

Question1: Choose the most appropriate answer (only one):

يتم اختيار الاجابة الصحيحة في ورقة الاسئلة.

تلقى درجة السؤال إذا تم اختيار أكثر من إجابة، وإذا حدث ذلك من دون قصد أكتب حرف الاجابه الصحيحة أمام السؤال في ورقة الاسئلة.
يتم حل جميع المسائل الرقمية في كراسة الاجابة واختيار الاجابة النهائية في ورقة الاسئلة، ولن تحسب درجة السؤال بغير ذلك.

1- What is a decision function?

- A. A function $d_j(x)$ which gives the correct answer in a pattern recognition problem with M classes.
- B. A function $d_j(x)$ which is defined for each class in a classification problem, so that an input pattern x is assigned to class k if $d_k(x) = \max\{d_1(x); \dots; d_M(x)\}$.
- C. A function $d_j(x)$ which specifies the class conditional probability $P(x|k)$ of an input example x given that it belongs to class j.

2- Which of the following statements is NOT true for a minimum distance classifier (MDC)?

- A. The MDC is specified completely by the prototype vectors for all M classes in a classification problem.
- B. The MDC minimizes the average loss of misclassifying the input patterns.
- C. The MDC is a special case of the Bayes optimal classifier.
- D. The MDC makes a decision boundary between two classes. This boundary is a line, plane or hyperplane where the two decision functions are equal.

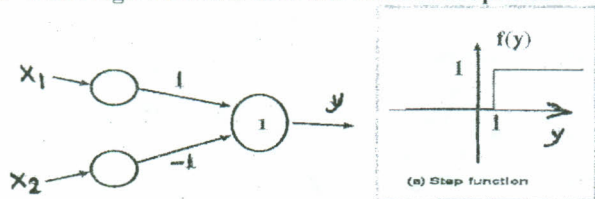
3- Many pattern recognition problems require the original input variables to be combined together to make a smaller number of new variables. These new input variables are called

- A. patterns
- B. features
- C. weights
- D. classes.

4- The process described in question 3 is:

- A. a type of pre-processing which is often called feature extraction.
- B. a type of pattern recognition which is often called classification.
- C. a type of post-processing which is often called winner-takes-all.

5- What logic function does this network represent?



- A. AND gate.
- B. NOR gate.
- C. AND NOT gate.
- D. XOR gate
- E. None of the above

6- The weight vectors of three processing units are given as follows:

$w_1 = [-1 \ -1.5 \ 0.5]^t$ $w_2 = [2 \ -2 \ 5.2]^t$ $w_3 = [1.5 \ 6 \ 4.3]^t$

An input vector $x = [-1.4 \ 2.3 \ 0.2]^t$ is presented to the network. The winning neuron has an output =

- A. $O_1 = 12.12$
- B. $O_3 = 12.12$
- C. $O_1 = 6.3$
- D. $O_3 = -1.3$

7- Adapt the weight vector of the winning unit in previous question with a learning rate of 0.5. What is the new weight vector?

- A. $w_{winner} = [-2.7 \ 4.65 \ 0.4]^t$
- B. $w_{winner} = [-1.2 \ 0.4 \ 0.35]^t$
- C. $w_{winner} = [0.05 \ 4.15 \ 2.25]^t$
- D. $w_{winner} = [0.3 \ 0.15 \ 2.7]^t$

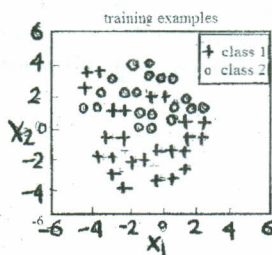
8- An artificial neural network may be trained on one data set and tested on a second data set, why?

- A. The error on the second data set provides a better (unbiased) estimate of the true generalization error.
- B. The error on the second data set is used to choose between lots of different possible systems.
- C. It's not important as training on the first data set indicates the generalization performance of the system.

9- What is classification?

- A. Deciding which features to use in a pattern recognition problem.
- B. Deciding which class an input pattern belongs to.
- C. Deciding which type of neural network to use.

10- A perceptron is trained on the data shown below, which has two classes (shown by the symbols '+' and 'o'). After how many iterations of training, the perceptron will converge and the decision line will reach a steady state?"



- A. few iterations.
- B. Large number of iterations.
- C. It'll not converge.

