ISTANBUL TECHNICAL UNIVERSITY, COMPUTER ENGINEERING

BLG527E Machine Learning Homework 4

Anıl Öztürk

504181504

January 7, 2019

QUESTION 3

3.A

$$P(R|W) = \frac{P(W|R)P(R)}{P(W)} \tag{0.1}$$

$$P(W|R) = P(W|S,R)P(S|R) + P(W|\sim S,R)P(\sim S|R) = \mathbf{0.91} \tag{0.2}$$

$$\begin{aligned} P(W) = P(W|S,R)P(S,R) + P(W|\sim S,R)P(\sim S,R) \\ +P(W|S,\sim R)P(S,\sim R) + P(W|\sim S,\sim R)P(\sim S,\sim R) = \mathbf{0.52} \end{aligned} \tag{0.3}$$

$$P(R|W) = \frac{0.91x0.4}{0.52} = \mathbf{0.7} \tag{0.4}$$

3.B

$$\begin{aligned} P(S) = P(S|W,R)\mathbf{P(W,R)} + P(S|\sim W,R)P(\sim W,R) \\ +P(S|W,\sim R)P(W,\sim R) + P(S|\sim W,\sim R)P(\sim W,\sim R) = \mathbf{0.52} \end{aligned} \tag{0.5}$$

$$P(W,R) = \frac{\mathbf{P(W|R)}}{P(R)} \tag{0.6}$$

$$P(W|R) = \frac{\mathbf{P(R|W)}P(W)}{P(R)} \tag{0.7}$$

$P(S)$ definition includes $P(R|W)$. So they are **dependent** to each other.

QUESTION 4

If we do default equal weighted standard bagging, we don't need the training data. We will only work with the test data.

Test data 1 **(1)**;

$$pred_1 = \frac{1+1+1+0+0+1+0}{7} = 0.57\textbf{(1)} \tag{0.8}$$

Test data 2 **(0)**;

$$pred_2 = \frac{0+1+0+1+0+0+0}{7} = 0.29\textbf{(0)} \tag{0.9}$$

We did two predictions and they are correct. We got a bagging outputs between 0 and 1 and rounded them to the closest label. In this situation we also would've take the mod of the model predictions as the label.

	Predicted	
Actual	**0**	**1**
0	1	0
1	0	1

QUESTION 5

State probabilities are;

- S1, S1, S1
- S1, S1, S2
- S1, S2, S1
- S1, S2, S2
- S2, S1, S1
- S2, S1, S2
- S2, S2, S1
- S2, S2, S2

$$\begin{aligned}
&(0.3X0.8X0.8)(0.7X0.7X0.2) + (0.3X0.8X0.2)(0.7X0.7X0.1) \\
&+(0.3X0.2X0.7)(0.7X0.1X0.2) + (0.3X0.2X0.3)(0.7X0.1X0.1) \\
&+(0.7X0.3X0.3)(0.1X0.1X0.1) + (0.7X0.7X0.2)(0.1X0.7X0.1) \\
&+(0.7X0.7X0.3)(0.1X0.7X0.2) + (0.7X0.3X0.7)(0.1X0.1X0.2) \\
&= \mathbf{0.024983}
\end{aligned} \tag{0.10}$$

QUESTION 6

FIRST LEVEL

$$x_1(1) = H[2,1] = -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} = 0.918 \tag{0.11}$$

$$x_1(0) = H[3,2] = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.97 \tag{0.12}$$

$$impurity(x_1) = \frac{3}{8}0.918 + \frac{5}{8}0.97 = \mathbf{0.95} \tag{0.13}$$

$$x_2(1) = H[3,1] = -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4} = 0.81 \tag{0.14}$$

$$x_2(0) = H[2,2] = -\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4} = 1 \tag{0.15}$$

$$impurity(x_2) = \frac{4}{8}0.81 + \frac{4}{8}1 = \mathbf{0.905} \tag{0.16}$$

$$x_3(1) = H[1,2] = -\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3} = 0.92 \tag{0.17}$$

$$x_3(0) = H[4,1] = -\frac{4}{5}\log_2\frac{4}{5} - \frac{1}{5}\log_2\frac{1}{5} = 0.722 \tag{0.18}$$

$$impurity(x_3) = \frac{3}{8}0.92 + \frac{5}{8}0.722 = \mathbf{0.796} \tag{0.19}$$

