Artificial Neural Networks
Examination, June 2005

Instructions

There are SIXTY questions. (The pass mark is 30 out of 60).

For each question, please select a maximum of ONE of the given answers (either A, B, C, D or E). You should select the one answer that represents the BEST possible reply to the question (in some cases, there may be no obvious "wrong" answers, but one answer should always be better than the others). Every time you select the correct answer, you will be awarded +1 mark. However, every time you select an incorrect answer, a penalty score will be subtracted from your total mark. This penalty depends on the number of possible answers to the question, as follows:

Number of possible answers	Score for correct answer	Score for incorrect answer
2	+1	-1
3	+1	$-\frac{1}{2}$
4	+1	$-\frac{1}{3}$
5	+1	$-\frac{1}{4}$

If you do not give any answer to a question, no marks will be added to your total score and there will be no penalty. If you give more than one answer to a question, this will be counted as an *incorrect* answer. So please be *very* careful, and make sure that ONLY one letter (A or B or C or D or E) is visible in each of your written answers. Please use the exam answer sheet provided and write your answers very clearly so that they can be read by an average examiner!

Advice: read all of the questions before you start to answer.
Tools required: calculator.

Questions

1 Pattern Recognition

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.

2 Pattern Recognition

What is a *pattern vector*?

A. A vector of weights $\mathbf{w} = [w_1, w_2, ..., w_n]^T$ in a neural network.

B. A vector of measured features $\mathbf{x} = [x_1, x_2, ..., x_n]^T$ of an input example.

C. A vector of outputs $\mathbf{y} = [y_1, y_2, ..., y_n]^T$ of a classifier.

3 Classical Pattern Recognition

For a minimum distance classifier with one input variable, what is the *decision boundary* between two classes?

A. A line.

B. A curve.

C. A plane.

D. A hyperplane.

E. A discriminant.

4 Classical Pattern Recognition

For a Bayes classifier with two input variables, what is the *decision boundary* between two classes?

A. A line.

B. A curve.

C. A plane.

D. A hypercurve.

E. A discriminant.

5 Classical Pattern Recognition

Design a minimum distance classifier with three classes using the following training data:

$$\text{Class 1: } \begin{bmatrix} -1.0 \\ -0.5 \end{bmatrix}, \begin{bmatrix} -1.0 \\ -1.5 \end{bmatrix} \qquad \text{Class 2: } \begin{bmatrix} 1.0 \\ 0.5 \end{bmatrix}, \begin{bmatrix} 1.0 \\ -0.5 \end{bmatrix} \qquad \text{Class 3: } \begin{bmatrix} -1.0 \\ 0.5 \end{bmatrix}, \begin{bmatrix} -1.0 \\ 1.5 \end{bmatrix}$$

Then classify the test vector $[0.5, -1]^T$ with the trained classifier. Which class does this vector belong to?

A. Class 1.

B. Class 2.

C. Class 3.

6 Classical Pattern Recognition

The decision function for a minimum distance classifier is $d_j(\mathbf{x}) = \mathbf{x}^T\mathbf{m}_j - \frac{1}{2}\mathbf{m}_j^T\mathbf{m}_j$ where $\mathbf{m}_j$ is the prototype vector for class ω_j. What is the value of the decision function for each of the three classes in question 5 for the test vector $[0, -0.5]^T$?

A. $d_1(\mathbf{x}) = -1.5$, $d_2(\mathbf{x}) = -0.5$, $d_3(\mathbf{x}) = -0.5$.

B. $d_1(\mathbf{x}) = -0.875$, $d_2(\mathbf{x}) = -0.375$, $d_3(\mathbf{x}) = -2.375$.

C. $d_1(\mathbf{x}) = -0.5$, $d_2(\mathbf{x}) = -0.5$, $d_3(\mathbf{x}) = -1.5$.

D. $d_1(\mathbf{x}) = -0.375$, $d_2(\mathbf{x}) = -0.875$, $d_3(\mathbf{x}) = -0.875$.

7 Classical Pattern Recognition

Give the equation of the decision line between classes 2 and 3 for the minimum distance classifier in question 5.

