National Institute of Technology Mizoram
Mid- Semester Examination, Odd Semester (2022-23)
Machine Learning (CSL – 1702)

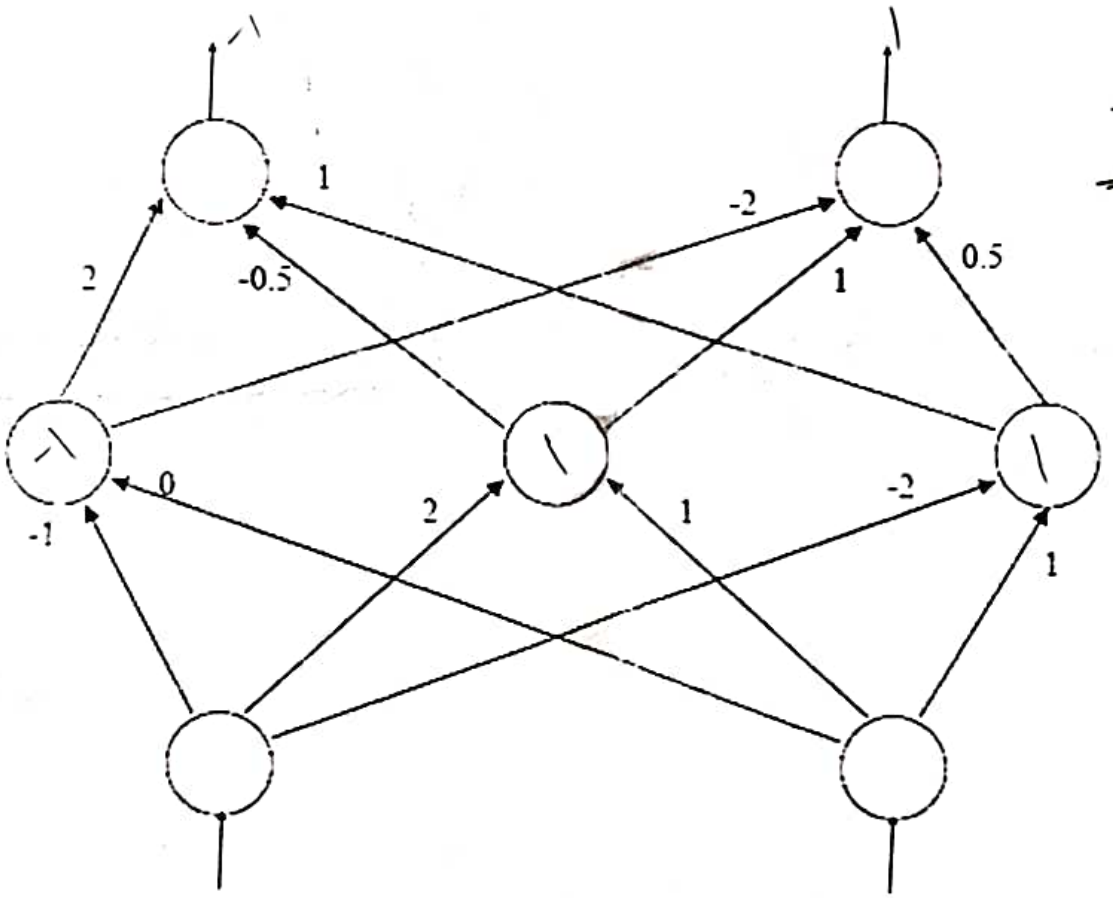
Semester - 5th **Full Marks - 30** **Duration - 1:30 hours**

Answer all 3(Three) Questions. All Questions carry same marks
(3 * 10 = 30 Marks)

21/30

1. (A) Realize the following function using MP neurons implemented for AND, OR, NOT gates in the class : $w = xyz + x'y$

   (B) Compute the output of the hidden-layer and the output-layer neurons with the hard limiter function for the input (0.5, 1.5) if applied to the network given in the following figure.



5 + 5

OR

(A) Illustrate why XOR logic is linearly non-separable.

(B) Discuss a general model of learning agents

5 + 5

2 (A) Compare inductive learning and deductive learning

(B) Consider the Hebbian learning rule for an ANN with a single PE which is an LTU. There are four inputs, x_1, x_2, x_3, and x_4 to this PE. The corresponding weight vector is $w = [w_1, w_2, w_3, w_4]^T$. Assume that this ANN is to be trained using the following three input vectors: $[1.5, 0.5, 0.5, 0]^T$, $[-0.5, 1, 0, 1.5]^T$, $[-1, 0, -1, -0.5]^T$. Assume $\eta=1$ and $w(1)=[1, 0.5, -1, 0]^T$. Train this network for one epoch and show intermediate updated weights and actual outputs along with the final.

5 + 5

OR

(A) (i)Explain "Turing Test" (2 X)

(ii) Compare meaningful learning and rote learning

(B) Compare supervised, unsupervised, and reinforcement learnings with neat diagrams (1 X)

(5 + 5)

(1 X) 3(A) Explain the working of an artificial neuron taking an analogy from the natural neurons.

(B) Explain the model of a simple perceptron network and its learning principle

(5 +5)

OR

. (A) Derive learning rules of a backpropagation algorithm.

(B) Give an outline of the backpropagation algorithm.

