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 :** | **14CS3074** | **Duration :** | **3hrs** |
| **Sub. Name :** | **ADVANCED DATA MINING** | **Max. marks :** | **100** |

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Q. No.** | **Sub Div.** | **Questions** | **Course**  **Outcome** | **Marks** |
| 1. | a. | Explain the knowledge discovery process. | CO1 | 10 |
| b. | Elucidate the typical framework of data warehouse. | CO1 | 10 |
| (OR) | | | | |
| 2. | a. | Why preprocess data? | CO1 | 15 |
| b. | Pen down the various strategies for data reduction. | CO1 | 5 |
| 3. | a. | Assume, you are a technical consultant for a chain of restaurants. The CEO of the company wants to know the current trends, customer reviews, competitors business patterns and company's growth for the past 5 years. In this case, how will you generalize and consolidate data in multidimensional space? | CO1 | 15 |
|  | b. | Draw three-tier data warehousing architecture. | CO1 | 5 |
| (OR) | | | | |
| 4. |  | Draw and explain the multidimensional model and its different forms. | CO2 | 20 |
| 5. |  | Elucidate support vector machines with nessesary diagrams. | CO3 | 20 |
| (OR) | | | | |
| 6. |  | Explain K-nearest neighbor classifiers with performance metrics and examples. | CO3 | 10 |
| 7. |  | Explain the process of finding frequent itemset using candidate generated from the given table.   |  |  | | --- | --- | | **TID** | **List of Item ID** | | T100 | I1,I2,I5 | | T101 | I2,I4 | | T102 | I2,I3 | | T103 | I1,I2,I4 | | T104 | I1,I3 | | T105 | I2,I3 | | T106 | I1,I3 | | T107 | I1,I3 | | T108 | I1,I2,I3,I5 | | T109 | I1,I2,I3 | | CO4 | 20 |
| (OR) | | | | |
| 8. |  | Explain decision tree algorithm using the table given below.   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Outlook | Temperature | Humidity | Wind | Play-tennis | | sunny | hot | high | weak | No | | sunny | hot | high | strong | No | | overcast | hot | high | weak | Yes | | rain | mild | high | weak | Yes | | rain | cool | normal | weak | Yes | | rain | cool | normal | strong | No | | overcast | mild | high | strong | Yes | | sunny | mild | high | weak | No | | sunny | cool | normal | weak | Yes | | rain | mild | normal | weak | Yes | | sunny | mild | normal | strong | Yes | | overcast | hot | normal | weak | Yes | | overcast | cool | normal | strong | Yes | | rain | mild | high | strong | No | | CO4 | 20 |
|  | | **Compulsory:** |  |  |
| 9. |  | Describe ensemble methods. | CO4 | 20 |