A. $x_2 = 0$

B. $x_2 = -2x_1 - \frac{1}{2}$

C. $x_2 = 2x_1 + \frac{1}{2}$

8 Classical Pattern Recognition

Give the equation of the decision line between classes 1 and 3 for the minimum distance classifier in question 5.

A. $x_2 = 0$

B. $x_2 = -2x_1 - \frac{1}{2}$

C. $x_2 = 2x_1 + \frac{1}{2}$

9 Training and Testing

Is the following statement true or false? "An outlier is an input pattern that is very different from the typical patterns of the same class".

A. TRUE.

B. FALSE.

10 Training and Testing

What is *generalization*?

A. The ability of a pattern recognition system to approximate the desired output values for pattern vectors which are not in the test set.

B. The ability of a pattern recognition system to approximate the desired output values for pattern vectors which are not in the training set.

C. The ability of a pattern recognition system to extrapolate on pattern vectors which are not in the training set.

D. The ability of a pattern recognition system to interpolate on pattern vectors which are not in the test set.

11 Biological Neurons

Is the following statement true or false? “In the human brain, roughly 70% of the neurons are used for input and output. The remaining 30% are used for information processing.”

A. TRUE.

B. FALSE.

12 Biological Neurons

Which of the following statements is the best description of supervised learning?

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.”

13 Artificial Neural Networks

Is the following statement true or false? “Artificial neural networks are parallel computing devices consisting of many interconnected simple processors.”

A. TRUE.

B. FALSE.

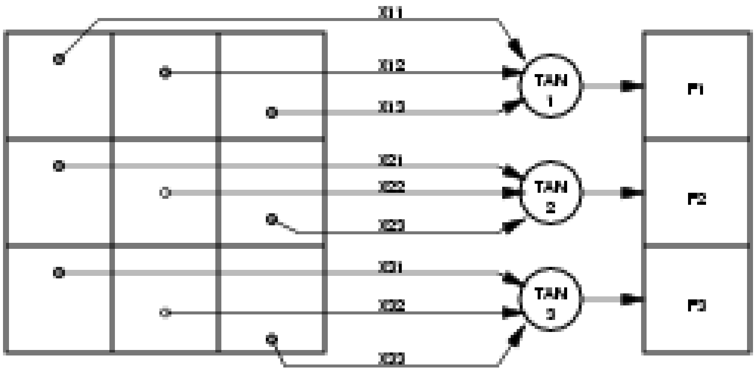
14 Artificial Neural Networks

Is the following statement true or false? “Knowledge is acquired by a neural network from its environment through a learning process, and this knowledge is stored in the connections strengths (neurons) between processing units (weights).”

A. TRUE.

B. FALSE.

15 Artificial Neural Networks



What is this network?

A. A single-layer feedforward neural network.

B. An auto-associative neural network.

C. A multi-layer feedforward neural network.

D. A Hopfield neural network.

16 Artificial Neural Networks

A neuron with 4 inputs has the weight vector $\mathbf{w} = [1, 2, 3, 4]^T$ and a bias $\theta = 0$ (zero). The activation function is linear, where the constant of proportionality equals 2 — that is, the activation function is given by $f(net) = 2 \times net$. If the input vector is $\mathbf{x} = [4, 8, 5, 6]^T$ then the output of the neuron will be

A. 1.

B. 56.

C. 59.

D. 112.

E. 118.

17 Artificial Neural Networks

Which of the following types of learning can used for training artificial neural networks?

A. Supervised learning.

B. Unsupervised learning.

C. Reinforcement learning.

D. All of the above answers.

E. None of the above answers.

18 Artificial Neural Networks

Which of the following neural networks uses supervised learning?

A. Simple recurrent network.

B. Self-organizing feature map.

C. Hopfield network.

D. All of the above answers.

E. None of the above answers.

19 Artificial Neural Networks

Which of the following algorithms can be used to train a single-layer feedforward network?

A. Hard competitive learning.

B. Soft competitive learning.

C. A genetic algorithm.

D. All of the above answers.

E. None of the above answers.

20 Artificial Neural Networks

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

A. There is no difference.

B. The Delta Rule is defined for step activation functions, but the Perceptron Learning Rule is defined for linear activation functions.

C. The Delta Rule is defined for sigmoid activation functions, but the Perceptron Learning Rule is defined for linear activation functions.

