A) Define the followings: [8]

1- Define AI agent.
An agent perceives its environment through sensors and acts upon it through actuators (or effectors, depending on whom you ask). The Agent takes its action(s), which changes the environment, and process repeats.

2- What is back propagation and why do we use it?

Algorithm for training neural networks with hidden layers. Calculate error for output layer and propagate error back layer wise and update weights.

3- How supervised and reinforcement learning are related?

Both Supervised and RL get feedback:
Supervise Learning get feedback in form of Labeled Class
RL get feedback in form Reward.

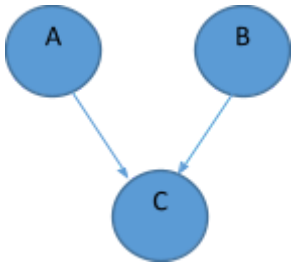
4- Define the role of activation function and give the name of popular activation function.

Activation function is non-linear function which is used activate/ trigger particular neuron. Sigmoid, tanh, RELU are mostly commonly used activation functions.

5- Differentiate between Markov decision process and reinforcement learning.

In MDP transition probability and rewards are known. In RL transitions and rewards are unknown and it is estimated through interacting with Environment

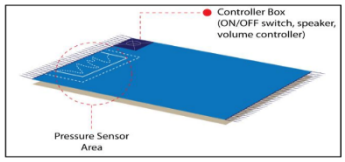
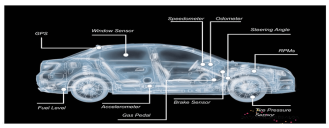
6- Given the figure below, determine the followings:
   a) Whether A and B are conditionally independent given C is observable (NO)
      No. If C is observed than probability A can effect the probability of B so they are not independent.
   b) Whether A and B are conditionally independent given C is unobservable (YES)
      Yes. They are conditional independent if C is unobserved. Probability of A does not affect probability of B.



7- What is the concept of gradient descent in neural network?

Gradient descent is used to minimize the loss function of neural network. We choose any starting point in weight space and then move to a neighboring point that is downhill, repeating until we converge on the minimum possible.

8- Explain the difference between impurity and abnormality of a Dataset?

Applications	PEAS	Environment Type	Agent type
 AI-enabled Prayer mat			
 Ambulance drone			
 AI-enabled smart shoes			
 Autonomous car			

Q No2:- Solution: Date:

Chart	Degree	x	y
A	34	8.29	5.59
B	45	7.07	7.07
C	55	5.735	8.191
D	48	6.691	7.43
E	49	6.56	7.54

$\bar{X} = 6.86$, $\bar{Y} = 7.1642$

x	y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$
8.29	5.59	1.43	-1.5742	2.0449	-2.2511
7.07	7.07	0.21	-0.0942	0.0441	-0.0197
5.735	8.191	-1.125	1.0268	1.2656	-1.155
6.691	7.43	-0.169	0.2658	0.028561	-0.044
6.56	7.54	-0.3	0.3758	0.09	-0.11274

$\sum = 3.473$ $\sum = 3.58254$

$$m = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2} = \frac{-3.58254}{3.473}$$

$$\boxed{m = -1.0315}$$

$$y = mx + c$$

$$7.1642 = (-1.0315)(6.86) + c$$

$$7.1642 = -7.07609 + c$$

$$7.1642 + 7.07609 = c$$

$$\boxed{c = 14.24}$$

Hence

$$\boxed{y = (-1.0315)x + 14.24}$$

d

Q no 3, Solution

Date:

(a)