11-What is the biggest difference between Widrow & Hoff's Rule and the Delta Learning Rule for learning in a single-layer feedforward network?

- A. The Delta Rule is defined for step activation functions, but the Widrow & Hoff's Learning Rule is defined for linear activation functions.
- B. The Delta Rule is defined for sigmoid activation functions, but the Widrow & Hoff's Learning Rule is defined for linear activation functions.
- C. The Delta Rule is defined for linear activation functions, but the Widrow & Hoff's Learning Rule is defined for step activation functions.

12- Is the following statement true or false? "The XOR problem can be solved by a multi-layer perceptron, but a multi-layer perceptron with bipolar step activation functions cannot learn to do this."

- A. TRUE
- B. FALSE.

13- Which of the following statements are true for typical neurons in the human brain?

- A. Electrical potential is summed in the neuron.
- B. When the potential is bigger than a threshold, the neuron fires a pulse through the axon.
- C. The neurons are connected to each other by axons, synapses and dendrites.
- D. All of the above answers.

14- If a perceptron is given a training set which is linearly separable, then what will happen during training?

- A. The weights of the network will reach a steady state, where the decision boundary gives the best possible separation of the data.
- B. The weights of the network will reach a steady state, but the decision line may give a sub-optimal (almost optimal) separation of the data.
- C. The weights of the network will not reach a steady state, and the decision line will not stop moving.
- D. The weights of the network will not reach a steady state, therefore a decision line cannot exist.

15- Regions where no class membership of an input pattern can be uniquely determined based on the response of the classifier.

- A. Decision boundaries
- B. Indecision regions
- C. Decision hyperplane
- D. All

16- Assume continuous unipolar sigmoid func. in next fig.(part of backpropagation network), $\delta_A = ??$

- A. 0.0468
- B. 0.055
- C. 0.034
- D. 0.05

17- What is the value of $\delta_c = ??$

- A. 0.19
- B. 0.046
- C. 0.09
- D. 0.22

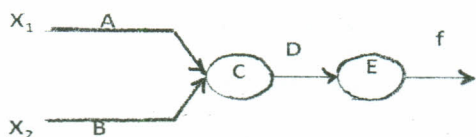
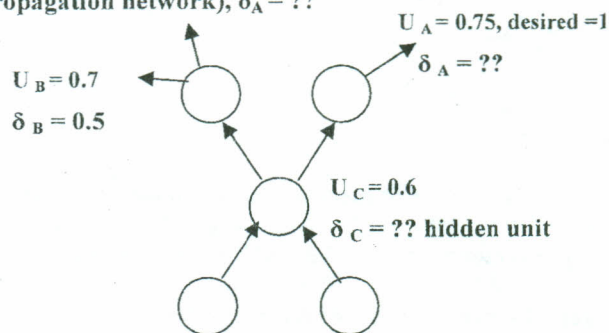


fig.4

18-For the McCulloch-Pitts Neuron model shown in the Fig.4, to get the function the weights and the thresholds A,B,C,D,E are:

- A. 1, 1, 1,-1, 0
- B. -1, 1,1,1,1
- C. 1, 0,1,1,1
- D. 1, 1, 2,-1, 0

19- What is the credit assignment problem in the training of multi-layer feedforward networks?

- A. The problem of adjusting the weights for the output layer.
- B. The problem of defining an error function for linearly inseparable problems.
- C. The problem of avoiding local minima in the error function.
- D. The problem of adjusting the weights for the hidden layers.

20- A perceptron with a unipolar binary function has two inputs with weights $w_1 = 0.2$ and $w_2 = -0.5$, and $w_3 = -0.2$ (where w_3 is the weight for bias which is -1). For a given training example $x = [0, 1]^T$, the desired output is 1. Using the learning rule $c = (d - y)x$, where learning rate = 0.2. What are the new values of the weights and threshold after one step of training with the input?

- A. $w_1 = 0.2, w_2 = -0.3, w_3 = -0.4$.
- B. $w_1 = 0.4, w_2 = -0.5, w_3 = -0.4$.
- C. $w_1 = 0.2, w_2 = -0.5, w_3 = -0.2$.
- D. $w_1 = 0.2, w_2 = -0.3, w_3 = 0$.

21- What is a feature in a pattern classification problem?

