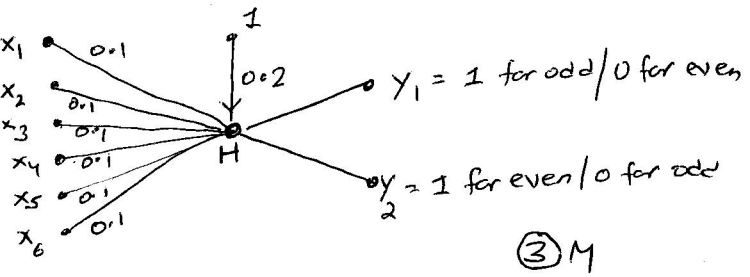
Q.2 (16 M)



(3) M

$x_1 x_2 x_3 x_4 x_5 x_6 = 101011$

$H_{in} = (0.1)(4) + 0.2 = 0.6$

$H_{out} = \dfrac{1 - e^{-2\times0.6}}{1 + e^{-2\times0.6}} = 0.5370$ (2M)

$Y_{1\text{-in}} = Y_{2\text{-in}} = 0.5370 \times 0.1 = 0.0537$

$Y_{1\text{-out}} = Y_{2\text{-out}} = \dfrac{1 - e^{-2\times0.0537}}{1 + e^{-2\times0.0537}} = 0.0536$ (2M)

error vector at output layer

$E_{2\text{-}1} = (1 - y_{2\text{-out}}^2)(y_{d1} - y_{2\text{-out}})$

$= (1 - .0536^2)(0 - .0536) = -0.0536$ (1.5M)

$E_{2\text{-}1} = (1 - y_{1\text{-out}}^2)(y_{d1} - y_{1\text{-out}})$

$= (1 - .0536^2)(1 - .0536) = 0.9436$ (1.5M)

$H = (1 - H_{out}^2)[E_{2\text{-}1} + E_{2\text{-}2}] \times 0.1$

$= (1 - .537^2)[.9436 - .0536] \times 0.1 = 0.0633$ (3M)

$W_{H\text{-}Y_1} = \eta E_{2\text{-}1} \cdot H_{out} = .9436 \times .5370$

$= 0.5067$ (1M)

$\Delta W_{H\text{-}Y_2} = \eta E_{2\text{-}2} \cdot H_{out} = -0.0536 \times .5370$

$= -0.2087$ (1M)

$\Delta b = \eta E_H \cdot 1 = 0.0633$ (1M)

Q.1

Note: Because of typo mistake $\Delta W = 0$ so marks for Q.1 reduced from 8 to 5

$x_1 = 1 \quad x_2 = 0$

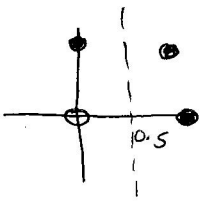
$net = 1 \times 1 + 0 \times 0 - 0.5 = 0.5$

$0 = f(net) = f(0.5) = 1$, desired $o/p = 1$

$\Delta W = \eta(t - o) = 0$ so no change in weight (3M)

separating line

$x_1 - 0.5 = 0$ i.e $x_1 = 0.5$



(2M)

(a)

$W = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}[0\ 1] + \begin{bmatrix}-1\\1\\1\\1\end{bmatrix}[1\ 0] + \begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}[1\ 1]$

$= \begin{bmatrix}0 & 2\\0 & 0\\0 & 0\\2 & 2\end{bmatrix}$ (4M)

(b)

S_1, mistaken last bit $= [1\ 1\ 1\ -1]$

$o/p = [1\ 1\ 1\ -1]\begin{bmatrix}0 & 2\\0 & 0\\0 & 0\\2 & 2\end{bmatrix} = [-2\ 0]$

$f[-2\ 0] = [0\ 1]$ correct (2M)

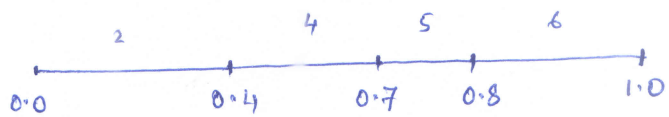
(c) $[1\ 1]\begin{bmatrix}0 & 0 & 0 & 2\\2 & 0 & 0 & 2\end{bmatrix} = [2\ 0\ 0\ 4]$