1: $\sqrt{(32-8)^2+(10-20)^2} = \sqrt{576+100} = 26$

2: $\sqrt{(32-4)^2+(10-18)^2} = \sqrt{784+64} = 29.12$

3: $\sqrt{(32-5)^2+(10-16)^2} = \sqrt{729+36} = 27.65$

4: $\sqrt{(32-10)^2+(10-15)^2} = \sqrt{484+25} = 22.56$

5: Normal

Min values

Signal 5 is Normally classified

(b) Solution :- K = 3

$P_1 = \sqrt{(44-40)^2+(5.8-6)^2} = \sqrt{4+0.04} = 2.0099$

$P_2 = \sqrt{(44-44)^2+(5.8-6.11)^2} = \sqrt{\quad} = 0.31$

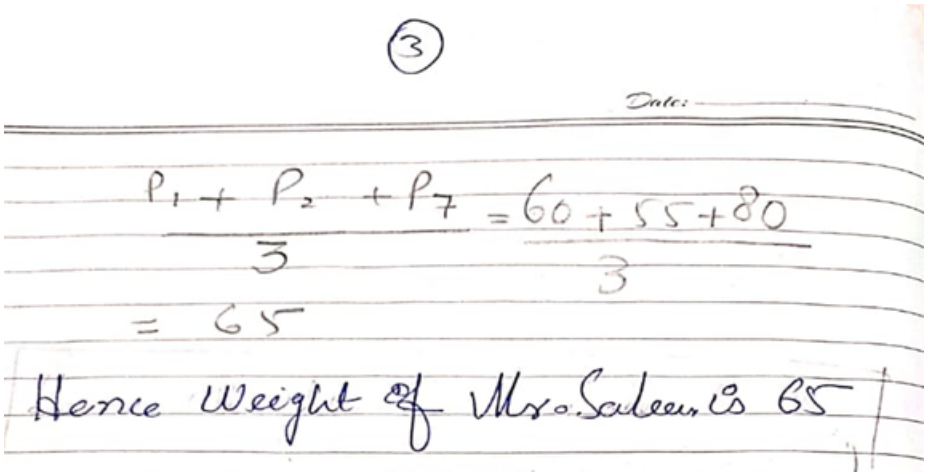
$P_3 = \sqrt{(44-36)^2+(5.8-5.9)^2} = \sqrt{64+0.01} = 8.0006$