D. The Delta Rule is defined for linear activation functions, but the Perceptron Learning Rule is defined for step activation functions.

E. The Delta Rule is defined for sigmoid activation functions, but the Perceptron Learning Rule is defined for step activation functions.

21 Artificial Neural Networks

Why are linearly separable problems interesting to neural network researchers?

A. Because they are the only problems that a neural network can solve successfully.

B. Because they are the only mathematical functions that are continuous.

C. Because they are the only mathematical functions that you can draw.

D. Because they are the only problems that a perceptron can solve successfully.

22 Perceptrons

A perceptron with a unipolar step function has two inputs with weights $w_1 = 0.5$ and $w_2 = -0.2$, and a threshold $\theta = 0.3$ (θ can therefore be considered as a weight for an extra input which is always set to -1). For a given training example $\mathbf{x} = [0, 1]^T$, the desired output is 1. Does the perceptron give the correct answer (that is, is the actual output the same as the desired output)?

A. Yes.

B. No.

23 Perceptrons

The perceptron in question 22 is trained using the learning rule $\triangle\mathbf{w} = \eta\,(d - y)\,\mathbf{x}$, where $\mathbf{x}$ is the input vector, η is the learning rate, $\mathbf{w}$ is the weight vector, d is the desired output, and y is the actual output. What are the new values of the weights and threshold after one step of training with the input vector $\mathbf{x} = [0, 1]^T$ and desired output 1, using a learning rate $\eta = 0.5$?

A. $w_1 = 0.5, w_2 = -0.2, \theta = 0.3$.

B. $w_1 = 0.5, w_2 = -0.3, \theta = 0.2$.

C. $w_1 = 0.5, w_2 = 0.3, \theta = -0.2$.

D. $w_1 = 0.5, w_2 = 0.3, \theta = 0.7$.

E. $w_1 = 1.0, w_2 = -0.2, \theta = -0.2$.

24 Perceptrons

The Perceptron Learning Rule states that "for any data set which is linearly separable, the Perceptron Convergence Theorem is guaranteed to find a solution in a finite number of steps."

A. TRUE.

B. FALSE.

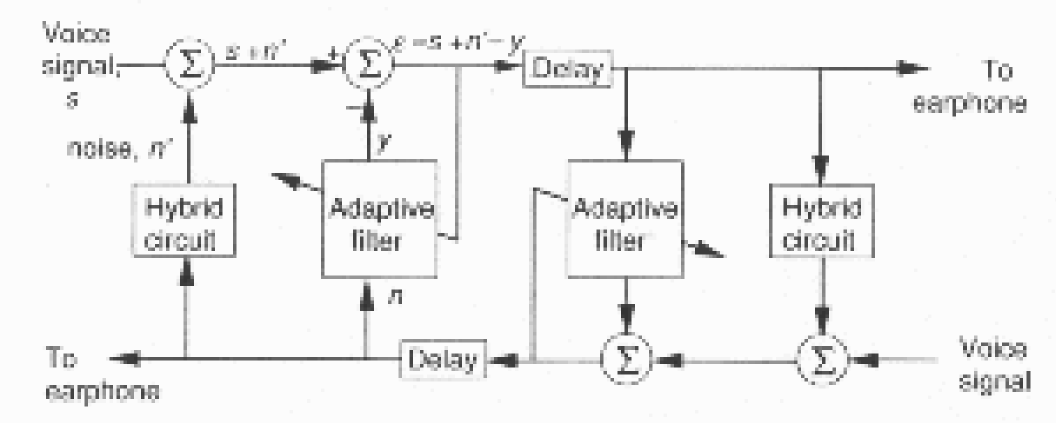
25 Perceptrons

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.

26 Adaline



The Adaline neural network can be used as an adaptive filter for echo cancellation in telephone circuits. For the telephone circuit given in the above figure, which one of the following signals carries the corrected message sent from the human speaker on the left to the human listener on the right? (Assume that the person on the left transmits an outgoing voice signal and receives an incoming voice signal from the person on the right.)

A. The outgoing voice signal, s.

B. The delayed incoming voice signal, n.

C. The contaminated outgoing signal, $s + n'$.

D. The output of the adaptive filter, y.

E. The error of the adaptive filter, $\varepsilon = s + n' - y$.

27 Multi-Layer Feedforward Networks

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 adapting the neighbours of the winning unit.

