

EEL6825 – Fall 98  
Exam 2      November 9, 1998

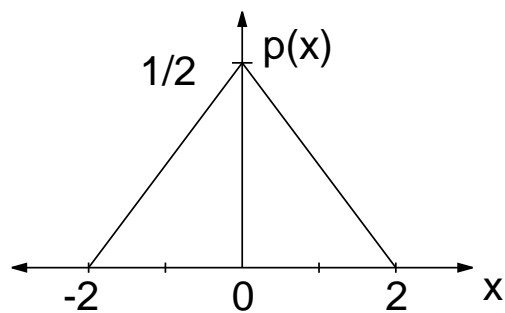
**NAME:** \_\_\_\_\_

This exam is open-book and calculator. You may use any books or papers that you like. There are four problems and an extra credit in this exam, you have two full class periods. State your assumptions and reasoning for each problem. Justify your steps and clearly indicate your final answers.

**Note: No lectures the rest of this week. Don't forget about your projects!**

<b>1</b>	<b>/25</b>
<b>2</b>	<b>/25</b>
<b>3</b>	<b>/25</b>
<b>4</b>	<b>/25</b>
<b>TOTAL</b>	

1. (25 points) You are given two data samples from some unknown 1-D probability distribution. The points are  $x^{(1)} = 1$  and  $x^{(2)} = -1$ . Answer the following questions regarding the Parzen Windows estimation of the probability density.
  - (a) (15 points) Derive expressions and sketch  $\hat{p}(x)$  for volumes of  $v = 1$ ,  $v = 2$ , and  $v = 3$ . Label all key points in the sketches.



- (b) (10 points) Suppose the true probability distribution is given above. Which of the three estimates (computed in 1a) is best in terms of minimum absolute error, given by

$$ERROR = \int_{-\infty}^{+\infty} |p(x) - \hat{p}(x)|$$

Show all of your work.

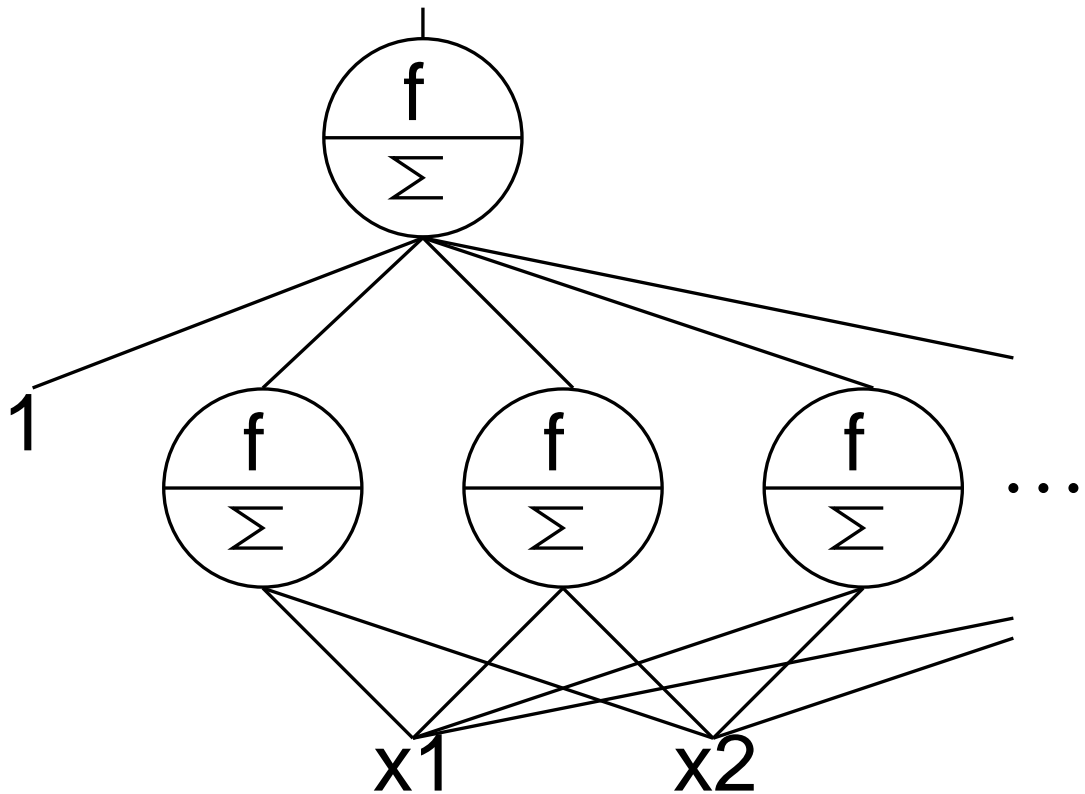


2. (25 points) You are given two data samples from some unknown 1-D probability distribution. As in problem 1, these points also happen to be  $x^{(1)} = 1$  and  $x^{(2)} = -1$ . Answer the following questions regarding nearest neighbor estimation and classification.
- (a) (10 points) Derive an expression and sketch  $\hat{p}(x)$  using the nearest neighbor volumetric technique with  $k = 2$ . Label all key points in the sketch.

