VIT

Vellore Institute of Technology

School of Information Technology and Engineering

Winter Semester 2022-2023

Continuous Assessment Test – II

Programme Name: MCA

Course Name & code: Data Mining and Business Intelligence - ITA5007

Class Number (s): 0528, 0296, 0530 Slot: C2+TC2

Faculty Name (s) (Dr. E.P.Ephzibah, Dr.Harshita Patel and Dr.S.Jagadeesan)

Exam Duration: 90 Min. Maximum Marks: 50

Answers all the Questions (5*10=50 Marks)

1. Use ID3 algorithm to construct a decision tree from the given data. Age, Competition, and Type are the input attributes. The Class is the output attribute with class labels Up and Down. Draw the generated decision tree with appropriate labels and node information.**(10)**

S.No	Age	Competition	Type	Class(Profit)
1	Old	Yes	Software	Down
2	Old	No	Software	Down
3	Old	No	Hardware	Down
4	Mid	Yes	Software	Down
5	Mid	Yes	Hardware	Down
6	Mid	No	Hardware	Up
7	Mid	No	Software	Up
8	New	Yes	Software	Up
9	New	No	Hardware	Up
10	New	No	Software	Up

2. The table given below has five weeks' sales data in Rupees (thousands) that deals with one dependent (y) and one independent variable (x). Implement a linear regression model on the data to find the line of regression. Predict the 7^{th} and 12^{th} week sales. **(10)**

S.No	Week(x)	Sales in Rupees (thousands) (y)
1	1	1.2
2	2	1.8
3	3	2.6
4	4	3.2
5	5	3.8

3. The table given below lists the training instances. Each training instance has two input attributes x1, x2, and one output attribute with class labels 1 and 0. Classify the new incoming instance t1 = (3, 7), with k=3 using K-Nearest Neighbour algorithm. Give your observations for assigning an even number to the parameter k. **(10)**

Training Instance	X1	X2	Output
I1	7	7	0
I2	7	4	0
I3	3	4	1
I4	1	4	1

4. Consider the training samples in the dataset given below. Let the test instance be X= (Slow, Rarely, No). Find the most appropriate class label for the given record, X using the Naive Bayes classifier. **(10)**

Swim	Fly	Crawl	Class label
Fast	No	No	Fish
Fast	No	Yes	Animal
Slow	No	No	Animal
Fast	No	No	Animal
No	Short	No	Bird
No	Short	No	Bird
No	Rarely	No	Animal
Slow	No	Yes	Animal
Slow	No	Yes	Fish
Slow	No	Yes	Fish
No	Long	No	Bird
Fast	No	No	Bird

5. The following table shows six transactions by customers at grocery store. Let minimum support is 33.34% and minimum confidence is 60%. Find all frequent item sets and the generated strong rules using the Apriori algorithm.**(10)**

Transaction Id	Items
T1	HotDogs,Buns,Ketchup
T2	Hotdogs, Buns
T3	HotDogs,Coke,Chips
T4	Chips,Coke
T5	Chips,Ketchup
T6	HotDogs,Coke,Chips