Name: _____

Email: _____

Final Exam

Statistical Pattern Recognition

Department of Computer Engineering
Sharif University of Technology
Spring 2008

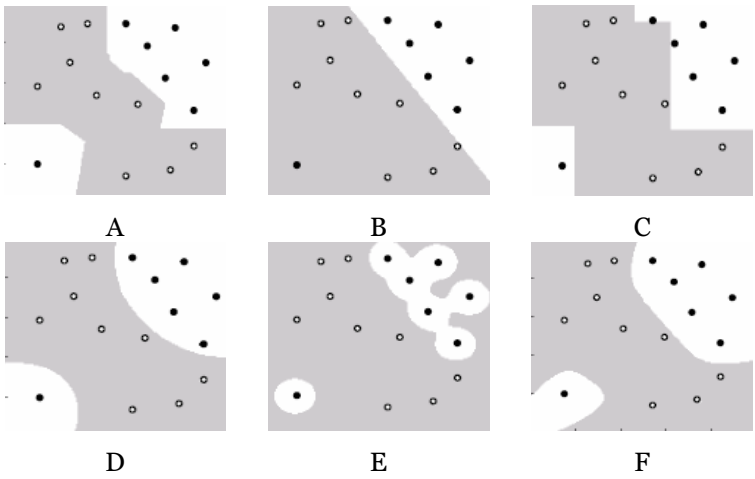
1. You must work out all the problems.
2. Show all your work in the space provided to justify your answers.

No	Max	Score
Problem 1	30	
Problem 2	20	
Problem 3	20	
Problem 4	15	
Problem 5	10	
Problem 6	15	
Total	100+10	

Good Luck!

1.Classification Methods

Consider a classification problem with two real-valued inputs. For each of the following algorithms, specify all of the separators bellow that it could have generated and explain why. If it could not have generated any of the separators, explain why not.



A B C

D E F

1. 1-nearest neighbor

2. Standard perceptron algorithm (one layer)

3. SVM with linear kernel

4. SVM with Gaussian kernel (sigma = 0.25)

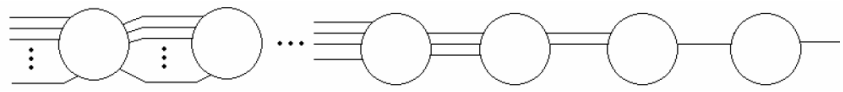
5. SVM with Gaussian kernel (sigma = 1)

6. Neural network with no hidden units and one output unit

7. Neural Network with 4 hidden units and one output unit

2. Neural Network

(a) You are given the following neural network with sigmoid units. There are n units in the network, and each unit has k inputs, with a weight of 1/k on each input. k ranges from n to 1. The first unit on the left has n inputs. There are no threshold inputs. The inputs to the network, i.e. to the leftmost neuron, are all 1. For large n, what is output of the rightmost neuron to the nearest tenth?

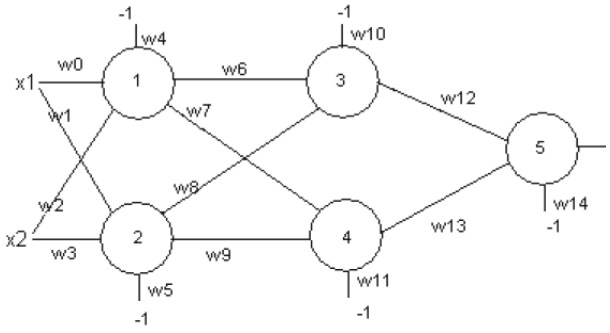


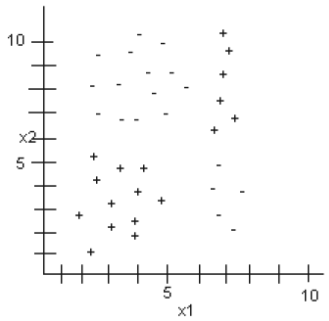
Sigmoid function used is:

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

(b) You are given the following perceptron network. Each perceptron, in contrast with the standard version will return -1 if the total of its weighted input is bellowing threshold and +1 if the total of its weighted input is above threshold. Assign the weights to properly classify the given in the following graph.

