Chandigarh College of Engineering and Technology (Degree Wing), Chandigarh
Department of Computer Science and Engineering
Sessional Test -1 (April-2022)

Subject (Code):Data Mining and Analysis (CS605C)   Semester: VIth
Name of Teacher: Dr. Ankit Gupta   Date (Session): 08.04.2022(M)
Time: 1:30 Hrs   Maximum Marks 30

Note: 1. Attempt all questions. 2. All question carry equal marks.

Q.N	Questions	CO
1.	a. Suppose that the minimum and maximum values for some attribute "income" are $10,000 and $195,000, respectively. We would like to map income to the range [20.0, 30.0]. Calculate the transformed value of $80000 by min-max normalization. [2] b. Explain Attribute Oriented Induction and its implementation with suitable example. [3]	CO1
2.	Discuss various Schemas used for Multidimensional model with the help of suitable example.	CO1
3.	a. "Data Warehouse deals with Archived Data". Comment. [2] → *Historical data* b. Define *Support* and *"Confidence"*. How they affect the data mining process. [3] → *Are all pattern intresting*	CO1
4.	a. What is a Concept Hierarchy? Explain with the help of suitable example. [3] b. Compare Symmetric binary attributes and Asymmetric binary attributes with the help of suitable example. [2]	CO1
5.	Suppose that a group of 1,500 people was surveyed. The gender of each person was noted. Each person was polled as to see whether their preferred type of reading material was fiction or nonfiction. Thus, we have two attributes, gender and preferred reading. The observed frequency (or count) of each possible joint event is summarized in the contingency table shown in Table male female Total Fiction 250 200 450 Non fiction 50 1000 1050 Total 300 1200 1500 Find out whether the two attributes are correlated or not. *Since our computed value is greater Thus attributes are correaleted* (Assume that for 1 degree of freedom, the chi-square value needed to reject the hypothesis at the 0.001 significance level is 10.828.)	CO1
6.	Explain the Architecture of Data Warehouse. Also Discuss various Data Cube measures. → *Multidimensional data model*	CO1