- A. An output variable. B. An input variable. C. A hidden variable or weight.

22- What is a decision boundary in a pattern classification problem with two input variables?

- A. A histogram defined for each image in the training set.
 B. A line or curve which separates two different classes.
 C. A plane or hypercurve defined in the space of possible inputs.

23- Which of the following statements is true for the winner takes all learning method?

- A. The learning rate is a function of the distance of the adapted units from the winning unit.
 B. The weights of the winning unit k are adapted by $\Delta w_k = \eta (x - w_k)$, where x is the input vector.
 C. The weights of the neighbours j of the winning unit are adapted by $\Delta w_j = \eta_j (x - w_j)$, where " $\eta_j < \eta$ " and $j \neq k$.

24- Is the following statement true or false? "Clustering can be useful when we want to analyse a large number of pattern vectors and identify groups of patterns with similar features."

- A. TRUE. B. FALSE.

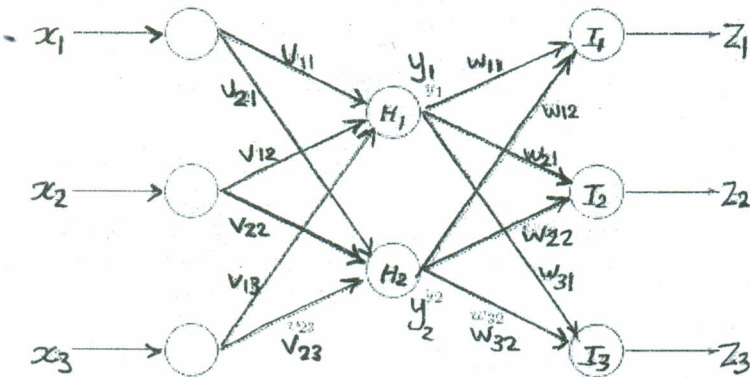
25- Which of the following statements is the best description of Hebbian's learning rule?

- A. "If a particular input stimulus is always active when a neuron fires then its weight should be increased."
 B. "If a stimulus acts repeatedly at the same time as a response then a connection will form between the neurons involved. Later, the stimulus alone is sufficient to activate the response."
 C. "The connection strengths of the neurons involved are modified to reduce the error between the desired and actual outputs of the system."

26- Is the following statement true or false? "A multi-layer feedforward network with linear activation functions is more powerful than a single-layer feedforward network with linear activation functions."

- A. TRUE. B. FALSE.

27- A training pattern, consisting of an input vector $x = [x_1 \ x_2 \ x_3]^T$ and desired outputs $t = [t_1 \ t_2 \ t_3]^T$, is presented to the following neural network. What is the usual sequence of events for training the network using the backpropagation algorithm?



- A. (1) calculate $z_k = f(I_k)$, (2) update w_{kj} , (3) calculate $y_j = f(H_j)$, (4) update v_{ji} .
 B. (1) calculate $y_j = f(H_j)$, (2) update v_{ji} , (3) calculate $z_k = f(I_k)$, (4) update w_{kj} .
 C. (1) calculate $y_j = f(H_j)$, (2) calculate $z_k = f(I_k)$, (3) update v_{ji} , (4) update w_{kj} .
 D. (1) calculate $y_j = f(H_j)$, (2) calculate $z_k = f(I_k)$, (3) update w_{kj} , (4) update v_{ji} .

28- After training, the units in the neural network of previous question have the following weight vectors:

$$v_1 = \begin{bmatrix} -2.0 \\ 2.0 \\ -2.0 \end{bmatrix}, v_2 = \begin{bmatrix} 1.0 \\ 1.0 \\ -1.0 \end{bmatrix}, w_1 = \begin{bmatrix} 1.0 \\ -3.5 \end{bmatrix}, w_2 = \begin{bmatrix} 0.5 \\ -1.2 \end{bmatrix} \text{ and } w_3 = \begin{bmatrix} 0.3 \\ 0.6 \end{bmatrix}$$

Assume that all units have unipolar sigmoid activation functions. If the network is tested with an input vector $x = [2 \ 3 \ 1]^T$, then the output of the second hidden neuron y_2 will be

