Introduction to Machine Learning, Fall 2014 - Exercise session IV

Rodion "rodde" Efremov 013593012

November 21, 2014

Problem 1 (6 points)

(a)

We seek to compute

$$P(Y \mid \mathbf{X}) = \frac{P(\mathbf{X} \mid Y)P(Y)}{P(\mathbf{X})},$$

where

$$P(\mathbf{X}) = \sum_{i=0}^{2} P(\mathbf{X} | Y = i) P(Y = i).$$

(0, 0)

If $\mathbf{X} = (0,0)$,

$$P((0,0)) = \sum_{i=0}^{2} P((0,0) | Y = i)P(Y = i)$$
$$= 0.2 \times 0.4 + 0.6 \times 0.3 + 0.1 \times 0.3$$
$$= 0.08 + 0.18 + 0.03$$
$$= 0.29.$$

$$P(Y = 0 | (0,0)) = \frac{P((0,0) | Y = 0)P(Y = 0)}{P((0,0))}$$

$$= \frac{0.2 \times 0.4}{0.29}$$

$$= \frac{0.08}{0.29}$$

$$\approx 0.276,$$

$$\begin{split} P(Y=1 \,|\, (0,0)) &= \frac{P((0,0) \,|\, Y=1) P(Y=1)}{P((0,0))} \\ &= \frac{0.6 \times 0.3}{0.29} \\ &= \frac{0.18}{0.29} \\ &\approx 0.621, \end{split}$$

$$\begin{split} P(Y=2 \,|\, (0,0)) &= \frac{P((0,0) \,|\, Y=2) P(Y=2)}{P((0,0))} \\ &= \frac{0.1 \times 0.3}{0.29} \\ &= \frac{0.03}{0.29} \\ &\approx 0.103. \end{split}$$

(0, 1)

If $\mathbf{X} = (0, 1)$,

$$P((0,1)) = \sum_{i=0}^{2} P((0,1) | Y = i)P(Y = i)$$
$$= 0.4 \times 0.4 + 0.1 \times 0.3 + 0.3 \times 0.3$$
$$= 0.16 + 0.03 + 0.09$$
$$= 0.28.$$

$$\begin{split} P(Y=0\,|(0,1)) &= \frac{P((0,1)\,|\,Y=0)P(Y=0)}{0.28} \\ &= \frac{0.4\times0.4}{0.28} \\ &= \frac{0.16}{0.28} \\ &\approx 0.571, \end{split}$$

$$\begin{split} P(Y=1 \,|\, (0,1)) &= \frac{P((0,1) \,|\, Y=1) P(Y=1)}{0.28} \\ &= \frac{0.1 \times 0.3}{0.28} \\ &= \frac{0.03}{0.28} \\ &\approx 0.107, \end{split}$$

$$P(Y = 2 | (0,1)) = \frac{P((0,1) | Y = 2)P(Y = 2)}{0.28}$$
$$= \frac{0.3 \times 0.3}{0.28}$$
$$= \frac{0.09}{0.28}$$
$$\approx 0.321.$$

(0, 2)

If $\mathbf{X} = (0, 2)$,

$$P((0,2)) = \sum_{i=0}^{2} P((0,2) | Y = i)P(Y = i)$$

$$= 0.0 \times 0.4 + 0.1 \times 0.3 + 0.2 \times 0.3$$

$$= 0.03 + 0.06$$

$$= 0.09,$$

$$P(Y = 0 | (0,2)) = \frac{P((0,2) | Y = 0)P(Y = 0)}{0.09}$$
$$= \frac{0.0 \times 0.4}{0.09}$$
$$= 0,$$

$$P(Y = 1 | (0,2)) = \frac{P((0,2) | Y = 1)P(Y = 1)}{0.09}$$

$$= \frac{0.1 \times 0.3}{0.09}$$

$$= \frac{0.03}{0.09}$$

$$\approx 0.333,$$

$$P(Y = 2 | (0,2)) = \frac{P((0,2) | Y = 2)P(Y = 2)}{0.09}$$

$$= \frac{0.2 \times 0.3}{0.09}$$

$$= \frac{0.06}{0.09}$$

$$\approx 0.666.$$

If $\mathbf{X} = (1, 0)$,

$$P((1,0)) = \sum_{i=0}^{2} P((1,0) | Y = i)P(Y = i)$$
$$= 0.1 \times 0.4 + 0.1 \times 0.3 + 0.4 \times 0.3$$
$$= 0.04 + 0.03 + 0.12$$
$$= 0.19,$$

$$P(Y = 0 | (1,0)) = \frac{P((1,0) | Y = 0)P(Y = 0)}{0.19}$$

$$= \frac{0.1 \times 0.4}{0.19}$$

$$= \frac{0.04}{0.19}$$

$$\approx 0.211,$$

$$\begin{split} P(Y=1 \,|\, (1,0)) &= \frac{P((1,0) \,|\, Y=1)P(Y=1)}{0.19} \\ &= \frac{0.1 \times 0.3}{0.19} \\ &= \frac{0.03}{0.19} \\ &\approx 0.158, \end{split}$$

$$P(Y = 2 | (1,0)) = \frac{P((1,0) | Y = 2)P(Y = 2)}{0.19}$$

$$= \frac{0.4 \times 0.3}{0.19}$$

$$= \frac{0.12}{0.19}$$

$$\approx 0.632.$$

(1, 1)