x_3 has the lower impurity

SECOND LEVEL

X3 = 1

$$x_1(1) = H[0,0] = 0 \tag{0.20}$$

$$x_1(0) = H[1,2] = -\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3} = 0.92 \tag{0.21}$$

$$impurity(x_1) = \mathbf{0.92} \tag{0.22}$$

$$x_2(1) = H[0,1] = 0 \tag{0.23}$$

$$x_2(0) = H[1,1] = 1 \tag{0.24}$$

$$impurity(x_2) = \mathbf{0.66} \tag{0.25}$$

x_2 has the lower impurity

X3 = 0

$$x_1(1) = H[2,1] = -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} = 0.918 \tag{0.26}$$

$$x_1(0) = H[2,0] = 0 \tag{0.27}$$

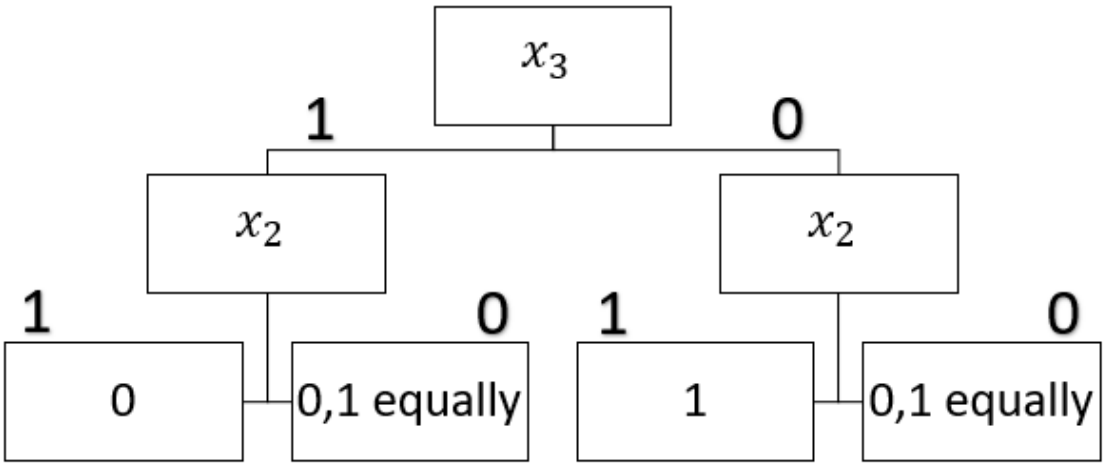
$$impurity(x_1) = \mathbf{0.5508} \tag{0.28}$$

$$x_2(1) = H[3,0] = 0 \tag{0.29}$$

$$x_2(0) = H[1,1] = 1 \tag{0.30}$$

$$impurity(x_2) = \frac{2}{5}1 = \mathbf{0.4} \tag{0.31}$$

x_2 has the lower impurity



QUESTION 8

8.A

$$e1_m = 0.22 \tag{0.32}$$

$$e2_m = 0.18 \tag{0.33}$$

$$\sigma = std(e1_m - e2_m) = 0.1516 \tag{0.34}$$

$$T_{score} = \frac{(e1_m - e2_m)\sqrt{5}}{\sigma} \tag{0.35}$$

$$T_{score} = \frac{(0.04)\sqrt{5}}{0.1516} = \mathbf{0.5857} \tag{0.36}$$

If T_{score} value falls between the $-T_{\alpha/2}$ and $T_{\alpha/2}$, these classifiers are similar.

8.B

$$e_m = \frac{e1_m + e2_m}{2} = 0.2 \tag{0.37}$$

$$SSb = n_1(e1_m - e_m)^2 + n_2(e2_m - e_m)^2 = \mathbf{0.004} \tag{0.38}$$

$$SSw = (e11 - e1_m)^2 + (e12 - e1_m)^2 + ... + (e25 - e2_m)^2 + = \mathbf{0.056} \tag{0.39}$$

$$\frac{SSb/(L-1)}{SSw/(L(K-1))} \sim F_{L-1,L(K-1)} \tag{0.40}$$

$$F = \mathbf{0.5714} \tag{0.41}$$

If F value is lesser than $F_{1,8}$ for given α value, these classifiers are similar.