C. The problem of defining an error function for linearly inseparable problems.

D. The problem of avoiding local minima in the error function.

E. The problem of adjusting the weights for the hidden layers.

28 Multi-Layer Feedforward Networks

Is the following statement true or false? “The generalized Delta rule solves the credit assignment problem in the training of multi-layer feedforward networks.”

A. TRUE.

B. FALSE.

29 Multi-Layer Feedforward Networks

In the backpropagation algorithm, how is the error function usually defined?

A. $\frac{1}{2}\sum_k(weight_j \times input_j)$ for all inputs j.

B. $\frac{1}{2}\sum_k(target_j - output_j)^2$ for all inputs j.

C. $\frac{1}{2}\sum_k(target_j - output_j)$ for all inputs j.

D. $\frac{1}{2}\sum_k(target_j - output_j)^2$ for all outputs j.

E. $\frac{1}{2}\sum_k(target_j - output_j)$ for all outputs j.

30 Multi-Layer Feedforward Networks

A common technique for training MLFF networks is to calculate the generalization error on a separate data set after each epoch of training. Training is stopped when the generalization error starts to decrease. This technique is called

A. Boosting.

B. Momentum.

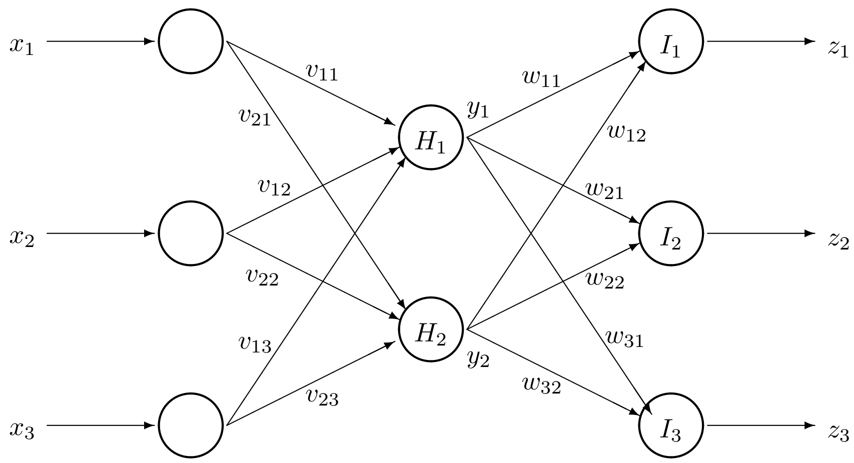
C. Hold-one-out.

D. Early stopping.

E. None of the above answers.

31 Multi-Layer Feedforward Networks

A training pattern, consisting of an input vector $\mathbf{x} = [x_1, x_2, x_3]^T$ and desired outputs $\mathbf{t} = [t_1, t_2, t_3]^T$, 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 $y_j = f(H_j)$, (2) calculate $z_k = f(I_k)$, (3) update v_{ji}, (4) update w_{kj}.

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

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

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

32 Multi-Layer Feedforward Networks

For the same neural network given in question 31, the input vector to the network is $\mathbf{x} = [x_1, x_2, x_3]^T$, the vector of hidden layer outputs is $\mathbf{y} = [y_1, y_2]^T$, the vector of actual outputs is $\mathbf{z} = [z_1, z_2, z_3]^T$, and the vector of desired outputs is $\mathbf{t} = [t_1, t_2, t_3]^T$. The network has the following weight vectors:

$$\mathbf{v}_1 = \begin{bmatrix} 0.4 \\ -0.6 \\ 1.9 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1.2 \\ 0.5 \\ -0.7 \end{bmatrix}, \quad \mathbf{w}_1 = \begin{bmatrix} 1.0 \\ -3.5 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} 0.5 \\ -1.2 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_3 = \begin{bmatrix} 0.3 \\ 0.6 \end{bmatrix}.$$

Assume that all units have sigmoid activation functions given by

$$f(x) = \frac{1}{1 + \exp(-x)}$$

and that each unit has a bias $\theta = 0$ (zero). If the network is tested with an input vector $\mathbf{x} = [1.0, 2.0, 3.0]^T$ then the output y_1 of the first hidden neuron will be

A. -2.300

B. 0.091

C. 0.644

D. 0.993

E. 4.900

(*Hint:* on some calculators, $\exp(x) = e^x$ where $e = 2.7182818$)

33 Multi-Layer Feedforward Networks

Assuming exactly the same neural network and the same input vector as in the previous question, what is the activation I_2 of the second output neuron?

