Assignment 3

Chen Siyi A0194556R   Gao Haochun A0194525Y   He Yuan A0211297H
Wang Pei A0194486M   Wang Zi A0194504E

11/3/2020

Question 1

1.a

P(senior) = 1/2	P(junior) = 1/2
P(marketing | senior) = 3/5	P(marketing | junior) = 1/5
P([31 to 35] | senior) = 2/5	P([31 to 35] | junior) = 1/5
P(>40k | senior) = 3/5	P(>40k | junior) = 1/5

$$P(senior) \times P(marketing \mid senior) \times P([31 to 35] \mid senior) \times P(> 40k \mid senior) = \frac{18}{250}$$

$$P(junior) \times P(marketing \mid junior) \times P([31 to 35] \mid junior) \times P(> 40k \mid junior) = \frac{1}{250}$$

Thus, the naive Bayes model should predict senior.

1.b

Laplace smoothed (k = 3) probability:

P(senior) = 1/2	P(junior) = 1/2
P(sales | senior) = 4/14	P(sales | junior) = 4/14
P([36 to 40] | senior) = 5/17	P([36 to 40] | junior) = 3/17
P(<30k | senior) = 3/14	P(<30k | junior) = 4/14

$$P(senior) \times P(sales \mid senior) \times P([36 to 40] \mid senior) \times P(< 30k \mid senior) = \frac{15}{1666}$$

$$P(junior) \times P(sales \mid junior) \times P([36 to 40] \mid junior) \times P(< 30k \mid junior) = \frac{6}{833} < \frac{15}{1666}$$

Thus, the naive Bayes model should predict senior.

Question 2

2.a

Temperature:

$$odds\ ratio = e^{-0.453} = 0.6357$$

Holding other variables constant, when the temperature increase by 1, the odds of playing tennis will decrease by 36.43%.

Wind:

$$odds\ ratio = e^{0.2365} = 25.4445$$

Holding other variables constant, the odds of playing tennis when the wind is weak is 25.4445 times as that when the wind is strong.

2.b

Null deviance:

$$\hat{\pi} = 12/20 = 3/5$$

$$g(Y_1, Y_2, ..., Y_{20}) = (3/5)^{12}(2/5)^8$$

$$-2ln(L(\beta_0, -)) = -2ln(g(Y_1, Y_2, ..., Y_{20})) = 26.92$$

AIC:

Number of parameters = 2

Number of levels of dependent variable = 2

$$AIC = -2LL + 2x(2 + (2-1)) = 17.211$$

2.c

There are 8 *NO* and 12 *YES* responses on playing tennis. Thus, the number of pairs with different target values is 96 (= 8 x 12). Among these 96 pairs, 5 pairs are discordant (listed below).

Pair	Index1	Index2
1	10	4
2	10	9
3	10	12
4	10	20
5	15	4

Since there are 0 ties, the rest 91 (= 96 - 5) are concordant pairs.

Percent Concordant:

$$91 \div 95 \times 100\% = 94.79\%$$

Percent Discordant:

$$5 \div 95 \times 100\% = 5.21\%$$

2.d

The table below is sorted by the given predict value.

Target	Predicted Value
No	0.0093
No	0.0146
No	0.0228
No	0.0832
No	0.1249
No	0.1249
Yes	0.3572
No	0.5790
Yes	0.5948
Yes	0.6839
Yes	0.6976
No	0.8511
Yes	0.9340
Yes	0.9822
Yes	0.9822
Yes	0.9886
Yes	0.9886
Yes	0.9927
Yes	0.9927
Yes	0.9954

There are 3 possible splitting cases:

Case 1: predict value <= 0.1249 and predict value > 0.1249

In this case, 2 targets with *NO* value will be misclassified.

Accuracy:

$$(20 - 2) \div 20 = 0.9$$

Case 2: predict value <= 0.5790 and predict value > 0.5790

In this case, 1 target with *NO* value and 1 target with *YES* value will be misclassified.

Accuracy:

$$(20 - 1 - 1) \div 20 = 0.9$$

Case 3: predict value <= 0.8511 and predict value > 0.8511

In this case, 4 targets with *YES* value will be misclassified.

Accuracy:

$$(20 - 4) \div 20 = 0.8$$

Thus, let the threshold be 0.1249 and the highest accuracy we can get from this model is 0.9.

2.e

Let x_1 be temperature variable, x_2 be the wind variable (x_2 equals 1 if wind is weak, and 0 otherwise).

$$output = 10.7385 - 0.4530x_1 + 3.2365x_2 = 10.7385 - 0.4530 \times 25 + 3.2365 \times 1 = 2.65$$

Thus, the predicted value is:

$$p = \frac{1}{1 + e^{-2.65}} = 0.9340$$

Since this predicted value is larger than 0.1249, we will play tennis on Day 21.

Question 3

3.a

$$\mathbf{V}\vec{x} = \begin{bmatrix} 1 & -0.5 & 2 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 10 \\ 25 \end{bmatrix} = \begin{bmatrix} 46 \\ 34 \end{bmatrix}$$

$$\vec{y} = \begin{bmatrix} \sigma(46) \\ \sigma(34) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{W}\vec{y} = \begin{bmatrix} -1 & 5 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 3$$

Hence, $o = \sigma(3) = 0.9526$

3.b

$$\delta_o = (d - o)o(1 - o) = (1 - 0.9526) \times 0.9526 \times (1 - 0.9526) = 2.1402 \times 10^{-3}$$

Update the first row of $\mathbf{W}$:

$$\mathbf{W_1}^T \leftarrow \mathbf{W_1}^T + \eta\delta_o\vec{y} = \begin{bmatrix} -1 \\ 5 \\ -1 \end{bmatrix} + 1 \times 2.1402 \times 10^{-3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.998 \\ 5.002 \\ -0.998 \end{bmatrix}$$

Hence, the updated weight is $\begin{bmatrix} -0.998 \\ 5.002 \\ -0.998 \end{bmatrix}$, and the bias value for the output unit is updated to -0.998.