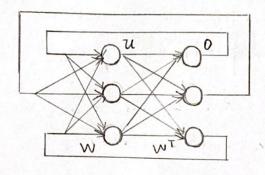
1. (a) 
$$P_1 = [1] - 1 + 1]^T$$
,  $P_2 = [-1] - 1]^T$   
 $W = \sum_{i=1}^{N} x_i y_i^T \implies W = \begin{bmatrix} 1 \\ -1 \end{bmatrix} [-1] - 1] = \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$ 



Let input be 
$$X_1 = P_1 = [1 - 1 \ 1]^T$$
;  
 $O(1) = (P(W^TX)) = P(\begin{bmatrix} -1 \ -1 \ -1 \end{bmatrix} \begin{bmatrix} -1 \ -1 \end{bmatrix} \begin{bmatrix} -1 \ -1 \end{bmatrix} \begin{bmatrix} -1 \ -1 \end{bmatrix} = P_1$   
 $= (P([-3 \ 3 \ -3]^T)) = [-1 \ 1 \ -1]^T = P_1$   
Let input be  $X_2 = [1 \ 1 \ 1]^T \neq P_1 \neq P_2$ :  
 $O(1) = (P([-1 \ 1 \ -1]) \begin{bmatrix} -1 \ 1 \ -1 \end{bmatrix} \begin{bmatrix} -1 \ 1 \end{bmatrix} = P_2$   
 $= (P([-1 \ 1 \ -1]) = [-1 \ 1 \ -1]^T = P_2$ 

(b) 
$$Q_1 = [1 - 1 - 1]^T$$
;  $Q_2 = [1 1 1]^T$ 

in old to the state of the stat
in pht on t put $ \begin{array}{c cccc} \hline \Gamma & -1 & -1 & T & T & T & T & T & T & T & T & T & $
$\begin{array}{c c} & & & \\ \hline & & & \\ \hline \end{array}$
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$
m 11x- (11)2
$f(x) = \Phi w = \sum_{i=1}^{N} w_i \cdot \exp(-\frac{x_i}{2\sigma^2})$ , assume $c_i = [1 - 1 - 1]^T$ .
$C_2 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$ , $D = a707$
D- 1 exp(-8) 1 aovoss5
$f(x) = \Phi w = \sum_{j=1}^{m} w_j \cdot \exp(-\frac{  x - c_j  ^2}{2\sigma^2}), \text{ assume } c_1 = [1 - 1 - 1]^T, \\ c_2 = [1 - 1 - 1]^T, \\ d = a_{j=1}^{m} v_j \cdot \exp(-8) = \begin{bmatrix} 1 & a_{000} & 0 \\ 0 & 000 & 0 \end{bmatrix} = \begin{bmatrix} 1 & a_{000} & 0 \\ 0 & 000 & 0 \end{bmatrix} = \begin{bmatrix} 1 & a_{000} & 0 \\ 0 & 000 & 0 \end{bmatrix}$
[ No. 2017 NO. 12. ] - [ NO. 2017 NO.
$W = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = (\Phi^T \Phi)^{-1} \Phi^T d = \begin{bmatrix} 1 & aovosss \\ avosss \end{bmatrix} \begin{bmatrix} 1 & aovosss \\ aovosss \end{bmatrix} \begin{bmatrix} 1 & aovosss \end{bmatrix}$
([0.000]]] [ ][0.000]]]
[ 1 00067] [ 1 -0006]] [ 1 0.0005]
$= \begin{bmatrix} 1 & 0 & 000 & 67 \end{bmatrix} \Phi^{T} d = \begin{bmatrix} 1 & -0 & 000 & 61 \end{bmatrix} \begin{bmatrix} 1 & 0 & 000 & 52 \end{bmatrix} d$
[#########] 이 교육하는 아니다 전 하는 11 등로 하는 11 등로 하다 하고 있는 12 등에 하는 12 등에
$= \begin{bmatrix} 1 & -avousss \\ -avousss & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1.000355 \\ -1.000355 \end{bmatrix}$
- 21/2000 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -
if x=[1-1-1] : f(x)=(1.000355 x exp(0)+(-1.000355 x exp(-8))
= 1,000355 - 0.000335 = 1

network overfits the training data. To solve this problem we could use a validation set to test the trained network and adjust the weights based on the validation results, and then apply the validated network to the training set.

input autput

$$Q_{1}=[1-1-1]^{T}$$
 $Q_{2}=[1-1-1]^{T}$ 
 $Q_{3}=[1-1-1]^{T}$ 
 $Q_{3}=[1-1-1]^{T}$ 
 $Q_{3}=[1-1-1]^{T}$ 
 $Q_{3}=[1-1-1]^{T}$ 
 $Q_{4}=[1-1-1]^{T}$ 
 $Q_{5}=[1-1-1]^{T}$ 
 $Q_{5}=[1-1]^{T}$ 
 $Q_{5}=[1-1]^{T}$ 

2. (a) 3 ingredients: ① competition ② cooperation ③ weight adaptation

(b) step 1: initialization. Randomly select weight vector vij(0) which satisfies vij(0) are different from each other and have small magnitudes.