$f[2\ 0\ 0\ 4] = [1\ -1\ -1\ 1]$

it is BAM (2M)

Q4. $f(x_1, x_2) = 1 + x_1^2 + x_1 x_2$, $x_1, x_2 \in [0, 10]$

	GEN #1.	Fitness	Rank	Selection Prob.	Mating Pool		GEN #2	fitness
		83.50 → E2.	—	—	—		8.3 | 3.8	101.43
1.>	5.5 | 9.5						5.5 | 9.5	83.50
2.>	7 | 4.1	78.70	4	0.4	—		8.65 | 5.15	120.37
3.>	8.3 | 3.8	101.43 → E1	—	—	7 | 4.1	(2)	5.35 | 3.05	45.94
4.>	3.7 | 2	22.09	3	0.3	3.7 | 2	(4)	9 | 4.65	123.85
5.>	1.2 | 4.9	8.32	1	0.1	3 | 3	(6)	5 | 3.55	43.75
6.>	3 | 3	19.00	2	0.2	7 | 4.1	(2)		518.84
		313.04	10	1.0				



Pair 1 → $P_1 = \begin{bmatrix} 7 \\ 4.1 \end{bmatrix}$, $P_2 = \begin{bmatrix} 3.7 \\ 2 \end{bmatrix}$.

$C_1 = 0.5 P_1 + 0.5 P_2 = \begin{bmatrix} 5.35 \\ 3.05 \end{bmatrix}$ → 45.94 (fitness)

$C_2 = 1.5 P_1 - 0.5 P_2 = \begin{bmatrix} 8.65 \\ 5.15 \end{bmatrix}$ → 120.37

$C_3 = -0.5 P_1 + 1.5 P_2 = \begin{bmatrix} 2.05 \\ 0.95 \end{bmatrix}$ → 7.15

C2 & ~~C3~~ C1 taken (best two)

Pair 2 → $P_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $P_2 = \begin{bmatrix} 7 \\ 4.1 \end{bmatrix}$.

$C_1 = 0.5 P_1 + 0.5 P_2 = \begin{bmatrix} 5 \\ 3.55 \end{bmatrix}$ → 43.75 (fitness)

$C_2 = 1.5 P_1 - 0.5 P_2 = \begin{bmatrix} 1 \\ 2.45 \end{bmatrix}$ → 4.45

$C_3 = -0.5 P_1 + 1.5 P_2 = \begin{bmatrix} 9 \\ 4.65 \end{bmatrix}$ → 123.85

∴ C_3 & C_1 taken. (best two)

Marking Scheme :

- Identifying Elites in GEN #1 → 2 M
- Rank & Selection Prob. computation → 2 M
- Forming the Mating Pool → 4 M
- Obtaining 4 children in GEN# 2 → 8 M

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

Neural Networks & Fuzzy Logic (BITS F312) [1st Semester, 2019-2020]

Mid-Semester Exam - Part A (closed book)

Max Time- 40 min | Max Marks - 45 | Date: 03.10.2019

NAME	ID NO.

Answer in the space provided. You may use the last page of this question paper for rough work.

1. Name a person associated with publishing first time use of Backpropagation for learning word embeddings. GEOFFREY HINTON [1]

2. Write name of the humanoid robot who has been granted citizenship of Saudi Arabia in October 2017. SOPHIA [1]

3. The crux of the test, named after someone who is considered to be the father of theoretical computer science and AI is: Can a machine fool a human into thinking they are chatting with another person? Name the person. [1]
ALAN TURING

4. Name the network based on the concept "once a neuron repeatedly excited another neuron, the threshold of excitation of the latter decreased". HEBB [1]

5. A 25x25 Associative network is able to store maximum how many orthogonal bipolar vectors of 25 dimensions? 24 [1]

6. In which kind of learning the program is provided feedback in terms of rewards and punishments as it navigates its problem space? REINFORCEMENT [1]

