Assignment 10

Pegah Khazaei

23 May 2020

Chapter 18

Problem 1

Part a

Given h, construct a binary tree. Let the root node ($x_1 = 0$?) and all the nodes at depth i of the form ($x_{i+1} = 0$?). The height of this tree is d and it has 2^d leaves. It is obvious that we can allocate one leaf to any possible combination of values for $x_1, x_2, ..., x_d$ with the leaf value being $h(x) = h((x_1, x_2, ..., x_d))$.

Part b

The previous example shows that we can shatter $\{0,1\}^d$. Therefore $VCdim = 2^d$.

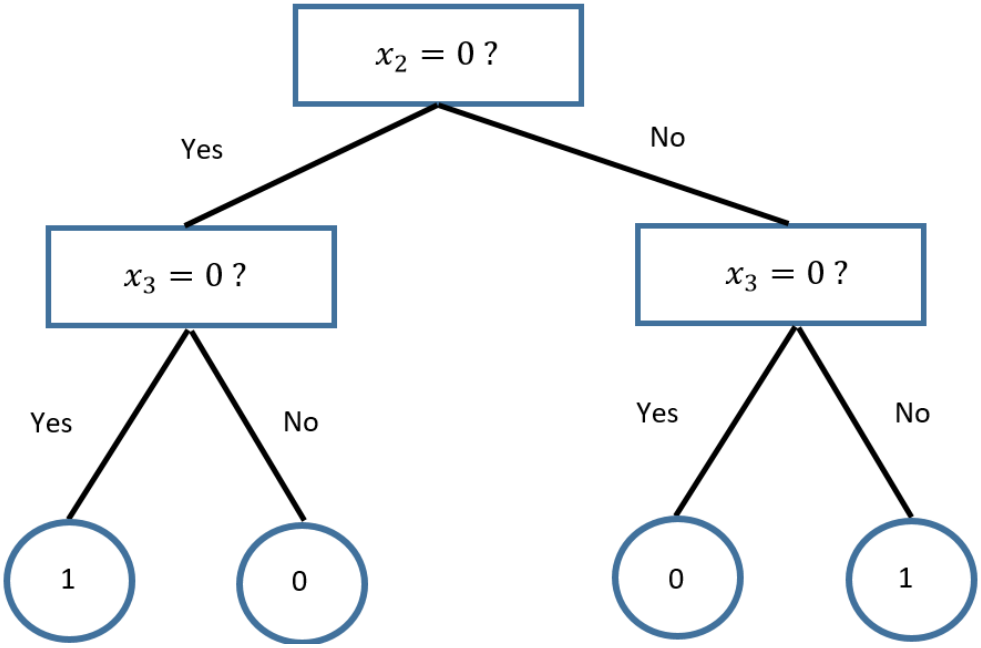
Problem 2

Part a

The information gain for the feature 1 (namely if we choose $x_1 = 0$? as the root) is $H(\frac{1}{2}) - \left(\frac{3}{4}H(\frac{2}{3}) + \frac{1}{4}H(0)\right) \approx 0.22$ in which H is the binary entropy.

The information gain for the feature 2, as well as the feature 3 is $H(\frac{1}{2}) - \left(\frac{1}{2}H(\frac{1}{2}) + \frac{1}{2}H(\frac{1}{2})\right) \approx 0$. Therefore we should pick $x_1 = 0$? as the root and it means that first three examples go down one subtree. If we pick $x_2 = 0$? as the next node $((1,1,1),1)$ and $((1,1,0),0)$ are in the same subtree and we cannot classify them. So we have four examples in the training set, and it follows that the error is $\frac{1}{4}$.

Part b


$x_2 = 0$?
Yes
No
$x_3 = 0$?
$x_3 = 0$?
Yes
No
Yes
No
1
0
0
1