$P_4 = \sqrt{(44-37)^2+(5.8-5.8)^2} = \sqrt{49+0} = 7$

$P_5 = \sqrt{(44-39)^2+(5.8-5.3)^2} = \sqrt{25+0.25} = 5.024$

$P_6 = \sqrt{(44-40)^2+(5.8-5.6)^2} = \sqrt{16+0.04} = 4.0049$

$P_7 = \sqrt{(44-42)^2+(5.8-5.9)^2} = \sqrt{4+0.09} = 2.022$



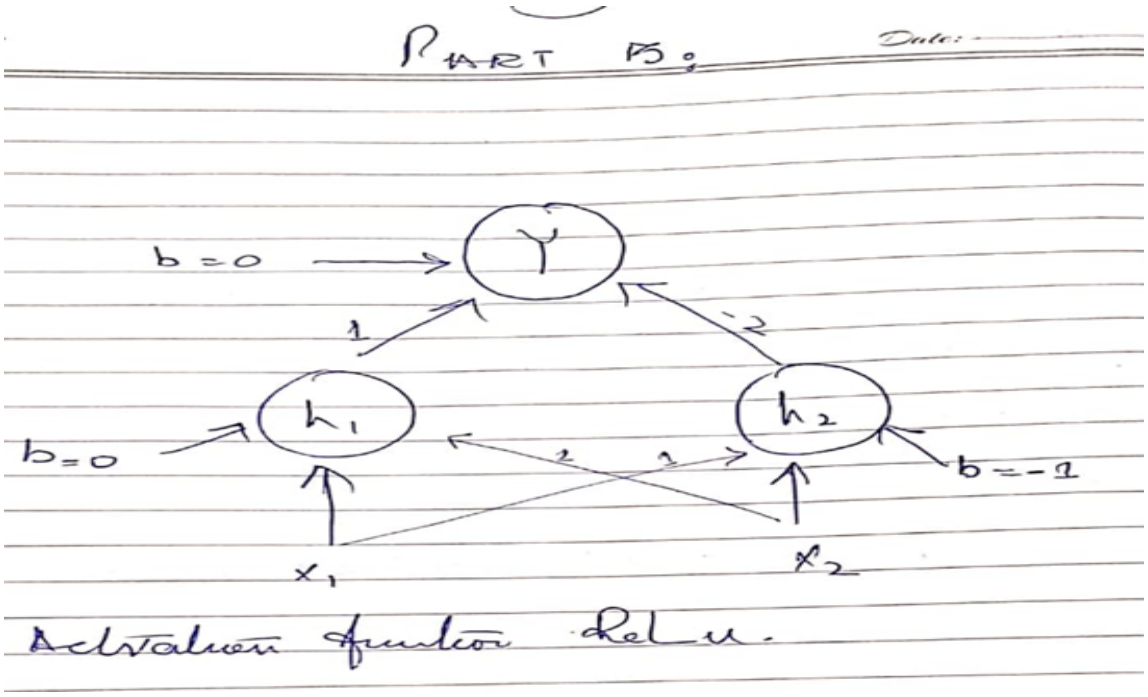
Q 4 a

The outputs of the hidden layer are

$$H_j = g\left(\sum_k W_{k,j} I_k\right) = c\sum_k W_{k,j} I_k + d$$

The final outputs are

$$O_i = g\left(\sum_j W_{j,i} H_j\right) = c\left(\sum_j W_{j,i}\left(c\sum_k W_{k,j} I_k + d\right)\right) + d$$



20.11 XOR (in fact any Boolean function) is easiest to construct using step-function units. Because XOR is not linearly separable, we will need a hidden layer. It turns out that just one hidden node suffices. To design the network, we can think of the XOR function as OR with the AND case (both inputs on) ruled out. Thus the hidden layer computes AND, while the output layer computes OR but weights the output of the hidden node negatively. The network shown in Figure S20.3 does the trick.

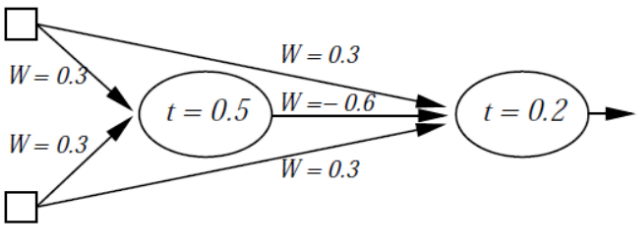


Figure S20.3 A network of step-function neurons that computes the XOR function.

Q No 4:- Solution Date:

PART C:

at h_3:

$3 * (-0.33) - 5 * 0.56 - 0.5 = -4.29$

After Activation function

$= 0.0135$

At h_4:

