VEER SURENDRA SAI UNIVERSITY OF TECHNOLOGY (VSSUT), ODISHA
Odd Mid Semester Examination for Academic Session 2025-26

COURSE NAME: B.Tech. IT   SEMESTER: 5th

BRANCH NAME: CSE

SUBJECT NAME: **Data Mining and Data Warehousing**

                                                 TIME: **90 Minutes**

FULL MARKS: **30**

Answer **All** Questions.

The figures in the right hand margin indicate Marks. *Symbols carry usual meaning.*

Q1.	Answer all Questions.	[2 × 3]
a)	Differentiate classification and prediction task.	- CO1
b)	Define Support, Support count, and Confidence with respect to Frequent itemset and Association rule?	- CO2
c)	What is posteriori probability in the context of Naïve-Bayes classification?	- CO3

Q2.		[8]
		- CO1
	A. Explain these Data Mining Functionalities: i) Classification, ii) Cluster Analysis, iii) Prediction, iv) Concept description.	4
	B. Discuss architecture of Data mining system with its all components. Use suitable block diagram.	4
	OR	
		- CO1
	C. Explain the KDD process with suitable example.	4
	D. Discuss Supervised machine learning and Unsupervised machine learning with suitable example.	4

Q3.		[8]
		- CO2
	A. Find frequent itemset and strong association rule from the give data (Table D) using FP growth algorithm, If the minimum confidence threshold is, say, 90%, and minimum support is 30%.	8
	OR	
		- CO2
	B. Find frequent itemset and strong association rule from the give data (below) using Apriori algorithm, If the minimum confidence threshold is, say, 90%, and minimum support is 30%.	8

Table D

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, I3

Q4.   [8]
                                                                                       - CO3
                                                                                       8

A. Use the data given in Dataset (in Table 1) and design a Decision tree classification model. Here, the "age", "income", "student", "credit_rating" are input attributes and "buys_computer" is the output attribute. Then, classify following data in to buys_computer "yes"/"no": X = (age 31..40, Income = low, Student = no, Credit_rating = Fair)

Table 1

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31...40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31...40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31...40	medium	no	excellent	yes
31...40	high	yes	fair	yes
>40	medium	no	excellent	no

OR

                                                                                       - CO3
                                                                                       8

B. Design a regression model using KNN from the given data and predict House Price Index from the given data (Age = 48, Loan = USD 150,000). Consider k = 3.

Age	Loan	House Price Index
25	$40,000	135
35	$60,000	256
45	$80,000	231
20	$20,000	267
35	$120,000	139
52	$18,000	150
23	$95,000	127
40	$62,000	216
60	$100,000	139
48	$220,000	250
33	$150,000	264