Step 2: sampling. Randomly select a input sample x from training samples.

Step 3: competition. Find a minning neuron which satisfies:  $i(x) = arg min || x - w_j(n) ||, j=1,2,...,N$ 

Step 4: cooperation and adaptation. Adjust the weight of the minning neuron as:

where hji'(x) =  $\exp(-\frac{dji^2}{20tn})$  is the neighborhood function,  $dji' = 11 \times j - \times i \times 11$  is the distance between winning neuron t and activated neuron t.  $J(n) = J_0 \cdot \exp(-\frac{n}{T_1})$ , n is the iteration time, t is the time constant, t is the initialization value. learning rate  $J(n) = J_0 \cdot \exp(-\frac{n}{T_2})$ 

Step 5: iteration. Go back to Step 2 until the stopping criterion is satisified:

O reach the pre-defined iteration number

@ no noticeable change in the network

cc) Som can tind the prototype of the input patterns

• Self-organizing phase: 
$$\eta(n) = \alpha_1 \exp(-\frac{n}{1000})$$
,  $h_{ji} = \exp(-\frac{d_{ji}^2}{2\sigma^2(n)})$ 

· convergence phase;

non) = 0.01, neuron) is the minning nouron

hii = { 0. otherwise

 $T_1 = \frac{1000}{1000}$ .  $T_0 = \text{radius of two}$ Lattice.

3. (a) for the linearly separable 2-class pattern, the decision boundary is given by: 
$$g(x) = w^{T}x + b$$

where w' is the weight,  $x = xp + r \frac{w}{||w||}$ , xp is the projection of x to the decision boundary, r is the distance, b is the bias.  $g(x) = w'x + b = w'xp + b + r \frac{w'w}{||w||} = g(xp) + r||w|| = r||w||$   $r = \frac{g(x)}{||w||}$ , to maximize the margin is equal to minimize the weight w.

The distance from support vector  $x^{(s)}$  to the hyperplane is:  $r = \begin{cases} \frac{9(x^{(s)})}{11 \text{ wil}} = \frac{1}{11 \text{ wil}}, & \text{for } d^{(s)} = +1 \\ \frac{9(x^{(s)})}{11 \text{ wil}} = \frac{-1}{11 \text{ wil}}, & \text{for } d^{(s)} = -1 \end{cases}$ 

The optimization can be termulated as: given samples { x(i), d(i)}, i=1,2,..., N, tind the optimal w and b such that:

min. 
$$J = \frac{1}{2}w^Tw$$
  
S.t.  $d(i) \times [w^TX(i) + b] \ge 1$ 

By using lagrange multiplier theory, we have:  $J(w,b,\alpha) = \frac{2}{2}w^{T}w - \frac{2}{1-1}\alpha(i)\left[d(i)xEw^{T}x(i)+b1-1\right]$ where dis is the lagrange multiplier and  $\alpha(i) \ge 0$ .

$$\frac{\partial J(w,b,d)}{\partial w} = w - \sum_{i=1}^{N} \alpha_{ii} d_{ii} \chi_{ii} = 0$$

$$\frac{\partial J(w,b,d)}{\partial b} = -\sum_{i=1}^{N} \alpha_{ii} d_{ii} \chi_{ii} = 0$$

=> 
$$w = \underset{i=1}{\overset{\circ}{\sum}} a_{ii} d_{ii} x_{ii}$$
  
 $\underset{i=1}{\overset{\circ}{\sum}} a_{ii} d_{ii} = 0$ 

with KKI conditions we know that: dii)[dii)x[w\xii)+b]-1]=0
for non-zero dii), (dii)x[w\xii)+b]-1) must be 0.