- (b) (15 points) Assume that points -1 and 1 are from class  $\omega_1$ . Class  $\omega_2$  points are 3 and 5. Assume  $P(\omega_1) = 5/6$  and  $P(\omega_2) = 1/6$ . Derive a classifier using the estimated distributions using the 2-NN volumetric technique. Your answer should make obvious what the classification will be for each possible value of  $x$ .



3. (25 points)



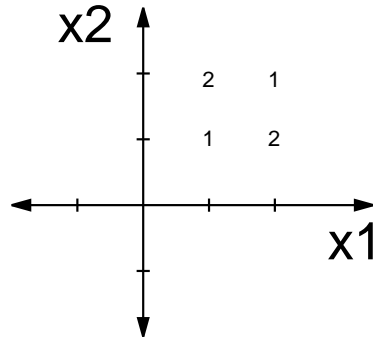
In doing an assignment, a student forgot to include the bias at the input layer. The student did remember to include the bias for the second layer– the resulting architecture is shown in the above figure. An arbitrary number of hidden units may be added but no additional hidden layers can be included. Assume the sigmoid activation function of the neural network to be:

$$f(a) = \begin{cases} 1 & \text{if } a > 0 \\ -1 & \text{else} \end{cases}$$

Answer the following questions regarding the use of this neural network to solve non-separable two-class classification problems.

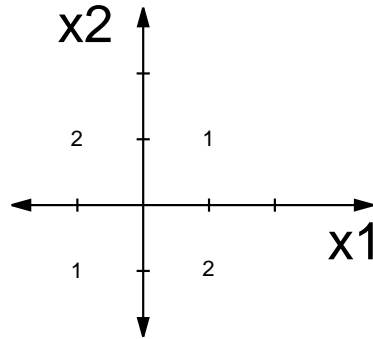


(a) (10 points)



Assuming that four data points are given as shown above, can this problem be solved using this single hidden-layer architecture with no bias in the input layer? Justify why or why not. If possible, provide all of the necessary weight values for the architecture with the minimum number of hidden units. The final output of your neural network should be +1 for class 1 and -1 for class 2.

(b) (10 points)



Assuming that four data points are given as shown above, can this problem be solved using this single hidden-layer architecture with no bias in the input layer? Justify why or why not. If possible, provide all of the necessary weight values for the architecture with the minimum number of hidden units. The final output of your neural network should be +1 for class 1 and -1 for class 2.

- (c) (5 points) Comment on the characteristics of the data sets that can be exactly classified using this architecture.



4. (25 points) Short Answer.

(a) (5 points)

Class  $\omega_1$  points are:

$$\begin{bmatrix} -1 \\ -1 \\ +1 \end{bmatrix} \begin{bmatrix} -1 \\ +1 \\ -1 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ -1 \end{bmatrix}$$

Class  $\omega_2$  points are:

$$\begin{bmatrix} +1 \\ +1 \\ -1 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ +1 \end{bmatrix} \begin{bmatrix} -1 \\ +1 \\ +1 \end{bmatrix}$$

Find any weight vector  $w$  such that  $w^T x > 0$  for all class  $\omega_1$  points and  $w^T x < 0$  for all class  $\omega_2$  points. Justify your answer.

(b) (5 points)

Class  $\omega_1$  points are:

$$\begin{bmatrix} -1 \\ -1 \\ +1 \end{bmatrix} \begin{bmatrix} -1 \\ +1 \\ -1 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ -1 \end{bmatrix}$$

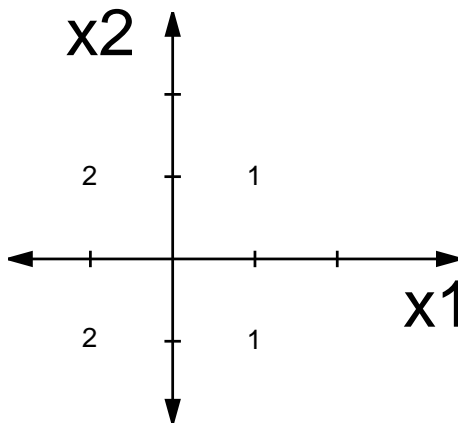
Class  $\omega_2$  points are:

$$\begin{bmatrix} +1 \\ +1 \\ -1 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ +1 \end{bmatrix} \begin{bmatrix} -1 \\ +1 \\ +1 \end{bmatrix}$$

Classify the following point using 3-Nearest Neighbor voting.

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

- (c) (5 points) Sketch the classification boundaries for 2-NN voting classification for this two-class classification problem. Clearly mark the class  $\omega_1$ , class  $\omega_2$ , reject regions.



(d) (5 points) Which methods generally have lower resubstitution errors, parametric or nonparametric methods? Explain.

(e) (5 points) We typically do not need to make any assumptions about our data when we use a nonparametric method (e.g. 1-NN). This is a tremendous advantage for nonparametric methods. While they may require a bit more computation, what is the MAJOR disadvantage of nonparametric methods?