|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Roll No. |  |  |  |  |  |  |  |  |  |  |  |  |

****

**PRESIDENCY UNIVERSITY**

**Bengaluru**

|  |
| --- |
| **End - Term Examinations –JANUARY 2025** |
| **Date:**13 – 01-2025 **Time:**09:30 am – 12:30 pm |

|  |  |  |
| --- | --- | --- |
| **School:** SOCSE | **Program:** B .Tech -ISE/ISB/ISR | |
| **Course Code :** CSE2021 | **Course Name :** Data Mining | |
| **Semester**: V | **Max Marks**: 100 | **Weightage**: 50% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CO - Levels** | **CO1** | **CO2** | **CO3** | **CO4** | **CO5** |
| **Marks** | **24** | **24** | **24** | **24** | **24** |

**Instructions:**

1. *Read all questions carefully and answer accordingly.*
2. *Do not write anything on the question paper other than roll number.*

**Part A**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Answer ALL the Questions. Each question carries 2marks. 10Q x2M=20M** | | | | |
| **1** | Describe the definition of data mining? | **2 Marks** | **L1** | **CO1** |
| **2** | Define Data cube? | **2 Marks** | **L1** | **CO1** |
| **3** | What is constraint based association mining? | **2 Marks** | **L2** | **CO2** |
| **4** | What is concept description of hierarchies? | **2 Marks** | **L2** | **CO2** |
| **5** | Define association rules, with example? | **2 Marks** | **L2** | **CO2** |
| **6** | Derive the formula for correlation and covariance analysis. | **2 Marks** | **L2** | **CO3** |
| **7** | Explain Data Integration? | **2 Marks** | **L1** | **CO3** |
| **8** | How many types of clustering and explain? | **2 Marks** | **L2** | **CO4** |
| **9** | Explain the various types of attributes? | **2 Marks** | **L2** | **CO4** |
| **10** |  |  |  |  |

**Part B**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Answer ALL Questions Total 80 Marks.** | | | | | |
| **11.** | **a.**  **b.** | Explain Various applications of data mining? Elucidate the functionalities of data mining with suitable real time examples? | **10 Marks**  **10 Marks** | **L2**  **L2** | **CO1**  **CO1** |
| **Or** | | | | | |
| **12.** | **a.**  **b.** | Explain Data reduction , define Attribute Subset Selection on {A1, A2, A3, A4, A5, A6} given data sets .  Suppose that the minimum and maximum values for the attribute income are $12,000 and $98,000, respectively. We would like to map income to the range [0.0,1.0]. By min-max normalization, a value of $73,600 for income is transformed find the normalization. Apply  i. Min-Max  ii. Z-Score  iii. Decimal Scaling | **10 Marks**  **10 Marks** | **L2**  **L2** | **CO1**  **CO2** |
|  |  |  |  |  |  |
| **13.** | **a.** | For the given transactional data find the interesting pattern. Use Apriori algorithm have min sup= 2 and min –conf= 50%. Determine the association and efficiency of the algorithm. | **20 Marks** | **L2** | **CO2** |
| **Or** | | | | | |
| **14.** | **a.** | Given a samples of S, where S={(5, Y), (2, N), (6, Y), (7, N), (9, N), (1, N), (10, Y), (15, Y)}. If S has to be partitioned into two intervals S1 and S2 using two possible split points 5 and 9. Compute the best split point. | **20 Marks** | **L2** | **CO2** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **15.** | **a.** | Apply Decision tree induction for following dataset using ID3, C4.5, generate root and child nodes for the given table. | **20 Marks** | **L3** | **CO3** |
| **Or** | | | | | |
| **16.** | **a.** | Consider the following table presents a datasetof10 objects, with attributes, check whether the customer is satisfied or not. classify the data using following properties {Color=Red, Origin=Domestic, and Type=Formal} using Navie bayes algorithm. | **20 Marks** | **L3** | **CO3** |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **17.** | **a.** | What do you understand by Classification by Backpropagation? Solve the given network using initial values. Find the new weights of the network.   |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **X1** | **X2** | **X3** | **W14** | **W15** | **W24** | **W25** | **W34** | **W35** | **W46** | **W56** | **04** | **05** | **06** | | **2** | **3** | **4** | **1.0** | **1.1** | **1.3** | **1.4** | **1.6** | **1.7** | **2.0** | **2.1** | **0.4** | **0.2** | **0.3** | | **20 Marks** | **L3** | **CO4** |
| **Or** | | | | | |
| **18.** | **a.** | Use the similarity matrix in Table 3 to perform single link and complete link hierarchical clustering. Show your results by drawing a dendogram. The dendogram should clearly show the order in which the points are merged. | **20 Marks** | **L3** | **CO5** |

**\*\*\*\*\* BEST WISHES \*\*\*\*\***