A. 0.353

B. 0.387

C. 0.596

D. 0.662

E. 0.674

34 Multi-Layer Feedforward Networks

For the hidden units of the network in question 31, the generalized Delta rule can be written as

$$\triangle v_{ji} = \eta \delta_j x_i$$

where $\triangle v_{ji}$ is the change to the weights from unit i to unit j, η is the learning rate, δ_j is the error term for unit j, and x_i is the i^{th} input to unit j. In the backpropagation algorithm, what is the error term δ_j?

A. $\delta_j = f'(H_j)(t_k - z_k)$.

B. $\delta_j = f'(I_k)(t_k - z_k)$.

C. $\delta_j = f'(H_j) \sum_k \delta_k w_{kj}$.

D. $\delta_j = f'(I_k) \sum_k \delta_k w_{kj}$.

where $f'(net)$ is the derivative of the activation function $f(net)$.

35 Multi-Layer Feedforward Networks

For the output units of the network in question 31, the generalized Delta rule can be written as

$$\triangle w_{kj} = \eta \delta_k y_j$$

where $\triangle w_{kj}$ is the change to the weights from unit j to unit k, η is the learning rate, δ_k is the error term for unit k, and y_j is the j^{th} input to unit k. In the backpropagation algorithm, what is the error term δ_k?

A. $\delta_k = f'(H_j)(t_k - z_k)$.

B. $\delta_k = f'(I_k)(t_k - z_k)$.

C. $\delta_k = f'(H_j) \sum_k \delta_k w_{kj}$.

D. $\delta_k = f'(I_k) \sum_k \delta_k w_{kj}$.

where $f'(net)$ is the derivative of the activation function $f(net)$.

36 Multi-Layer Feedforward Networks

Which of the following equations best describes the generalized Delta rule with momentum?

A. $\triangle^p w_{kj}(t+1) = \eta \delta_k y_j + \alpha f'(H_j) y_j(t)$

B. $\triangle^p w_{kj}(t+1) = \alpha \delta_k y_j(t)$

C. $\triangle^p w_{kj}(t+1) = \eta \delta_k y_j + \alpha \triangle^p w_{kj}(t)$

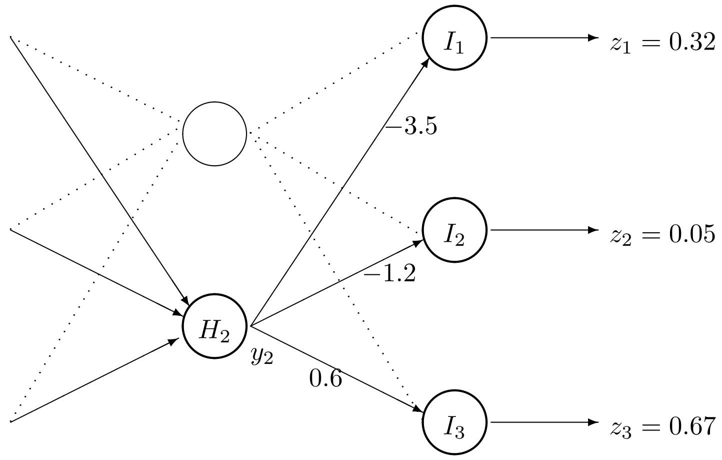
D. $\triangle^p w_{kj}(t+1) = \eta \delta_k y_j(t)$

E. $\triangle^p w_{kj}(t+1) = \eta \delta_k y_j + \alpha \delta_k y_j(t)$

where $\triangle w^p_{kj}(t)$ is the change to the weights from unit j to unit k at time t when a training pattern p is presented to the network, η is the learning rate, α is the momentum coefficient, δ_k is the error term for unit k, and y_j is the j^{th} input to unit k.