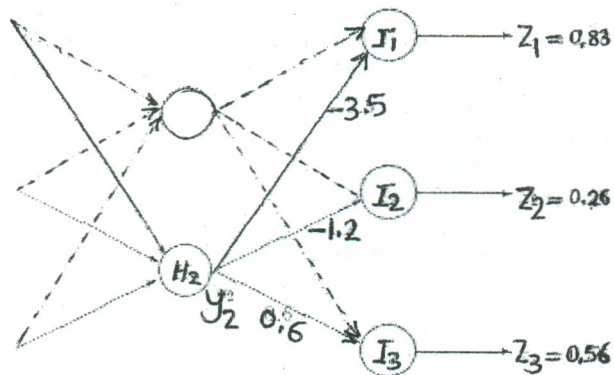
- A. 0.018 B. 0.500 C. 0.982 D. 1 E. 4

29- For the same neural network described in previous question(q27-q28), the output of the second output neuron z_2 will be (assuming same activation function, input and weights)

- A. 0.715 B. -0.928 C. -0.087 D. 0.283 E. 0.527



d- The following figure shows part of the neural network described in questions 27 and 28. In this question, a new input pattern is presented to the network and training continues as follows. The actual outputs of the network are given by $z = [0.83 \ 0.26 \ 0.56]^T$ and the corresponding target outputs are given by $t = [0.58 \ 0.70 \ 0.20]^T$. The weights w_{12} , w_{22} and w_{32} are also shown below on figure. What is the error for each of the output units?



- A. $\delta_{\text{output 1}} = -0.0425$, $\delta_{\text{output 2}} = 0.3256$, and $\delta_{\text{output 3}} = -0.1584$.
- B. $\delta_{\text{output 1}} = -0.0353$, $\delta_{\text{output 2}} = 0.0847$, and $\delta_{\text{output 3}} = -0.0887$.
- C. $\delta_{\text{output 1}} = 0.0425$, $\delta_{\text{output 2}} = -0.3256$, and $\delta_{\text{output 3}} = 0.1584$.
- D. $\delta_{\text{output 1}} = 0.0353$, $\delta_{\text{output 2}} = -0.0847$, and $\delta_{\text{output 3}} = 0.0887$.

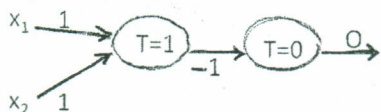
31- For the hidden units of the same previous network q 30, What is the error for hidden unit 2 given that its activation for the pattern being processed is currently $f(y_2) = 0.69$?

- A. $\delta_{\text{hidden 2}} = -0.0721$
- B. $\delta_{\text{hidden 2}} = -0.0067$
- C. $\delta_{\text{hidden 2}} = 0.0067$
- D. $\delta_{\text{hidden 2}} = 0.0721$

32- Which of the following applications is the most obvious example of a classification problem?

- A. Stock market prediction.
- B. Box pushing with a mobile robot.
- C. Handwritten character recognition
- D. Learning to steer the ALVINN autonomous vehicle.

33- Output "O" of next McCulloch model represents:



- A. OR logic gate
- B. AND logic gate
- C. NOR logic gate
- D. NAND gate

34- A multi-layer feedforward network has 5 input units, one hidden layer with 4 units, and 3 output units. How many weights does this network have?

- A. 12
- B. 20
- C. 27
- D. 32
- E. 40

35- A perceptron with a unipolar step function has two inputs with weights $w_1 = 0.2$ and $w_2 = -0.5$ and threshold $= -0.2$ (i.e. weight for an extra input $= -1$). For a given training example $x = [1 \ 1]^T$, the desired output is zero. Does the perceptron give the correct answer (i.e. no error)?

- A. True
- B. False

36- The weights of the ANN to be trained are typically initialized atvalues.

- A. Small statistical
- B. Small random
- C. Large statistical
- D. Large random

37- The factor that strongly affects the ultimate solution of an ANN is:

- A. Initial weights
- B. Learning Factor
- C. Steepness Factor
- D. All

38- One of the problems of backpropagation algorithm is that the error minimization procedure may produce only a.....of the error function.

- A. Global minimum
- B. Local minimum
- C. Local maximum
- D. B or C

39-For comparison of networks with different numbers of training patterns P and having a different number of output neurons, it's fair enough to use:

- A. Quadratic error.
- B. Root mean square normalized error.
- C. Decision error.
- D. All previous answers.

*****End of questions*****