7. What is the name of the term coined by Statistician Nassim Nicholas Taleb for an event which is an outlier (it lies outside the realm of regular expectations)? [1]
BLACK SWAN

8. Which performance term in classification performance is 100 % when it totally avoids false alarms? SPECIFICITY [1]

9. For a test having perfect diagnostic ability, the area under ROC curve would be how much? [1]
1.0

10. In which activation function slope of the output for negative inputs is a learnable parameter? [1]
PReLU

11. Identify the term which tunes or selects the preferred level of model complexity so your models are better at predicting (generalizing). REGULARISATION [1]

12. For multi –class classification of images, which activation function is applied at the last layer of deep neural networks? SOFTMAX [1]

13. The single layer perceptron is a supervised learning algorithm that only works on what kind of data? LINEARLY SEPARABLE [1]

14. Which term used in BPA would result in faster convergence if it is used to upgrade weights based on the basic concept of changing the weight in a direction that is a combination of the current gradient and previous gradient? MOMENTUM [1]

15. When classification accuracy on the validation data computed at the end of each epoch saturates, then training is stopped to avoid overfitting. What is the name of this strategy? [1]

EARLY STOPPING

16. Anytime an algorithm is trying to fit noise in addition to the pattern, it is exhibiting what phenomenon? OVERFITTING [1]

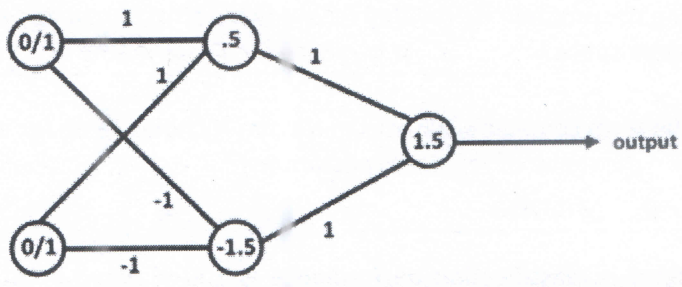
17. Which kind of network architecture would have weight connections from the input to each layer and from each layer to the successive layers? CASCADE FEEDFORWARD [1]

18. GAN is made of two networks, one is Generator. Write the name of another network. [1]

DISCRIMINATOR

19. Other than CPUs & GPUs, high performance computing machines for deep learning architectures use which processor units used as coprocessor? TPU [1]

20. Figure below shows an MLP with threshold given inside nodes in hidden and output layers. Activation function with threshold given is unipolar. Inputs and outputs are binary.



(a) Based on inputs and output of a particular node, what kind of Gate is being realized by (i) Upper node of hidden layer (ii) Lower node of hidden layer (ii) Output node?

(i) OR

(ii) NOT AND

(iii) AND

(b) The relation between inputs and output is governed by what Gate?

XOR [4]

21. The prevalence of certain disease is 0.2% . Probability of (Screen Positive | Disease) = 0.85, the probability of screening positive overall is 8% . Using a hypothetical population of 1,00,000 people, write value of all the enteries in the CONFUSION MATRIX. [12]

	Diseased	Not Diseased	
Test +	A = 170	B = 7,830	G = 8,000
Test -	C = 30	D = 91,970	H = 92,000
	E = 200	F = 99,800	T = 1,00,000

22. The objective is to determine the shortest distance between the straight line x+y=5 and the rectangular hyperbola xy=10. Write the expression of the corresponding Lagrangian, L. [3]

$(x_1 - x_2)^2 + (y_1 - y_2)^2 + \lambda_1 (x_1 + y_1 - 5) + \lambda_2 (x_2 y_2 - 10)$

23. Random walk method is inapplicable if the objective function has a discontinuity in the search space. True or False? FALSE [1]

24. In PSO algorithm, velocity update equation has the following three components:

INERTIA , cognitive and SOCIAL [2]

25. A craziness component is sometimes added to the standard PSO algorithm to achieve faster convergence. True or False? FALSE [1]

26. Making the three components in the velocity update equation in PSO independent of one another will lead to faster convergence. True or False? FALSE [1]

27. The probability that the GA string 10100101 will survive bit-wise mutation with P_m=5%, is

0·66 [2]

***