In Semester Examination-2,March 2022

Data Mining & Business Visualization

# YEAR/SEM: ST/ SEM IV DATE: 22/03/2022

**BRANCH: AIML TIME: 2.30 PM to 3.30 PM MARKS: 20**

## Subjective Questions Instructions:

1. All questions are compulsory.
2. Assume suitable data wherever necessary and state the assumptions made.
3. Figures to the right indicate full marks.

|  |  |  |
| --- | --- | --- |
| **Q2** | **Subjective Type Question (10 Marks)**  **Solve any 1 (5 Marks)** | MARKS |
| Q2.a | Illustrate the followings:  (i) FP tree Growth  (ii) Support and Confidence with example  (iii)Market Basket analysis | 5 |
| Q2.b | Illustrate how Apriori algorithm can be applied on the planning dataset to find frequent item set with minimum support count is 2 and minimum confidence be 80%.  Find all the frequent item set and also generate association rule using Apriori algorithm.For that consider the below dataset.     |  |  | | --- | --- | | TID | ITEM Bought | | T1 | 1 3 4 | | T2 | 2 3 5 | | T3 | 1 2 3 5 | | T4 | 2 5 | | 5 |
| **Q3** | **Solve any 1 (5 Marks)** |  |
| Q3.a | Predict the class label for an unknown sample ‘x’ using naïve Bayesian classification X=(age<=30,income=”med”, student=”Yes” credit rating=”fair”) The class label attribute is “Buys-computer” where the training tuples for all database is given below   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **ID** | **Age** | **Income** | **Student** | **Credit** | **Buys-computer** | | 1 | ≤30 | High | No | Fair | No | | 2 | ≤30 | High | No | Excellent | No | | 3 | 31- 40 | High | No | Fair | Yes | | 4 | > 40 | Med | No | Fair | Yes | | 5 | > 40 | Low | Yes | Fair | Yes | | 6 | > 40 | Low | Yes | Excellent | No | | 7 | 31- 40 | Low | Yes | Excellent | Yes | | 8 | ≤ 30 | Med | No | Fair | No | | 9 | ≤ 30 | Low | Yes | Fair | Yes | | 10 | > 40 | Med | Yes | Fair | Yes | | 11 | ≤ 30 | Med | Yes | Excellent | Yes | | 12 | 31- 40 | Med | No | Excellent | Yes | | 13 | 31- 40 | High | Yes | Fair | Yes | | 14 | > 40 | Med | No | Excellent | No | | 5 |

Q3 b. (i) Evaluate Information Gain, Entropy and Gain for “Age” for above table. 5

(ii) Explain ID3 Algorithm .Provide the importance of relationship of Information gain

and Entropy.