P P SAVANI UNIVERSITY

Sixth Semester of B. Tech. Examination

Dec 2022

SECE3031 Data Warehouse & Data Mining

3.12.2022, Saturday **Time: 01:00 p.m. To 03:30 p.m.** **Maximum Marks: 60**

Instructions:

1. The question paper comprises of two sections.
2. Section I and II must be attempted in separate answer sheets.
3. Make suitable assumptions and draw neat figures wherever required.
4. Use of scientific calculator is allowed.

SECTION – I

			CO	BTL
Q - 1	Answer the Following: (Any five)	**[05]**		
(i)	List out types of Data.		**1**	**2**
(ii)	In data mining, how many categories of functions are included?		**1**	**1**
(iii)	What is Extraction of Data?		**1**	**1**
(iv)	Define: noisy data		**1**	**2**
(v)	What is Pattern Evaluation?		**1**	**1**
(vi)	Define fact Table.		**2**	**1**
(vii)	Define star schema.		**2**	**1**
Q - 2 (a)	Explain methods of Data Computation?	**[05]**	**1**	**2**
Q - 2 (b)	A database has five transactions. Let min sup = 60% and min conf = 80% TID ITEMS 100 {C, O, O, K, E, Y} 200 {L, O, N, E, L, Y} 300 {T,A,K,E} 400 {M,O,N,K,I,E} 500 {B,A,K,E} Find all frequent itemset using FP-growth.	**[05]**	**1**	**5**

OR

Q - 2 (a)	Explain Co relation Analysis with example.	**[05]**	**1**	**5**
Q - 2 (b)	What do you mean by Data Processing? Explain with suitable application.	**[05]**	**1**	**2**
Q - 3 (a)	What is data cleaning? Discuss various ways of handling noisy values during data cleaning.	**[05]**	**1**	**1**
Q - 3 (b)	Define the term "Data Mining". With the help of a suitable diagram explain the process of knowledge discovery from databases.	**[05]**	**1**	**2**

OR

Q - 3 (a)	Discuss the constraint-based Association Rule Mining.	**[05]**	**1**	**2**
Q - 3 (b)	Discuss the issues of Data Mining.	**[05]**	**1**	**2**
Q - 4	Attempt any one.	**[05]**		
(i)	Explain OLAP operations in detail with suitable examples.		**2**	**3**
(ii)	Explain Datamining Architecture in detail.		**2**	**2**

SECTION – II

Q - 1	Answer the Following: (Any five)	**[05]**		
(i)	What is prediction?		**1**	**1**
(ii)	What is accuracy of a prediction?		**1**	**5**
(iii)	What are core points and border points?		**1**	**1**

(iv)	What is regression?		1	2
(v)	Define outlier analysis.		1	1
(vi)	Enlist types of clustering.		1	2
(vii)	Define K-means.		1	1
Q - 2 (a)	What is IDE3? Explain with example.	[05]	1	4
Q - 2 (b)	Explain how SVM classification works.	[05]	1	2
	OR			
Q - 2 (a)	Explain Lazy Learner with example.	[05]	1	4
Q - 2 (b)	What is backpropagation? Explain with suitable diagram.	[05]	1	2
Q - 3 (a)	Explain Bagging and Boosting.	[05]	1	2
Q - 3 (b)	Difference between linear and non-linear regression in detail	[05]	1	3
	OR			
Q - 3 (a)	Explain Density-Based Clustering	[05]	1	2
Q - 3 (b)	Define "clustering"? Mention any two applications of clustering.	[05]	1	6
Q - 4	Attempt any one.	[05]		
(i)	What is r2 in regression? Explain what is SSE, SSR, and SST in regression and how to find out the values of each term.		1	4
(ii)	Difference between classification and prediction.		1	4

****