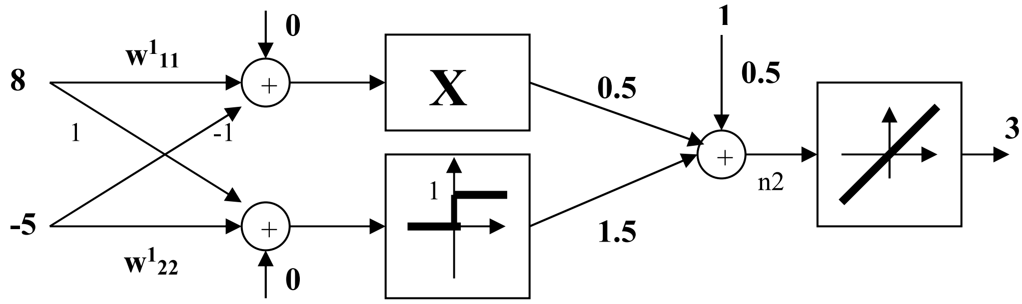
DEI/FCTUC
Machine Learning/Neuronal and Fuzzy Computation

Duration: **90 minutes**

Exam with consultation, except past exams and its solutions. It is not allowed the use of computers.

---

1. Consider the neural network of the figure with arbitrary weights w^1_{22} and w^2_1 to be computed in each case in order to obtain the desired output.



a) Which of the following activation functions are possible in the neuron **X** ?

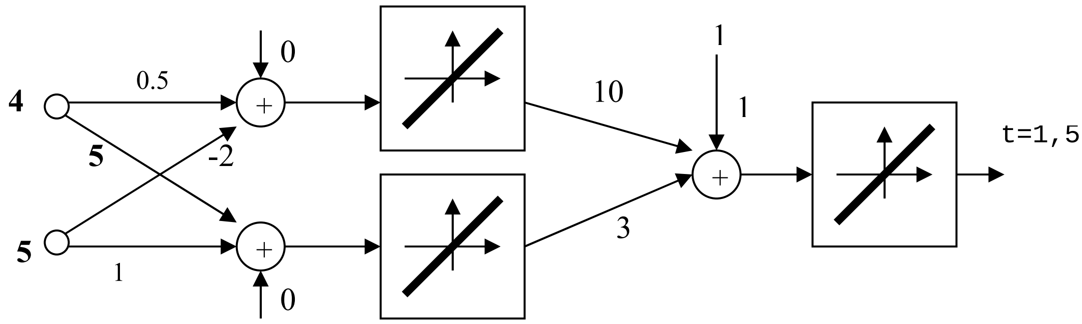
Function	Yes or No	If yes, compute a set of values for these weights. If not, justify.
Logsig	No	From X must come 2, in order to obtain 2x0.5+1.5+0.5=3.
Hardlims (symmetric)	No	From X must come 2, in order to obtain 2x0.5+1.5+0.5=3.
Purlin	Yes	$8-5w^1_{22}+0>0$ $8w^1_{11}+5+0=2$

b) For the possible cases compute a set of values for w^1_{11} and w^1_{22} allowing to obtain the desired result.

2. Consider four objects in the bidimensional space represented by the following four points given by their coordinates: O1 (2,-4), O2 (-1,2), O3 (1,3), O4 (-2,1). Project a binary perceptron to classify these objects into four classes, using a minimum number of neurons. Draw the complete perceptron and write the class code.

Use the delivered millimetric paper.

3. a) Make a complete iteration of the backpropagation algorithm to adjust the weight w^1_{21}, starting from its initial value 5. The target for the input is -6. Comment your results.



$F = e^2$ (objective function, criteria) $\qquad a = n^2$

$$\mathrm{w}^1_{12}(k+1) = \mathrm{w}^1_{12}(k) - \alpha \frac{\partial F}{\partial w_{12}}\Big|_k \qquad n^2 = a_1^1 \times 2 + a_2^1 \times 1 + 1$$

$$n_1^1 = 2\times(-3) + 5w_{12}^1 + 0$$

$$a_1^1 = n_1^1$$

1st step, forward

$n_1^1 = -3\times 2 - 2\times 5 + 0 = -16 \qquad a_1^1 = -16 \qquad n^2 = -2\times 16 + 1\times 26 + 1 = -5$

$n_2^1 = 3\times 2 + 4\times 5 + 0 = 26 \qquad a_2^1 = 26$

2nd step, calculation of the error: $e = t - a = 1,5 - (-5) = 6,5$

3rd step: retropropagation of the error and updating of the weight

$$\frac{\partial F}{\partial w^1_{12}} = \frac{\partial F}{\partial e} \times \frac{\partial e}{\partial a^2} \times \frac{\partial a^2}{\partial n^2} \times \frac{\partial n^2}{\partial a_1^1} \times \frac{\partial a_1^1}{\partial n_1^1} \times \frac{\partial n_1^1}{\partial w_{12}^1}$$

$$= 2e \times (-1) \times 1 \times 2 \times 1 \times 5$$

$$\frac{\partial F}{\partial w^1_{12}} = -20e = -130$$

$$w_{12}^1(1) = w_{12}^1(0) - \alpha \frac{\partial F}{\partial w^1_{12}}\Big|_0 = -2-(-130) = 128.$$

(for a learning coefficient equal to 1).

b) Discuss the advantages and disadvantages of the gradient method in the backpropagation algorithm.