$3 * (-0.76) - 5 * (-0.1) + 0.3 = 3.08$

and after Activation fun

$= 0.956$

At h_5:-

$3 * 0.22 - 5 * 0.38 + 0.6 = -0.64$

After Activation

$= 0.35$

at Output node O

$(-0.8) \times 0.0135 + (0.85) * 3.08 - 0.5 * 0.64 + 0.25$

$= 3.18$

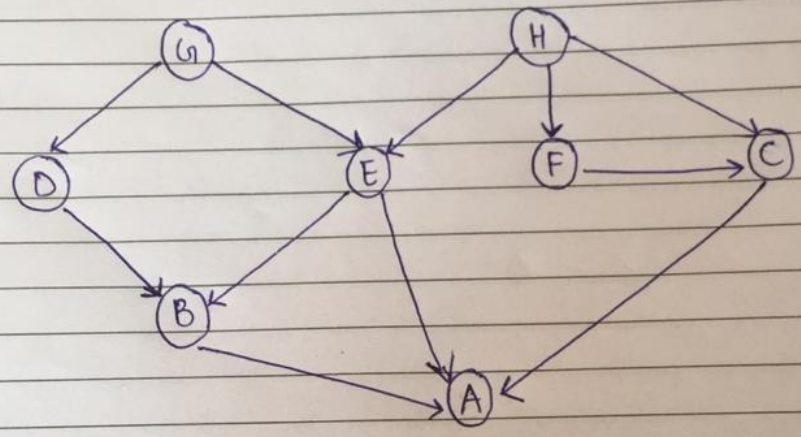
After Activation

$= 0.96$

Q 5

Q1) $P(A \mid B, C, E)\ P(B \mid D, E)\ P(C \mid F, H)\ P(D \mid G)\ P(E \mid G, H)$

$P(F \mid H)\ P(G)\ P(H)$



Q2) $P(OS = pass \mid ITC = pass) = \dfrac{P(OS = pass \mid ITC = pass)}{P(ITC = pass)}$

$= \dfrac{P(OS = pass \mid DS)\ P(DS \mid ITC = pass, OOP)\ P(OOP)\ \cancel{P(ITC = pass)}}{\cancel{P(ITC = pass)}}$

$= P(OS = pass \mid DS)\ P(DS \mid ITC = pass, OOP)\ P(OOP)$

DS	OOP	$P(OS = pass \mid DS)$	$P(DS \mid ITC = pass, OOP)$	$P(OOP)$	
pass	pass	$1 - 0.7 = 0.3$	0.2	0.4	0.024
pass	fail	$1 - 0.7 = 0.3$	0.3	0.6	0.054
fail	pass	$1 - 0.2 = 0.8$	$1 - 0.2 = 0.8$	0.4	0.256
fail	fail	$1 - 0.2 = 0.8$	$1 - 0.3 = 0.7$	0.6	0.336
					+ 0.67

$P(OS = pass \mid ITC = pass) = 0.67$ Answer.

M.M

Q 6:

Applications	Supervised LEARNING Methods With one line justification	Unsupervised Learning Methods With one line justification	Reinforcement Learning With one line justification
Hate Speech Detection	Yes		
Self-driving car			Yes
Rescues robots			Yes
Stock Market prediction	yes		
Bank transaction fraud detection	Yes		

1) $H(Pass) = -\frac{3}{6}\log\frac{3}{6} - \frac{3}{6}\log\frac{3}{6}$

$\Rightarrow 1$

2) $H(Pass \mid Grade) =$

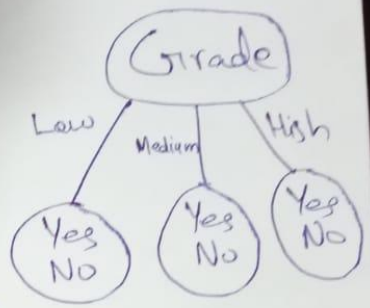
$H(Pass \mid Grade=Low) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$

$H(Pass \mid Grade=Medium) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$

$H(Pass \mid Grade=High) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$

Weighted Entropy $= (1)(2/6) + 1(2/6) + 1(2/6)$

Weighted Entropy $= 1$

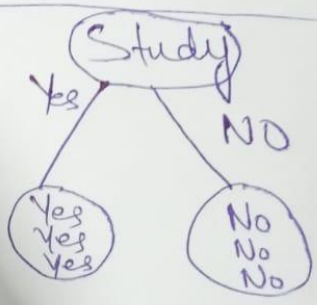


3) $H(Pass \mid Study) =$

$H(Pass \mid study=Yes) = -\frac{3}{3}\log\frac{3}{3} - \frac{0}{3}\log\frac{0}{3} = 0$

$H(Pass \mid Study=No) = -\frac{3}{3}\log\frac{3}{3} - \frac{0}{3}\log\frac{0}{3} 0$

weighted entropy $= 0$



4) "Study" will be the root node as it has lowest entropy and hence high info gain

Information gain of study = Parent entropy - weighted entropy

$= 1 - 0 = 1$

5) Final decision tree