Department of Computer Engineering
Academic Year 2021-2022
Term Test – II Subjective Questions

Date: 13/12/2021   **Maximum Marks: 15**

Course Name: Data Mining and Warehousing   **Course Code: DJ19CEC501**
Class: TE (A & B)   **Sem:** V
Maximum Marks: 15   **Duration**: 20 min

Instructions:

1. Please solve questions in order with clear and dark ink pens
2. Draw figures wherever required
3. Sign with Name at the end of each page

Q1	You are given the transaction data shown in the Table below from a fast food restaurant. There are 9 distinct transactions (order:100 – order:900) and each transaction involves between 2 and 4 meal items. There are a total of 5 meal items that are involved in the transactions. For simplicity we assign the meal items short names (I1 – I5) rather than the full descriptive names (e.g., Big Mac).	**[09]**

Transaction – ID	Items
100	I1, I2, I5
200	I2, I4
300	I2, I3
400	I1, I2, I4
500	I1, I3
600	I1, I3
700	I1, I3, I2, I5
800	I1, I3
900	I1, I2, I3

Minimum Support Count = 2

Minimum – Confidence = 70%.

Apply Apriori to find all frequent itemsets. Also, generate the association rules and mention the strong association rules.

Q2	What are Crawlers? Differentiate between Regular Crawlers and Focused Crawlers.	**[03]**
Q3a	Design a star schema for the census department of Indian government. The census department oversees finding official count of people living in the country, jobs,	**[03]**

Department of Computer Engineering
Academic Year 2021-2022
Term Test – II Subjective Questions

Date: 13/12/2021   **Maximum Marks: 15**

	economics status etc.	
	OR	
Q3b	What are various approaches to designing a DW? Discuss advantages and their disadvantages.	[03]