CS401 Final Exam — Fall 2005

Instructions: Answer all five (5) questions. The three you do best on will be worth 25 points each. The one you do fourth-best on will be worth 15 points. The one you do fifth-best on will be worth 10 points. Please show any relevant calculations. (And be sure to include your name!)

(I) Define *supervised* vs *unsupervised* learning, and give an example of one supervised learning algorithm and one unsupervised learning algorithm.

(II) What is the VC dimension of the concept class of triangles in a 2D input space? (Reminder: this is the maximum number of points that can be "shattered" by triangles.)

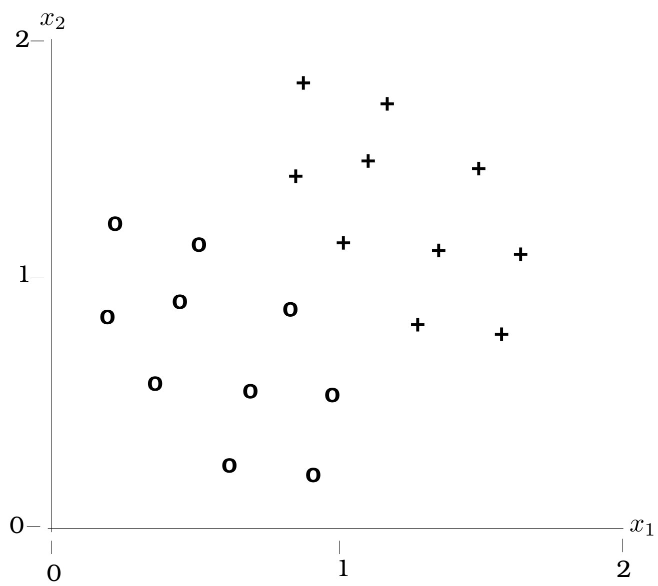
(III) Below is a set of p inputs $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(p)}$ in a 2D input space. These are divided into two classes, $\mathbf{x}^{(i)} \mapsto y^{(i)}$, with $y^{(i)} = 0$ marked as **o** below and $y^{(i)} = 1$ marked as **+** below.

Consider the linear discriminator $\hat{y} = \begin{cases} 0 & \text{if } \mathbf{w} \cdot \mathbf{x} < \theta \\ 1 & \text{if } \mathbf{w} \cdot \mathbf{x} > \theta \end{cases}$

where w_1, w_2, and θ constitute the three modifiable parameters of the learning machine.

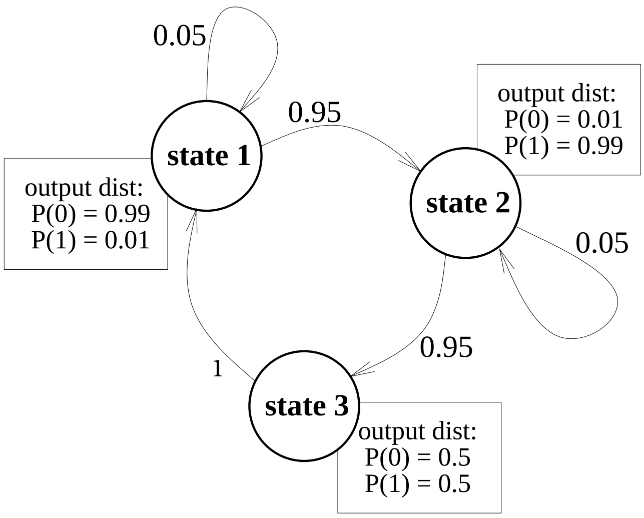
(a) In the diagram below, circle the input points which would be classified with **+** by the setting $(w_1, w_2, \theta) = (1, 0, 1)$.

(b) Give a value for (w_1, w_2, θ) which classifies the points correctly, according to the labels with which they are marked in the diagram below.



(IV) Briefly describe *cross validation* and how it can be used to avoid *overfitting*.

(V) Consider the below Hidden Markov Model. This model has three states, and two output symbols "0" and "1". The initial state distribution is $(\pi(1), \pi(2), \pi(3)) = (1, 0, 0)$, in other words the machine always starts in state 1. The state transition probabilities, and the output symbol probabilities associated with each state, are shown below.



(a) Give a length-ten sequence of states *and* a corresponding sequence of observable output symbols which you choose to be *likely* to be produced by this machine.

state sequence:										
output symbol sequence:										

(b) Give a length-ten sequence of states *and* a corresponding sequence of observable output symbols which you choose to be *unlikely but not impossible* to be produced by this machine.

state sequence:										
output symbol sequence:										