VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)
CHENNAI

Reg. Number:	

Continuous Assessment Test(CAT) – II - April 2024

Programme	:	Master of Computer Applications	Semester	:	Winter 2023-24
Course Code & Course Title	:	PMCA507L-Machine Learning	Class Number	:	CH2023240501386
Faculty	:	Dr.B.Saleena	Slot	:	B2+TB2
Duration	:	1 ½ Hours	Max. Mark		50

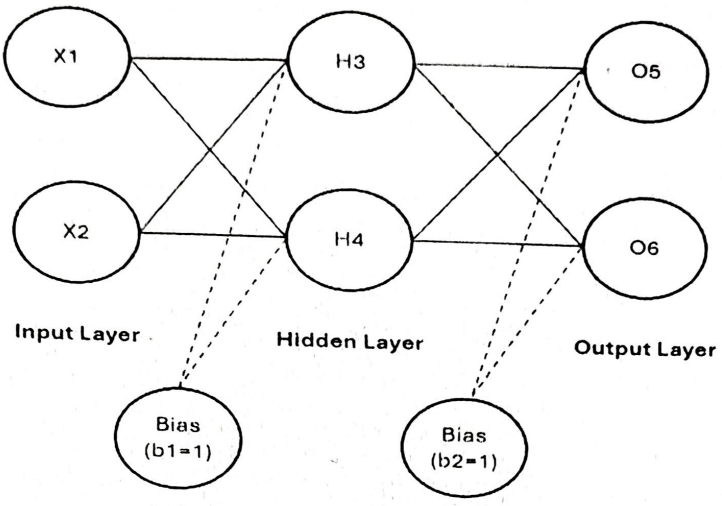
General Instructions:

- Write only your registration number on the question paper in the box provided and do not write other information.
- Only non-programmable calculator without storage is permitted

Answer ALL questions

<table>
<tr><th>Q. No</th><th>Sub Sec.</th><th>Description</th><th>Marks</th></tr>
<tr><td>1.</td><td></td><td>Consider the following data which belongs to 2 classes.
<table>
<tr><td>x_1</td><td>1</td><td>-1</td><td>-1</td><td>1</td><td>2</td><td>0</td><td>-2</td><td>0</td></tr>
<tr><td>x_2</td><td>1</td><td>1</td><td>-1</td><td>-1</td><td>0</td><td>2</td><td>0</td><td>-2</td></tr>
<tr><td>Class</td><td>Blue</td><td>Blue</td><td>Blue</td><td>Blue</td><td>Red</td><td>Red</td><td>Red</td><td>Red</td></tr>
</table>
Mapping Function:
$$\Phi\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{cases}\begin{pmatrix}6-x_1+(x_1-x_2)^2\\6-x_2+(x_1-x_2)^2\end{pmatrix} & \text{if } \sqrt{x_1^2+x_2^2} \geq 2\\ \begin{pmatrix}x_1\\x_2\end{pmatrix} & \text{otherwise}\end{cases}$$
(a) Justify with a graph that the dataset is not linearly separable (2 marks)
(b) Use the above mapping Function to map the above dataset to a new feature space so that a separating hyperplane can be identified. Illustrate the step-by-step procedure to find the hyperplane equation and plot the graph. (10 marks)
(c) Identify if the point (x1,x2) =(-1,2) belongs to red or blue class? (3 marks)</td><td>15</td></tr>
<tr><td>2.</td><td></td><td>Consider the Market Basket Data in the below table that illustrates the mechanism of frequent itemset mining. Assume minimum support=60% and minimum confidence=80%. Find all frequent itemsets using Apriori Algorithm. Show the step-by-step illustrations and derive the confidence rules.
<table>
<tr><th>Transaction ID</th><th>Items bought</th></tr>
<tr><td>101</td><td>~~Tomato,~~ ~~Potato,~~ ~~Onion~~</td></tr>
<tr><td>102</td><td>~~Tomato,~~ ~~Potato,~~ ~~Brinjal~~, Pumpkin</td></tr>
<tr><td>103</td><td>~~Tomato~~, ~~Potato,~~ ~~Onion~~, ~~Chilly~~</td></tr>
<tr><td>104</td><td>Lemon, Tamarind, Chilly</td></tr>
<tr><td>105</td><td>~~Tomato;~~ Potato, ~~Brinjal,~~</td></tr>
<tr><td>106</td><td>~~Potato~~, ~~Brinjal~~, ~~Onion~~, Chilly</td></tr>
</table></td><td>10</td></tr>
</table>

3.		Consider a neural network shown in the below figure. Assume the neurons use the sigmoid activation function to perform forward and backward passes. The inputs are X1=0.05 and X2=0.10 and Target outputs are O5=0.01 and O6=0.99 and the learning rate =0.5	15



Following are the weights assigned to the layers and the bias.

W13=0.15	W14=0.20	W23=0.25	W24=0.30
W35=0.40	W36=0.45	W45=0.50	W46=0.55
b1=0.35	b2=0.60		

a) Calculate the output of hidden layers and output layers using Forward propagation. **(7 Marks)**
b) Compute the new weights W35, W45,W36, and W46 using the back propagation algorithm for one iteration. **(8 marks)**

Illustrate the step-by-step procedure for solving the above problem.

4.		(a) Bring out the trade-off between bias and variance with a neat diagram. Which ensemble techniques can be used to reduce bias and Variance? **(5 Marks)** (b) How can ensemble techniques be used to reduce overfitting in decision trees? Justify your answer **(5 Marks)**	10

**********All the best **********