Weight	W0	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
Value															

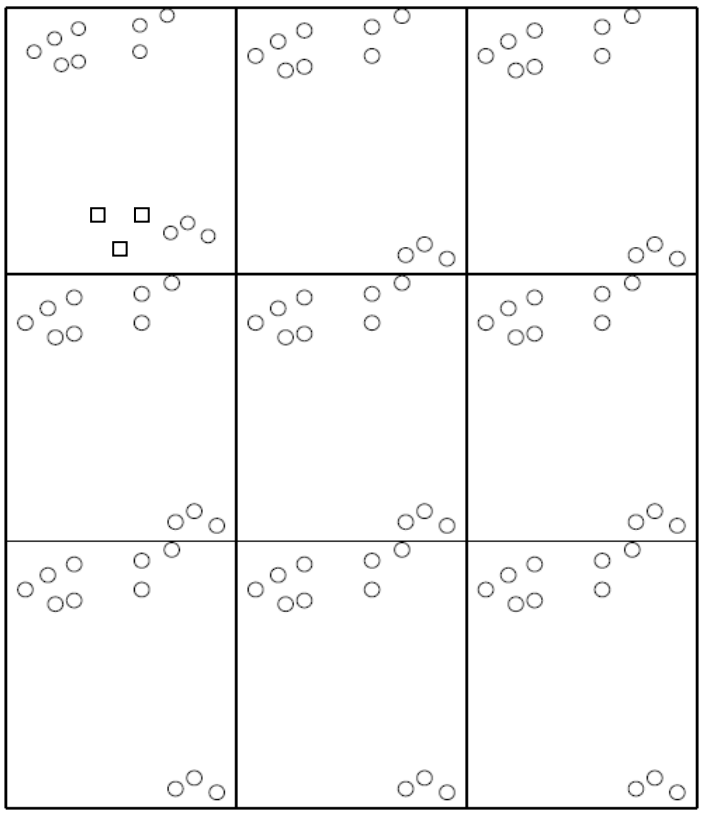




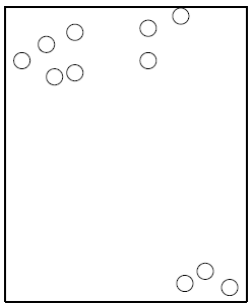
3. Clustering

(a) Run k-means manually for the following dataset. Circles are data points and squares are the initial cluster centers. Draw the cluster centers and the decision boundaries that define each cluster. Use as many pictures as you need until convergence.

Note: Execute the algorithm such that if a mean has no points assigned to it, it stays where it is for that iteration.



(b) Draw (approximately) what a Gaussian mixture model of three Gaussians with the same initial centers as for the k-means problem would converge to.



(c) Is the classification given by the mixture model the same as the classification given by k-means? Why or why not?

4. SVM

(a) What are the values for the λ_i and the offset w_0 that would give the maximal margin linear classifier for the two data points $x_1=0$, $y_1=1$ and $x_2=4$, $y_2=-1$? You should be able to find the answer without deriving it from the dual Lagrangian.

5. HMM

(a) Consider an HMM defined by the traditional set of parameter:

$$\lambda = (N, M, \{a_{ij}\}, \{b_j(O_k)\}, \pi_i)$$

The following two arrays of probabilities can be obtained by dynamic programming:

$$\alpha_t(i) = P(O_1 \wedge \ldots \wedge O_t \wedge q_t = s_i \mid \lambda)$$

$$\beta_t(i) = P(O_{t+1} \wedge \ldots \wedge O_T \mid q_t = s_i, \lambda)$$

In the process of learning an HMM from data we need the following array of probabilities:

$$\gamma_t(i) = P(q_t = s_i \mid O_1 \wedge \ldots \wedge O_T, \lambda)$$

Define $\gamma_t(i)$ in a simple finite expression in terms of $\alpha_t(i)$, $\beta_t(i)$ and $P(O_1 \wedge \ldots O_T \mid \lambda)$.

6. Learning

Given three classes, S_1, S_2, and S_3. Find linear discriminant Function for the sample points:

$$\{(0,1,-1,2)\} \in S_1, \qquad \{(1,1,1,1),(2,1,1,1)\} \in S_2 \qquad (-1,1,0,-1) \in S_3$$

Using the perceptron fixed correction algorithm. Correction can be done for many or one at a time. Use $\underline{w}^{(k)}(0) = \underline{0}$ for all *k*.