(5 +5)

NATIONAL INSTITUTE OF TECHNOLOGY MIZORAM

End–Semester Examination, Odd Semester (2022-2023)

Machine Learning (CSL – 1702)

7th Semester Maximum Marks: 50 Time: 3 hours

Answer all 5 (Five) Questions. All Questions Carry the Same Marks
(5 × 10 = 50 Marks)

1.a) Compare classification and regression
b) What is a bias-variance dilemma
c) You are training a logistic regression model and notice that it does not perform well on test data. Could the poor performance be due to underfitting? Justify.
d) What is the interpretation of the values of Pearson's correlation?
e) When no linear relationship exists between two variables, what would the regression look like?
f) What do you mean by VC dimension?
g) What for the "kernel trick" is used?
h) What happens to the performance of the k-NN classifier for various k values?
i) What is an autoencoder?
j) What is the link between entropy and information content

1 × 10 = 10

2a). Consider this distance table given in Table 1 and perform hierarchical clustering using the Average linkage strategy

	A	B	C	D	E	F
A	0	1	7	8	1	2
B	1	0	2	7	8	2
C	7	2	0	4	5	4
D	8	7	4	0	2	9
E	1	8	5	2	0	12
F	2	2	4	9	12	0

Table-1

b) Consider the following dataset 5
(0,1),(1,1),(1,2),(1,3), (2,1),(3,4),(5,5),(6,5),(6,6),(7,6),(7,7)
Perform clustering using the k-means clustering algorithm with three centers (k=2) as (0,1) and (6,1). (Two iterations only). 5

OR

c) Write the algorithm for fuzzy c-means clustering. 5
d) Compare DB-index for clusterings {(1,2), (3,5,6,8)} and {(1,2), (3,5), (6,8)} 5

3. a) It is known that 1.5% of the population suffers from a particular disease. A blood test has a 95% chance of identifying the disease for a diseased individual and an 8% chance of falsely indicating that a healthy person has a disease.
(i) What is the probability that a random person has a positive blood test? 0.09305
(ii) if a blood test is negative, what's the probability that the person does not have the disease? 0.999 2 + 3

b) Consider a 2-class problem with $P(C_1) = 3/4$, $P(C_2) = 1/4$; a scalar feature x and three possible actions α_1, α_2, α_3 defined as:

100

a_1: choose C_1
a_2: choose C_2
a_3: do not classify
Let the loss matrix $\lambda(a_i | C_j)$ be:

	a_1	a_2	a_3
C_1	0	1	1/4
C_2	1	0	1/4

and let $P(x | C_1) = (2-x)/3$, $P(x | C_2) = 1/3$, $0 \le x \le 2$
(i)Which action to decide for a pattern x: $0 \le x \le 2$?
(ii) What is the proportion of patterns for which action a_3 is performed (i.e., "do not classify")?
(iii) Compute the total minimum risk

2 + 2 + 1

OR

c) Consider Table 2 and compute the parameters for a Naïve Bayes classifier for predicting parameters for predicting "Play Golf." 5

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No

Table 2

d) (i) Explain the item-to-item Collaborative Filtering algorithm
(ii) Explain one technique for classifier combination in ensemble learning.

3 + 2

4.a) Realize nonlinear XOR logic using a RBF network. 5
b) Give an outline of the backpropagation learning algorithm with a diagram 5

OR

c) Discuss a convolutional neural network used in deep learning. 5
d) Explain SVM for two-class classification. 5

5. a) Create a decision tree following the ID3 algorithm for data given in Table 2. 5
b) Consider the data given in Table 3. Determine a decision tree through the Gini index to find E based on four attributes A, B, C, and D. 5

1	4.8	3.4	1.9	0.2	positive
2	5	3	1.6	1.2	positive
3	5	3.4	1.6	0.2	positive
4	5.2	3.5	1.5	0.2	positive
5	5.2	3.4	1.4	0.2	positive
6	4.7	3.2	1.6	0.2	positive
7	4.8	3.1	1.6	0.2	positive
8	5.4	3.4	1.5	0.4	positive
9	7	3.2	4.7	1.4	negative
10	6.4	3.2	4.7	1.5	negative
11	6.9	3.1	4.9	1.5	negative
12	5.5	2.3	4	1.3	negative
13	6.5	2.8	4.6	1.5	negative
14	5.7	2.8	4.5	1.3	negative
15	6.3	3.3	4.7	1.6	negative
16	4.9	2.4	3.3	1	negative

Table-3

OR

c) Find the least squares regression line for the data given in Table 4

Fue	Distance	Damage
1	3.4	26.2
2	1.8	17.8
3	4.6	31.3
4	2.3	23.1
5	3.1	27.5
6	5.5	36.9
7	0.7	14.1
8	3.0	22.3
9	2.6	19.6
10	4.3	31.3
11	2.1	24.0
12	1.1	17.3
13	6.1	43.2
14	4.8	36.4
15	5.8	26.1

Table-4

d) For a given classifier, suppose the first ten predictions of our classifier and ten true observations are as follows:

Prediction	1	1	1	1	1	0	1	1	1	1
True label	1	1	1	1	0	0	0	1	1	1

Compute accuracy, precision, recall, specificity, and false positive rate.