4. The following Matlab code is intended to configure a CNN for classification of a dataset into a certain number of classes.

```
Layers = [imageInputLayer([42 42 1])
          convolution2dLayer(3, 1)
          maxPooling2dLayer (2, 'Stride',2)
          fullyConnectedLayer (4)
          softmaxLayer
```

classificationLayer];

a) How many classes exist in the problem ? Why ?
   Ans: 4 because the fully connected layer has 4 neurons.

b) Which is the number of inputs and outputs of the fully connected layer ? Justify in detail.

   42x42 with a filter 3x3 and stride =1 gives a feature matrix 40x40. Now with average pooling 2x2 with stride 2 gives a reduced feature matrix 20x20, that is given to the fully connected layer; so this layer has 400 inputs and 4 outputs.

5. A system has two inputs P1 e P2 and one output S1, with the following scales:

P1: [0 10]   P2: [-5 5]   S: [-2 2]

Define in the normalized interval [-1 1] three triangular fuzzy sets, with an adequate coverage property, for each of the variables (inputs and output). Label them as Negative (N) , Zero (Z), Positive (P).

Consider the following Mamdani rules:

Rule 1: IF P1 is Negative AND P2 is Zero THEN S1 is Negative $(N \wedge Z \rightarrow N)$
Rule 2: IF P1 is Negative AND P2 is Positive THEN S1 is Zero $(N \wedge P \rightarrow Z)$
Rule 3: IF P1 is Zero and P2 is Positive THEN S1 is Positive $(Z \wedge P \rightarrow P)$

Consider now the measured inputs

Input P1= 2,0   Input P2 = 1,5

Compute the fuzzy output of each rule and the aggregated fuzzy output, using the minimum for logic AND in the antecedents, and the Mamdani implication for inference. For defuzzification, compute it taking into account the shape of the aggregated fuzzy output and the meaning of the defuzzification operation.

Answer:

Rule 1 $(N \wedge Z \rightarrow N)$

For P1=2 (-0,6 normalized), the membership to N is 0.6
For P2=1.5 (0.32 normalized), the membership to Z is 0.7   minimum =0.6
The output of the rule is a negative trapezoid with height 0.6 resulting from min (0.6, 0.7)

Rule 2 $(N \wedge P \rightarrow Z)$

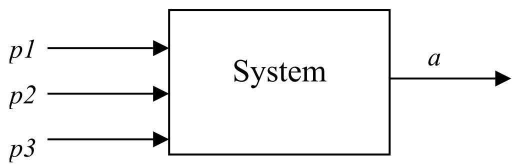
For P1=2 (-0.6 normalized), membership to a N is 0.6
For P2=1.5 (0.32 normalized), membership to P é 0.3   minimum = 0.3
The output of the rule is a zero trapezoid triangle with height 0.3.

Rule 3 ($Z \wedge P \rightarrow P$)

For P1=2 (-0,6 normalized), membership to Z is 0.6
For P2=1.5 (0.32 normalized), membership to P é 0.3   minimum =0.3
The output of the rule is a triangle of height 0.2 resulting from the product of P by 0.2.

The total output are two triangles resulting from the rule 1 and rule 3. If we apply the defuzification by the height, it becomes zero, normalized, or 0 denormalized; i.e, the total output is 0.

6. A system has three inputs (*p1, p2, p3*) and one output *a.*



A dataset from the input-output space was collected, then normalized in [-1; 1] in each dimension; after the subtractive clustering algorithm was applied, resulting into 3 clusters with centres:

C1=[-0.9; 1; -0.9; 0,5], C2=[0,9; -0,8; 0,2 ; -0,2], C3=[0; 0.02; 1; -0,8].

Write a set of fuzzy rules, of TSK zero order type, to model this system, using the clusters information, and defining appropriately the fuzzy sets in the universe of discourse of each variable. Use Gaussian membership functions in the antecedents with the same openness. Use the millimetric paper to draw them for the three inputs and the output.

~~**7.** Compare sucintamente e criticamente a aplicabilidade e a utilidade das redes neuronais artificiais e dos sistemas difusos para extrair conhecimento a partir de dados genéricos.~~