MASSACHVSETTS INSTITVTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science
6.034—Artificial Intelligence
Fall Semester, 1996

Practice Problems – Nov 14, 1996

Problem 1. Set Cover

Several plants in your vegetable garden are dying. You want to figure out what's wrong. Your Gardener's handbook gives the following information about possible symptoms and diseases that cause them.

disease	stunted growth (sg)	yellow leaves (yl)	sticky sap (ss)	wiling branches (wb)
bacterial wilt (BW)	x		x	x
leaf fungus (LF)	x	x		
fusarium wilt (FW)		x		x
clubroot (CR)				x

(a) Suppose the symptoms are given in the order sg, yl, ss, and wb. Use Reggia's set-covering algorithm to find the diagnosis.

Answer:

(b) Your handbook tells you that a diagnostic test is 90% accurate in testing the presence of BW. What is the probability that your plant has BW given the test result is positive? Assume the prior probability of BW is 0.2.

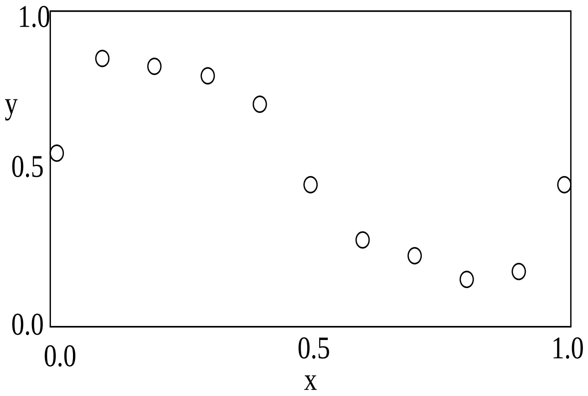
Answer:

(c) Assume $P(sg|LF) = P(sg|BW) = 0.7$. Use the noisy-or assumption to calculate $P(sg|LF,BW)$.

Answer:

Problem 2. Curve Fitting

Consider a test set of 11 data points obtained by sampling a sin function in the interval $(0, 1)$ and adding a small random noise.



(a) Suppose you are to fit the data set with a linear polynomial. Sketch the curve on the diagram. What do you expect the variance of the linear polynomial to be? Circle the correct answer.

A. LOW B. HIGH C. CAN'T TELL

(b) You could obtain a better fit by increasing the order of the polynomial.

Describe how the training error varies as you increase the order of the polynomial (say from 1 to 10).

Answer:

Describe how the predictive error varies as you increase the order of the polynomial (say from 1 to 10).

Answer:

(c) Suppose you do not know the underlying function you want to fit and you do not have a test set yet. Could you estimate the predictive error from the training samples alone? Why or why not?

Answer: