

|  |  |  |  |
| --- | --- | --- | --- |
| **III/IV B.Tech (Regular / Supplementary) DEGREE EXAMINATION** | | | |
| **January, 2022** | **Computer Science & Engineering** | | |
| **Fifth Semester** | **Data Warehousing and Data Mining** | | |
| **Time:** Three Hours | | **Maximum:** 50 Marks | |
| *Answer Question No.1 compulsorily.* | | | (10X1 = 10 Marks) |
| *Answer ONE question from each unit.* | | | (4X10=40 Marks) |
|  | | |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. | a) | What is a transactional database? | CO1 |  |
|  | b) | Define data cleaning. | CO1 |  |
|  | c) | Define concept Hierarchy. | CO1 |  |
|  | d) | What is Data Mart? | CO2 |  |
|  | e) | Define Min-Max Normalization. | CO2 |  |
|  | f) | How to compute support measure for an Itemset? | CO2 |  |
|  | g) | What is meant by frequent Itemset? | CO2 |  |
|  | h) | Define strong association rule. | CO3 |  |
|  | i) | What are the binary and nominal attributes? | CO3 |  |
|  | j) | What is a Dissimilarity matrix? | CO3 |  |
| **Unit -I** | | | | |
| 2. | a) | Define Data Mining. Describe the important tasks of data mining. | CO1 | 5M |
|  | b) | What are the types of data? Differentiate among structured and unstructured data with suitable examples. | CO1 | 5M |
| **(OR)** | | | | |
| 3. | What is data pre-processing? Explain why it is essential to pre-process the data before mining. Explain in detail about various preprocessing techniques. | | CO1 | 10M |
| **Unit -II** | | | | |
| 4. | a) | List out the OLAP operations and explain the same with an example. | CO2 | 5M |
|  | b) | Explain multidimensional data model with a neat diagram. | CO2 | 5M |
| **(OR)** | | | | |
| 5. | a) | Write short notes on summarization-based characterization. | CO2 | 5M |
|  | b) | What is Data mart? Describe the process of conversion from On-Line Analytical Processing to On-Line Analytical Mining. | CO2 | 5M |
| **Unit -III** | | | | |
| 6. | a) | Define Association rule mining and explain how Apriori algorithm works with suitable example. | CO3 | 5M |
|  | b) | How market basket analysis helps in finding the associations among different items? Explain with an example | CO3 | 5M |
| **(OR)** | | | | |
| 7. | a) | Elaborate the steps for improving the efficiency of Apriori algorithm. | CO3 | 5M |
|  | b) | Explain FP-Growth algorithm for suitability of mining frequent patterns | CO3 | 5M |
| **Unit -IV** | | | | |
| 8. | a) | Define clustering. Explain about cluster analysis. | CO4 | 5M |
|  | b) | Discuss K-means algorithm with a neat diagrams. | CO4 | 5M |
| **(OR)** | | | | |
| 9. | a) | Differentiate between agglomerative and divisive hierarchical clustering. | CO4 | 5M |
|  | b) | Explain briefly about outlier analysis. | CO4 | 5M |

****