37 Multi-Layer Feedforward Networks

The following figure shows part of the neural network described in question 31. A new input pattern is presented to the network and training proceeds as follows. The actual outputs of the network are given by $\mathbf{z} = [0.32, 0.05, 0.67]^T$ and the corresponding target outputs are given by $\mathbf{t} = [1.00, 1.00, 1.00]^T$. The weights w_{12}, w_{22} and w_{32} are also shown below.



For the output units, the derivative of the sigmoid function can be rewritten as

$$f'(I_k) = f(I_k)[1 - f(I_k)].$$

What is the error for each of the output units?

A. $\delta_{\texttt{output\_1}} = -0.2304$, $\delta_{\texttt{output\_2}} = 0.3402$, and $\delta_{\texttt{output\_3}} = -0.8476$.

B. $\delta_{\texttt{output\_1}} = 0.1084$, $\delta_{\texttt{output\_2}} = 0.1475$, and $\delta_{\texttt{output\_3}} = 0.1054$.

C. $\delta_{\texttt{output\_1}} = 0.1480$, $\delta_{\texttt{output\_2}} = 0.0451$, and $\delta_{\texttt{output\_3}} = 0.0730$.

D. $\delta_{\texttt{output\_1}} = 0.4225$, $\delta_{\texttt{output\_2}} = -0.1056$, and $\delta_{\texttt{output\_3}} = 0.1849$.

(Assume exactly the same weights, activation functions and bias values as described in question 31.)

38 Multi-Layer Feedforward Networks

For the hidden units, the derivative of the sigmoid function can be rewritten as

$$f'(H_j) = f(H_j)[1 - f(H_j)].$$

What is the error for hidden unit 2 given that its activation for the pattern being processed is currently $y_2 = 0.50$?

A. $\delta_{\texttt{hidden\_2}} = -0.4219$

B. $\delta_{\texttt{hidden\_2}} = -0.1321$

C. $\delta_{\texttt{hidden\_2}} = -0.0677$

D. $\delta_{\texttt{hidden\_2}} = 0.0481$

E. $\delta_{\texttt{hidden\_2}} = 0.1231$

(Assume exactly the same weights, activation functions and bias values as described in question 31, and exactly the same output vectors $\mathbf{t}$ and $\mathbf{z}$ as described in the previous question.)

39 Multi-Layer Feedforward Networks

Which of the following techniques is NOT a strategy for dealing with local minima in the backpropagation algorithm?

A. Train and test using the hold-one-out strategy.

B. Add random noise to the weights during training.

C. Add random noise to the training examples at each epoch of training.

D. Repeat the training process several times with new random weights, and pick the network that gives the best generalization performance.

E. Train and test a committee or ensemble of networks, then use the average of the network outputs.

40 Multi-Layer Feedforward Networks

Is the following statement true or false? "The neural network in question 31 could be used for data compression."

A. TRUE.

B. FALSE.

41 Multi-Layer Feedforward Networks

Is the following statement true or false? "The neural network in question 31 could be used for function approximation."

A. TRUE.

B. FALSE.

42 Multi-Layer Feedforward Networks

Which of the following statements is NOT true for an autoassociative feedforward network with a single hidden layer of neurons?

A. During training, the target output vector is the same as the input vector.

B. It is important to use smooth, non-decreasing activation functions in the hidden units.

C. The network could be trained using the backpropagation algorithm, but care must be taken to deal with the problem of local minima.

D. After training, the hidden units give a representation that is equivalent to the principal components of the training data, removing non-redundant parts of the input data.

E. The trained network can be split into two machines: the first layer of weights compresses the input pattern (encoder), and the second layer of weights reconstructs the full pattern (decoder).

43 Multi-Layer Feedforward Networks

Which of the following statements is NOT true for a simple recurrent network (SRN)?

A. The training examples must be presented to the network in the correct order.

B. The test examples must be presented to the network in the correct order.

C. This type of network can predict the next chunk of data in the series from the past history of data.

D. The hidden units encode an internal representation of the data in the series that precedes the current input.

E. The number of context units should be the same as the number of input units.

44 Associative Memory

How many hidden layers are there in an autoassociative Hopfield network?

A. None (0).

B. One (1).

C. Two (2).

45 Associative Memory

A Hopfield network has 20 units. How many adjustable parameters does this network contain?

