Reg. Number:	22BAI1266

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

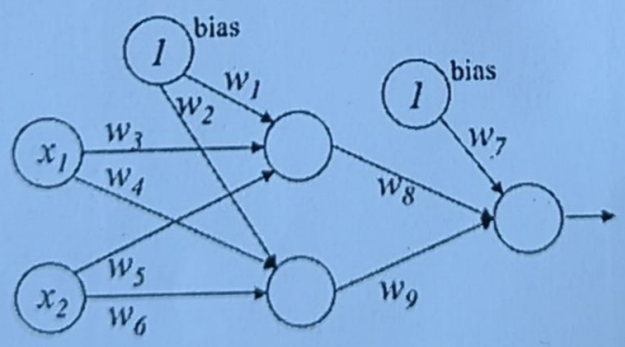
Continuous Assessment Test (CAT) – II - April 2024

Programme	:	B.Tech.(CSE)	Semester	:	Winter 23-24
			Code		BCSE209L
Course Code & Course Title	:	Machine Learning	Class Number	:	CH2023240501696
Faculty	:	Dr. Syed Ibrahim S P (50392)	Slot	:	C1+TC1
Duration	:	1½ 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.
- Use statistical tables supplied from the exam cell as necessary
- Use graph sheets supplied from the exam cell as necessary
- Only non-programmable calculator without storage is permitted

Answer all questions

Q. No	Sub Sec.	Description	Marks
1.		Consider a neural net for a binary classification which has one hidden layer as shown in the figure. We use a linear activation function f (z) = cz at hidden units and a sigmoid activation function g (z) $= \frac{1}{1+e^{-z}}$ at the output unit to learn the function from P(y=1|x,w) where x=(x_1, x_2) and w = (w_1, w_2,…, w_9).  1. What is the output P(y=1 |x,w) from the above neural net? Express it in terms of x_i, c and weights w_i. (5 Marks) 2. What is the final classification boundary? (3 Marks) 3. Draw a neural net with no hidden layer which is equivalent to the given neural net, and write weights w of this new neural net in terms of c and w_i. (2 Marks) 4. Is it true that any multi-layered neural net with linear activation function at hidden layers can be represented as a neural net without any hidden layer? Briefly explain your answer. (5 Marks)	15

<table>
<tr><td>2.</td><td></td><td>Consider a table with a single attribute “Wind” and Category “Rain”, where “wind” can take two attribute values – High and Low, and “Rain” has two classes – Yes and No. There are 10 entries in the table, and it is known that 8 entries in the table have wind = high. It is also known that 8 entries in the table also have rain = yes. Design a suitable logistic regression model.</td><td>10</td></tr>
<tr><td>3.</td><td></td><td>Table 1: Dataset
<table>
<tr><td>#</td><td colspan="4">Attribute</td></tr>
<tr><td></td><td>A1</td><td>A2</td><td>A3</td><td>A4</td></tr>
<tr><td>1</td><td>{{A1}}</td><td>{{A2}}</td><td>{{A3}}</td><td>{{A4}}</td></tr>
<tr><td>2</td><td>{{A5}}</td><td>{{A6}}</td><td>{{A7}}</td><td>{{A8}}</td></tr>
<tr><td>3</td><td>{{A9}}</td><td>{{A10}}</td><td>{{A11}}</td><td>{{A12}}</td></tr>
<tr><td>4</td><td>{{A13}}</td><td>{{A14}}</td><td>{{A15}}</td><td>{{A16}}</td></tr>
<tr><td>5</td><td>{{A17}}</td><td>{{A18}}</td><td>{{A19}}</td><td>{{A20}}</td></tr>
<tr><td>6</td><td>{{A21}}</td><td>{{A22}}</td><td>{{A23}}</td><td>{{A24}}</td></tr>
</table>
i) Consider the above dataset, {{A1 to A24}} represents continuous values range from 1.0 to 5.0. Choose two initial centroids optimally and form suitable Clusters based on feature values. [10 Marks].

ii) Assume, you have applied Fuzzy C-means clustering on the given data. Assume imaginary output and show that as a diagram to present it and provide its suitability in this data. [5 Marks].</td><td>15</td></tr>
<tr><td>4.</td><td></td><td>In online learning, we can update the decision boundary of a classifier based on new data without reprocessing the old data. Now for a new data point that is an outlier, which of the following classifiers are likely to be affected more severely?

MLP, NB, LR, SVM and any other classifier of your choice. Give explanation to your answer with suitable examples.</td><td>10</td></tr>
<tr><td colspan="4">**********All the best **********</td></tr>
</table>