PRIYADARSHINI COLLEGE OF ENGINEERING, NAGPUR
DEPARTMENT OF COMPUTER TECHNOLOGY
ACADEMIC SESSION: 2022-23(EVEN SEMESTER)
Question Bank for CAT-2

Subject	:	Data Warehousing and Mining(BTCT602T)	**Semester**	:	VI A&B
Subject teacher	:	Mrs.R.A.Khan/Ms.Shreyanshi Patel	**Date of Display**	:	03/04/2023
Unit	:	III, IV & V			

Course Outcomes: Upon the succesfull completion of the course, students will be to:

CO3	Explore and illustrate different classification and data clusturing techniques.
CO4	Apply different mining techniques for frequent itemset mining
CO5	Describe and compare various techniques of Web,Temporal and Spatial data mining.

<table>
<tr><th>Q.No.</th><th>Questions</th><th>Mapping with COs</th><th>Marks</th></tr>
<tr><td rowspan="2">1</td><td>a) Enlist various types of data in cluster analysis.</td><td>CO3</td><td>5</td></tr>
<tr><td>b) Explain k-means algorithm.</td><td>CO3</td><td>7</td></tr>
<tr><td rowspan="2">2</td><td>a)Write a detailed note on split algorithm based on gini index.</td><td>CO3</td><td>5</td></tr>
<tr><td>b)What do you mean by hierarchical clustering approach? Explain agglomerative and divisive hierarchical clustering.</td><td>CO3</td><td>7</td></tr>
<tr><td rowspan="2">3</td><td>a)Explain Bayesian classification with suitable example.</td><td>CO3</td><td>5</td></tr>
<tr><td>b)Explain Naive Bays algorithm with an example.</td><td>CO3</td><td>7</td></tr>
<tr><td rowspan="3">4</td><td>a) Consider the following dataset consisting of the scores of two variables on each of seven subjects. Design K Means clustering for the data set.
<table>
<tr><th>Subject</th><th>A</th><th>B</th></tr>
<tr><td>1</td><td>1</td><td>1</td></tr>
<tr><td>2</td><td>15</td><td>2</td></tr>
<tr><td>3</td><td>3</td><td>4</td></tr>
<tr><td>4</td><td>5</td><td>7</td></tr>
<tr><td>5</td><td>3.5</td><td>5</td></tr>
<tr><td>6</td><td>4.4</td><td>5</td></tr>
<tr><td>7</td><td>3.5</td><td>4.5</td></tr>
</table></td><td>CO3</td><td>7</td></tr>
<tr><td>b) What is Frequent pattern mining and Association Rules? What is the use of both? Explain.</td><td>CO4</td><td>6</td></tr>
<tr><td>c) Explain the technique for improving efficiency of FP growth algorithm.</td><td>CO4</td><td>7</td></tr>
<tr><td>5</td><td>A database has five transactions. Let min sup = 2.
TID items bought
T1 {A, B, C, D, E,}
T2 {B, C, D}
T3 {B, C, D, E}
T4 {A, B, C, D, E}
T5 {B, C, D, E}
Find all frequent itemsets using FP-growth Algorithm.</td><td>CO4</td><td>7</td></tr>
</table>

6	Consider Transactional data for an *AllElectronics* branch. *TID* *List of item IDs* T100 I1, I2, I5 T200 I2, I4 T300 I2, I3 T400 I1, I2, I4 T500 I1, I3 T600 I2, I3 T700 I1, I3 T800 I1, I2, I3, I5 T900 I1, I2, There are nine transactions in this database, that is, $\|D\| = 9$. Apply the Apriori algrithm for finding frequent itemsets in *D*. consider min_support=2.	CO4	7
7	a)Explain Apriori Algorithm in detail.	CO4	7
	b)How FP growth algorithm works? Explain.	CO4	7
8	a) Explain Market Basket Analysis for mining frequent pattern set and association rules with suitable example.	CO4	7
9	Explain Web Mining in detail.	CO5	7
	Describe various Graph properties of Web.	CO5	7
10	How to access accuracy of text retrieval in text mining system?	CO5	7
	Explain Spatial Data Mining in detail.	CO5	7
11	What do you mean by web Data mining? Explain various steps involved in WDM.	CO5	7
	How data mining is useful in customer relationship management in e-business world?	CO5	7
12	Explain the concept of visual web data mining in detail.	CO5	6
13	Differentiate temporal and spatial data mining in detail.	CO5	6
14	Discuss the challenges that occurred during knowledge discovery on the web.	CO5	6
15	Discuss i) Web Content Mining ii) Web Usage Mining iii) Web Structure Mining iv) Visual Web Data Mining	CO5	12

Mrs.R.A.Khan /Ms.Shreyanshi Patel
Subject Teacher

Dr. (Mrs.) N. M. Thakare
(HOD, CT)