A. 95

B. 190

C. 200

D. 380

E. 400

46 Associative Memory

Calculate the weight matrix for a Hopfield network to store the pattern $\begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}$.

A. $\mathbf{W} = \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{bmatrix}$

B. $\mathbf{W} = \begin{bmatrix} 0 & 1 & -1 & -1 \\ 1 & 0 & -1 & -1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & 1 & 0 \end{bmatrix}$

C. $\mathbf{W} = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$

D. $\mathbf{W} = \begin{bmatrix} 0 & -1 & 1 & -1 \\ -1 & 0 & -1 & 1 \\ 1 & -1 & 0 & -1 \\ -1 & 1 & -1 & 0 \end{bmatrix}$

47 Associative Memory

Calculate the weight matrix for a Hopfield network to store two patterns $\begin{bmatrix}1 & -1 & 1 & -1\end{bmatrix}$ and $\begin{bmatrix}-1 & -1 & -1 & 1\end{bmatrix}$.

A. $\mathbf{W} = \begin{bmatrix} 0 & 2 & 0 & -2 \\ 2 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ -2 & -2 & 0 & 0 \end{bmatrix}$

B. $\mathbf{W} = \begin{bmatrix} 0 & -2 & 0 & -2 \\ -2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ -2 & 2 & 0 & 0 \end{bmatrix}$

C. $\mathbf{W} = \begin{bmatrix} 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 2 \\ -2 & 0 & 2 & 0 \end{bmatrix}$

D. $\mathbf{W} = \begin{bmatrix} 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & -2 \\ -2 & 0 & -2 & 0 \end{bmatrix}$

48 Associative Memory

The trained Hopfield network in question 47 is used for recall. A pattern $\begin{bmatrix}-1 & -1 & 1 & -1\end{bmatrix}$ is presented to the network, then the nodes of the network are updated until a steady state is reached. What is the final state of the network?

A. $\begin{bmatrix}-1 & -1 & 1 & -1\end{bmatrix}$

B. $\begin{bmatrix}1 & -1 & 1 & -1\end{bmatrix}$

C. $\begin{bmatrix}-1 & -1 & -1 & 1\end{bmatrix}$

D. $\begin{bmatrix}-1 & 1 & -1 & 1\end{bmatrix}$

49 Unsupervised Learning

Is the following statement true or false? "Patterns within a cluster should be similar in some way."

A. TRUE.

B. FALSE.

50 Unsupervised Learning

Is the following statement true or false? "Clusters that are similar in some way should be far apart."

A. TRUE.

B. FALSE.

51 Unsupervised Learning

Which of the following statements is NOT true for *hard competitive learning* (HCL)?

A. There is no target output in HCL.

B. There are no hidden units in a HCL network.

C. The input vectors are often normalized to have unit length — that is, $\| \mathbf{x} \| = 1$.

D. The weights of the winning unit k are adapted by $\triangle \mathbf{w}_k = \eta \, (\mathbf{x} - \mathbf{w}_k)$, where $\mathbf{x}$ is the input vector.

E. The weights of the neighbours j of the winning unit are adapted by $\triangle \mathbf{w}_j = \eta_j \, (\mathbf{x} - \mathbf{w}_j)$, where $\eta_j < \eta$ and $j \neq k$.

52 Unsupervised Learning

Which of the following statements is NOT true for a self-organizing feature map (SOFM)?

A. The size of the neighbourhood is decreased during training.

B. The SOFM training algorithm is based on soft competitive learning.

C. The network can grow during training by adding new cluster units when required.

D. The cluster units are arranged in a regular geometric pattern such as a square or ring.

E. The learning rate is a function of the distance of the adapted units from the winning unit.

53 Genetic Algorithms

Which of the following statements is the best description of *reproduction*?

A. Randomly change a small part of some strings.

B. Randomly generate small initial values for the weights.

C. Randomly pick new strings to make the next generation.

D. Randomly combine the genetic information from 2 strings.

54 Genetic Algorithms

Which of the following statements is the best description of *mutation*?