Rennite Jiw, b, a) as follows:

$$J(w,b,d) = \frac{1}{2}w^{T}w - \sum_{i=1}^{N}\alpha_{ii}d_{ii})w^{T}x_{ii} - \sum_{i=1}^{N}\alpha_{ii}d_{ii}b + \sum_{i=1}^{N}\alpha_{ii}$$
with  $w = \sum_{i=1}^{N}\alpha_{ii}d_{ii}x_{ii}$ , we have

$$w^{T}w = \sum_{i=1}^{N} \alpha_{i}i)d_{i}i) \times (i) \sum_{j=1}^{N} \alpha_{j}j)d_{j}j \times (j) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_{i}i)d_{i}j \times (i) \times (j)$$

where Q(d) is the dual problem of the above optimization problem and can be solved when d(i) is found by QP solvers.

min. 
$$Q(d)$$
  
s.t.  $\angle A(i)d(i)=0$   
 $a(i)\geq 0$   
 $b^*=1-w^*x^{(i)}(i)$ 

(b) Simillar with (a), for linearly non-seperable 2-class patterns, we can tormulate the optimization problem as:

min.  $J = \frac{1}{2}w^Tw + C \stackrel{\times}{Z} 3(i)$  where  $\frac{1}{2}(i)$  denotes the slack s.t.  $d(i) \times [w^T \times (i) + b] \ge 1 - \frac{1}{2}(i)$  vaniable.

min. Q(a) = \( \frac{1}{2} \alpha(i) - \frac{1}{2} \frac{1}{2} \frac{1}{2} \alpha(i) \alpha(i) \alpha(i) \text{X}(i) \text{X}(

$$(x + w^* = \sum_{i=1}^{N} a(i)d(i) \times (i)$$
,  $b^* = 1 - w^* \times (i)$ 

(c) We can use the one-v.s.-all method to solve multi-class classification problem.

4. (a)
$$M_{X_{1}}(x) = \begin{cases}
1, & 0 \le x \le 0.3 \\
-\frac{10}{3}x + 2, & 0.3 \le x \le 0.6 \\
0, & x \ge 0.6
\end{cases}$$

$$M_{X_{2}}(x) = \begin{cases}
0, & 0 \le x \le 0.4 \\
\frac{10}{3}x - \frac{4}{3}, & 0.4 \le x \le 0.7 \\
1, & x \ge 0.7
\end{cases}$$

5. (a)
$$B = A \circ R = \begin{bmatrix} 0.1 \\ 0.8 \\ 0.4 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 & 0.1 \\ 0.4 & 0.8 & 0.2 \\ 0.1 & 0.2 & 0.9 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.4 & 0.8 & 0.2 \\ 0.1 & 0.2 & 0.4 \end{bmatrix} \begin{cases} min \\ max \end{cases}$$

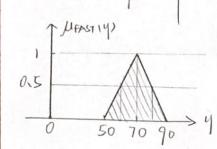
(b) 
$$X_0 = 22.5$$
,  $MWARM(X_0) = 5.5 - 0.2 \times 22.5 = 1$   
Rule 4: If TEMP is WARM THEN SPEED is FAST

$$= \sum M_{FAST}(y) = \begin{cases} a_0 \le y - 2.5 \\ 45 - a_0 \le y \end{cases}, \quad 70 \le y \le 90$$

$$= \sum M_{FAST}(y) = \begin{cases} a_0 \le y - 2.5 \\ 45 - a_0 \le y \end{cases}, \quad 70 \le y \le 90$$

$$= \sum M_{FAST}(y) = \begin{cases} a_0 \le y - 2.5 \\ 45 - a_0 \le y \end{cases}, \quad 70 \le y \le 90$$

$$= \sum M_{FAST}(y) = \begin{cases} a_0 \le y - 2.5 \\ 45 - a_0 \le y \end{cases}, \quad 70 \le y \le 90$$



speed<sub>co6</sub> = 
$$\frac{(90-50)\times\frac{1}{2}\times1\times(70)}{\frac{1}{2}\times190-50)\times1}$$
= 
$$\frac{20\times70}{20} = 70$$

(C) 
$$X_1 = 20$$
,  $M_{WARM}(X_1) = \alpha_1 \times 20 - 3$ ,  $\Sigma = 0.5$ .  
 $Speed_{COG} = \frac{15 \times 70}{15} > 70$ 

$$0.05y - 2.5 = 0.5$$
  
 $y = 60$   
 $45 - 0.05y = 0.5$   
 $y = 80$   
 $0.05y = 0.5$   
 $0.05y = 0.5$ 

EE7207

## NANYANG TECHNOLOGICAL UNIVERSITY SEMESTER 1 EXAMINATION 2016-2017 EE7207 – NEURAL AND FUZZY SYSTEMS

November/December 2016

Time Allowed: 3 hours

## INSTRUCTIONS

1. This paper contains 5 questions and comprises 5 pages.

All questions carry equal marks.

