ML.exer.chap18

samira moradzade

January 2022

1 exercise 18.2-1:

in first iteration we compute the information gain:

$$H(Y) = -\frac{1}{2}\log(\frac{1}{2}) - \frac{1}{2}\log(\frac{1}{2}) = 1$$

$$IG(X_1) = H(Y) - H(Y|X_1)$$
$$= 1 - \left[\left(\frac{3}{4}\right)\left(\left(-\frac{2}{3}\right)\log\left(\frac{2}{3}\right)\right) - \frac{1}{3}\log\frac{1}{3} + \frac{1}{4}\left(-0\log 0 - 1\log 1\right)\right]$$
$$= 1 - \left(\frac{3}{4}\right)\left[-\frac{2}{3}\log\left(\frac{2}{3}\right) - \frac{1}{3}\log\left(\frac{1}{3}\right)\right] > 0$$

$$IG(X_2) = H(Y) - H(Y|X_2)$$
$$= 1 - \left[\left(\frac{1}{2}\right)\left(\left(-\frac{1}{2}\right)\log\left(\frac{1}{2}\right)\right) - \frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\left(-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2}\right)\right]$$
$$= 1 - \left[\frac{1}{2}(-1) + \frac{1}{2}(-1)\right] = 0$$

$$IG(X_3) = H(Y) - H(Y|X_3)$$
$$= 1 - \left[\left(\frac{1}{2}\right)\left(\left(-\frac{1}{2}\right)\log\left(\frac{1}{2}\right)\right) - \frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\left(-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2}\right)\right]$$
$$= 1 - \left[\frac{1}{2}(-1) + \frac{1}{2}(-1)\right] = 0$$

so we choose $X_1 = 0$ for begin the tree. 1. for choosing the left node:

$$ID3\left(\left\{\left((1,1,1),1\right),\left((1,0,0),1\right),\left((1,1,0),0\right)\right\},\{x_2,x_3\}\right)$$

we have to compute the info. gain again

$$H(Y) = -\frac{2}{3}\log(\frac{2}{3}) - \frac{1}{3}\log(\frac{1}{3})$$

$$IG(X_2) = H(Y) - H(Y|X_2)$$
$$= H(Y) - \left[\frac{2}{3}\left(-\frac{1}{2}\log\frac{1}{2}\right) - \frac{1}{2}\log\frac{1}{2}\right)$$
$$= H(Y) - \frac{2}{3}$$

$$IG(X_3) = H(Y) - H(Y|X_3)$$
$$= H(Y) - \left[\frac{2}{3}\left(-\frac{1}{2}\log\frac{1}{2}\right) - \frac{1}{2}\log\frac{1}{2}\right)$$
$$= H(Y) - \frac{2}{3}$$

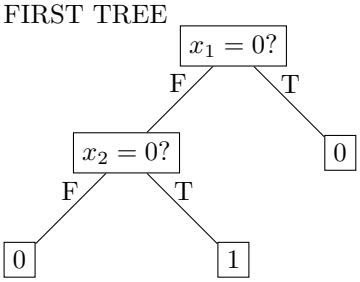
its possible to choose either X_2 or X_3 to have 2 different trees.

2. right node:

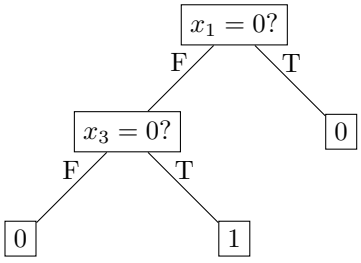
$$ID3\left(\left\{\left(\left(0,0,1\right),0\right)\right\},\left\{x_2,x_3\right\}\right)$$

the only possible label is 0. training error for FIRST tree is $\frac{1}{4}$ because the only mislabeled point is $\left(\left(1,1,1\right),1\right)$

FIRST TREE



SECOND TREE:



and the training error for the second tree is also $\frac{1}{4}$ because the only mislabeled point is $((1,0,0),1)$
so the training error for any tree with the 2 depth with ID3 is at least $\frac{1}{4}$

2 18.2-2:

we want to show the decision tree with the 0 training error.