A. Randomly change a small part of some strings.

B. Randomly pick new strings to make the next generation.

C. Randomly generate small initial values for the weights.

D. Randomly combine the genetic information from 2 strings.

55 Genetic Algorithms

Ranking is a technique used for

A. deleting undesirable members of the population.

B. obtaining the selection probabilities for reproduction.

C. copying the fittest member of each population into the mating pool.

D. preventing too many similar individuals from surviving to the next generation.

56 Genetic Algorithms

Is the following statement true or false? "A genetic algorithm could be used to search the space of possible weights for training a recurrent artificial neural network, without requiring any gradient information."

A. TRUE.

B. FALSE.

57 Applications

Is the following statement true or false? "Learning produces changes within an agent that over time enables it to perform more effectively within its environment."

A. TRUE.

B. FALSE.

58 Applications

Which application in intelligent mobile robots made use of a single-layer feedforward network?

A. Goal finding.

B. Path planning.

C. Wall following.

D. Route following.

E. Gesture recognition.

59 Applications

Which application in intelligent mobile robots made use of a self-organizing feature map?

A. Goal finding.

B. Path planning.

C. Wall following.

D. Route following.

E. Gesture recognition.

60 Applications

Which application in intelligent mobile robots made use of a genetic algorithm?

A. Goal finding.

B. Path planning.

C. Wall following.

D. Route following.

E. Gesture recognition.

Answers

1.		B	+1	$-\frac{1}{2}$	
2.		B	+1	$-\frac{1}{2}$	
3.		E	+1	$-\frac{1}{4}$	
4.		B	+1	$-\frac{1}{4}$	
5.		B	+1	$-\frac{1}{2}$	
6.		C	+1	$-\frac{1}{3}$	
7.		C	+1	$-\frac{1}{2}$	
8.		A	+1	$-\frac{1}{2}$	
9.		A	+1	-1	
10.		B	+1	$-\frac{1}{3}$	

11.		B	+1	-1	
12.		C	+1	$-\frac{1}{2}$	
13.		A	+1	-1	
14.		A	+1	-1	
15.		A	+1	$-\frac{1}{3}$	
16.		E	+1	$-\frac{1}{4}$	
17.		D	+1	$-\frac{1}{4}$	
18.		A	+1	$-\frac{1}{4}$	
19.		D	+1	$-\frac{1}{4}$	
20.		D	+1	$-\frac{1}{4}$	

21.		D	+1	$-\frac{1}{3}$	
22.		B	+1	-1	
23.		C	+1	$-\frac{1}{4}$	
24.		B	+1	-1	
25.		A	+1	-1	
26.		E	+1	$-\frac{1}{4}$	
27.		E	+1	$-\frac{1}{4}$	
28.		A	+1	-1	
29.		D	+1	$-\frac{1}{4}$	
30.		E	+1	$-\frac{1}{4}$	

31.		B	+1	$-\frac{1}{3}$	
32.		D	+1	$-\frac{1}{4}$	
33.		B	+1	$-\frac{1}{4}$	
34.		C	+1	$-\frac{1}{3}$	
35.		B	+1	$-\frac{1}{3}$	
36.		C	+1	$-\frac{1}{4}$	
37.		C	+1	$-\frac{1}{3}$	
38.		B	+1	$-\frac{1}{4}$	
39.		A	+1	$-\frac{1}{4}$	
40.		A	+1	-1	

41.		A	$+1$	-1	
42.		D	$+1$	$-\frac{1}{4}$	
43.		E	$+1$	$-\frac{1}{4}$	
44.		A	$+1$	$-\frac{1}{2}$	
45.		B	$+1$	$-\frac{1}{4}$	
46.		D	$+1$	$-\frac{1}{3}$	
47.		D	$+1$	$-\frac{1}{3}$	
48.		B	$+1$	$-\frac{1}{3}$	
49.		A	$+1$	-1	
50.		B	$+1$	-1	

51.		E	$+1$	$-\frac{1}{4}$	
52.		C	$+1$	$-\frac{1}{4}$	
53.		C	$+1$	$-\frac{1}{3}$	
54.		A	$+1$	$-\frac{1}{3}$	
55.		B	$+1$	$-\frac{1}{3}$	
56.		A	$+1$	-1	
57.		A	$+1$	-1	
58.		C	$+1$	$-\frac{1}{4}$	
59.		D	$+1$	$-\frac{1}{4}$	
60.		B	$+1$	$-\frac{1}{4}$	