let input be 
$$X = P_1 = [1 - 1 \ 1]^T$$

$$O(1) = Q(WX) = Q([-1] - 1 \ 1] = [-1 \ 1 - 1]^T$$

4. This is a closed-book examination.

1. (a) There are two vectors:

$$P_1 = [1 \quad -1 \quad 1]^T$$
 and  $P_2 = [-1 \quad 1 \quad -1]^T$ .

Design a bi-directional associative memory (BAM) neural network to store  $P_1$  and  $P_2$ . Sketch the architecture of the BAM neural network designed, compute the weights on the links between neurons, and test whether the designed BAM works.

weights on the links between neurons, and test whether the designed BAM works.

(8 Marks)

(b) There are two vectors:
$$Q_1 = \begin{bmatrix} 1 & -1 & -1 \end{bmatrix}^T \text{ and } Q_2 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T.$$

Design a Gaussian radial basis function (RBF) neural network to map  $Q_1$  and  $Q_2$  to 1 and -1, respectively. Sketch the RBF neural network architecture, determine centre vectors of hidden layer neurons and the weights on the links between neurons, and test whether the designed RBF neural network works.

$$f(x) = \Phi w = \sum_{j=1}^{m} w_{ji} \cdot \exp(-\frac{11x - c_{j}(8xMarks)}{2\sigma^{2}})$$

Note: Question No. 1 continues on page 2

EE7207

(c) In an application, it was found that the trained RBF neural network fitted the training data perfectly, but generalized badly on the unseen testing data. Discuss the causes of the above problem and ways to deal with the problem.

(4 Marks)

- 2. Self-organizing map (SOM) neural network is a powerful tool in data analysis.
  - (a) Describe the 3 ingredients of the SOM neural network learning.

    Competition cooperation adaptation (6 Marks)
  - (b) Describe SOM neural network learning procedure, and discuss parameter settings in different learning phases of the SOM neural network.

(9 Marks)

(c) Discuss the potentials of SOM neural networks in data visualization.

(5 Marks)

- 3. Support vector machine (SVM) is widely used in pattern classification.
  - (a) Describe the optimal separating hyperplane of linear SVM for linearly separable 2-class patterns, and formulate the solution of the optimal hyperplane as an optimization problem.

(7 Marks)

(b) Describe the optimal separating hyperplane of linear SVM for linearly non-separable 2-class patterns, and formulate the solution of the optimal hyperplane as an optimization problem.

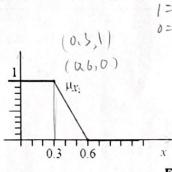
(7 Marks)

(c) SVM is inherently formulated for solving 2-class pattern classification problems. Discuss how to extend SVM to solve multi-class pattern classification problems.

(6 Marks)

$$\mathcal{L}_{X_1}(x) = \begin{cases} -\frac{19}{5}X_1 + 2 \\ -\frac{19}{5}X_1 + 2 \end{cases}$$
 EE7207

Consider the membership functions  $\mu_{X_1}(x)$  and  $\mu_{X_2}(x)$  in Figure 1, where  $X_1 = [0, \infty) \text{ and } X_2 = [0, \infty).$ 



10.5,1)

230=-1

1=030+6 a30=-1 0=060+6 a=-10 (0.4,0)

Figure 1

Determine their analytical expressions.

0.30=1 (4 Marks)  $\alpha = \frac{10}{3}$ 

- $a = -\frac{10}{5}$  b = 2
  - Suppose that an unknown system with input x and output y is modelled by the following fuzzy rules: (b) following fuzzy rules:

R1: IF x is  $X_1$  THEN  $y = a_{11} + a_{12}x$ R2: IF x is  $X_2$  THEN  $y = a_{21} + a_{22}x$ 

where  $X_1$  and  $X_2$  are fuzzy sets with membership functions  $\mu_{X_1}(x)$  and  $\mu_{X_2}(x)$  in Figure 1,  $a_{11}$ ,  $a_{12}$ ,  $a_{21}$  and  $a_{22}$  are unknown parameters.

Some experiments are carried out on the unknown system and N pairs of input output data in Table 1 are obtained.

**Table 1 Input-Output Data** 

Input x	<i>x</i> <sub>1</sub>	x <sub>2</sub>	 $x_N$
Output y	<i>y</i> <sub>1</sub>	y <sub>2</sub>	y <sub>N</sub>

Based on the data, a system of linear equations is obtained as follows:

$$\Lambda\theta = b$$
.

What is  $\theta$ ? Determine  $\Lambda$  and b using the given data.

(10 Marks)

Suppose that the optimal estimate of  $\theta$  in part 4(b) is  $\theta^*$  based on the least-squares (c) parameter identification scheme. Find  $\theta^*$  in terms of  $\Lambda$  and b.

(6 Marks)

EE7207

5. (a) Consider the fuzzy relation  $R: X \times Y \rightarrow [0, 1]$ .

$$R = \begin{array}{c|ccccccc} & y_1 & y_2 & y_3 \\ \hline x_1 & 0.7 & 0.3 & 0.1 \\ x_2 & 0.4 & 0.8 & 0.2 \\ x_3 & 0.1 & 0.2 & 0.9 \end{array}$$

and a fuzzy set A defined on X:

$$A = \{0.1/x_1, 0.8/x_2, 0.4/x_3\}.$$

Compute fuzzy set B = A o R, where o is the sup-min composition operator.

(6 Marks)

(b) Consider an air-conditioner with five control switches: COLD, COOL, PLEASANT, WARM and HOT with input temperature x in degrees Celsius (°C). The corresponding speeds of the motor controlling the fan speed y revolutions per minute (rpm) on the air-conditioner are MINIMUM, SLOW, MEDIUM, FAST and BLAST.

The rules governing the air-conditioner are as follows:

• RULE 1:		
IF TEMP	is	COLD
THEN SPEED	is	MINIMUM
• RULE 2:		
IF TEMP	is	COOL
THEN SPEED	is	SLOW
• RULE 3:		
IF TEMP	is	PLEASANT
THEN SPEED	is	MEDIUM
RULE 4:		
IF TEMP	is	WARM
THEN SPEED	is	FAST
• RULE 5:		
IF TEMP	is	HOT
THEN SPEED	is	BLAST

Note: Question No. 5 continues on page 5

EE7207

The membership functions for the fuzzy sets are respectively given as follows:

$\mu_{COLD}(x) = \begin{cases} -0.1x + 1 & 0 \le x \le 10\\ 0 & otherwise \end{cases}$	$\mu_{COOL}(x) = \begin{cases} 0.08x & 0 \le x < 12.5\\ 3.5 - 0.2x & 12.5 \le x < 17.5\\ 0 & otherwise \end{cases}$
$\mu_{PLESANT}(x) = \begin{cases} 0.4x - 6 & 15 \le x < 17.5 \\ 8 - 0.4x & 17.5 \le x < 20 \\ 0 & otherwise \end{cases}$	$\mu_{WARM}(x) = \begin{cases} 0.2x - 3.5 & 17.5 \le x < 22.5 \\ 5.5 - 0.2x & 22.5 \le x < 27.5 \\ 0 & otherwise \end{cases}$
$\mu_{HOT}(x) = \begin{cases} 0.2x - 5 & 25 \le x < 30 \\ 1 & 30 \le x \\ 0 & otherwise \end{cases}$	$\mu_{\text{MINIMUM}}(y) = \begin{cases} -\frac{1}{30}y + 1 & 0 \le y \le 30\\ 0 & \text{otherwise} \end{cases}$
$\mu_{SLOW}(y) = \begin{cases} 0.05y - 0.5 & 10 \le y < 30\\ 2.5 - 0.05y & 30 \le y < 50\\ 0 & otherwise \end{cases}$	$\mu_{\text{MEDIUM}}(y) = \begin{cases} 0.1y - 4 & 40 \le y < 50 \\ 6 - 0.1y & 50 \le y < 60 \\ 0 & otherwise \end{cases}$
$\mu_{FAST}(y) = \begin{cases} 0.05y - 2.5 & 50 \le y < 70 \\ 4.5 - 0.05y & 70 \le y < 90 \\ 0 & otherwise \end{cases}$	$\mu_{BLAST}(y) = \begin{cases} \frac{1}{30}y - \frac{7}{3} & 70 \le y \le 100\\ 0 & otherwise \end{cases}$

For a temperature of 22.5°C, determine the fan speed using the centre of gravity (COG) method.

(7 Marks)

(c) If the temperature is 20°C in part 5(b), what is the fan speed determined by using the mean-of-maxima (MOM) method?

(7 Marks)