If X = (1, 1),

$$P((1,1)) = \sum_{i=0}^{2} P((1,1) | Y = i)P(Y = i)$$

$$= 0.2 \times 0.4 + 0.1 \times 0.3 + 0.0 \times 0.3$$

$$= 0.08 + 0.03$$

$$= 0.11,$$

$$P(Y = 0 | (1,1)) = \frac{P((1,1) | Y = 0)P(Y = 0)}{0.11}$$

$$= \frac{0.2 \times 0.4}{0.11}$$

$$= \frac{0.08}{0.11}$$

$$\approx 0.727,$$

$$P(Y = 1 | (1,1)) = \frac{P((1,1) | Y = 1)P(Y = 1)}{0.11}$$

$$= \frac{0.1 \times 0.3}{0.11}$$

$$= \frac{0.03}{0.11}$$

$$\approx 0.273,$$

$$P(Y = 2 | (1,1)) = \frac{P((1,1) | Y = 2)P(Y = 2)}{0.11}$$
$$= \frac{0.0 \times 0.3}{0.11}$$
$$= 0.0.$$

(1, 2)

If
$$\mathbf{X} = (1, 2)$$
,

$$P((1,2)) = \sum_{i=0}^{2} P((1,2) | Y = i) P(Y = i)$$

= 0.1 × 0.4 + 0.0 × 0.3 + 0.0 × 0.3
= 0.04,

$$P(Y = 0 | (1,2)) = \frac{P((1,2) | Y = 0)P(Y = 0)}{0.04}$$
$$= \frac{0.1 \times 0.4}{0.04}$$
$$= 1.0,$$

$$P(Y = 1 | (1,2)) = \frac{P((1,2) | Y = 1)P(Y = 1)}{0.04}$$
$$= \frac{0.0 \times 0.3}{0.11}$$
$$= 0.0,$$

$$P(Y = 2 | (1,2)) = \frac{P((1,2) | Y = 2)P(Y = 2)}{0.04}$$
$$= \frac{0.0 \times 0.3}{0.11}$$
$$= 0.0.$$

Summary

	Y = 0	Y = 1	Y=2
(0,0)	0.276	0.621	0.103
(0,1)	0.571	0.107	0.321
(0,2)	0.0	0.333	0.667
(1,0)	0.211	0.158	0.631
(1,1)	0.727	0.273	0.0
(1,2)	1.0	0.0	0.0

The optimal Bayes classifier is

$$Y(\mathbf{X}) = \begin{cases} 0 & \text{if } \mathbf{X} \in \{(0,1), (1,1), (1,2)\} \\ 1 & \text{if } \mathbf{X} = (0,0) \\ 2 & \text{if } \mathbf{X} \in \{(0,2), (1,0)\}. \end{cases}$$

(b)

The error rate is

$$\sum_{x_1=0}^{1} \sum_{x_2=0}^{2} P(x_1, x_2) E(x_1, x_2),$$

where $E(x_1, x_2)$ is the error probability at point (x_1, x_2) , which is defined as the

X	$E(\mathbf{X})$	$P(\mathbf{X})$	$\mid E(\mathbf{X})P(\mathbf{X}) \mid$
(0,0)	0.379	0.29	0.110
(0,1)	0.428	0.28	0.120
(0,2)	0.333	0.09	0.030
(1,0)	0.369	0.19	0.070
(1,1)	0.273	0.11	0.030
$(1\ 2)$	0	0.04	0

So the

sum of two least probabilities.

error rate is 0.110 + 0.120 + 0.030 + 0.070 + 0.030 = 0.36.

(c)

For naïve Bayes classifier in this context, we have

$$P(\mathbf{X} \mid Y = y) = \prod_{i=1}^{2} P(X_i \mid Y = y) = P(X_1 \mid Y = y)P(X_2 \mid Y = y),$$

so we seek to compute

$$P(Y = y \mid \mathbf{X}) = \frac{P(\mathbf{X} \mid Y = y)P(Y = y)}{P(\mathbf{X})} = \frac{P(X_1 \mid Y = y)P(X_2 \mid Y = y)P(Y = y)}{P(\mathbf{X})},$$

where

$$P(\mathbf{X}) = \sum_{i=0}^{2} P(\mathbf{X} | Y = i) P(Y = i)$$
$$= \sum_{i=0}^{2} P(X_1 | Y = i) P(X_2 | Y = i) P(Y = i).$$

(0,0)

If X = (0, 0), and

$$P(\mathbf{X}) = \sum_{i=0}^{2} P(X_1 \mid Y = i) P(X_2 \mid Y = i) P(Y = i)$$

$$= (0.2 + 0.4 + 0.0) \times (0.2 + 0.1) \times 0.4$$

$$+ (0.6 + 0.1 + 0.1) \times (0.6 + 0.1) \times 0.3$$

$$+ (0.1 + 0.3 + 0.2) \times (0.1 + 0.4) \times 0.3$$

$$= 0.6 \times 0.3 \times 0.4 + 0.8 \times 0.7 \times 0.3 + 0.6 \times 0.5 \times 0.3$$

$$= 0.072 + 0.168 + 0.09$$

$$= 0.33.$$

$$\begin{split} P(Y=0 \,|\, (0,0)) &= \frac{P(X_1 \,|\, Y=0) P(X_2 \,|\, Y=0) P(Y=0)}{0.33} \\ &= \frac{(0.2+0.4+0.0) \times (0.2+0.1) \times 0.4}{0.33} \\ &= \frac{0.6 \times 0.3 \times 0.4}{0.33} \\ &= \frac{0.072}{0.33} \\ &\approx 0.218, \end{split}$$

$$\begin{split} P(Y=1 \,|\, (0,0)) &= \frac{P(X_1 \,|\, Y=1)P(X_2 \,|\, Y=1)P(Y=1)}{0.33} \\ &= \frac{(0.6+0.1+0.1)\times(0.6+0.1)\times0.3}{0.33} \\ &= \frac{0.8\times0.7\times0.3}{0.33} \\ &= \frac{0.168}{0.33} \\ &\approx 0.509, \end{split}$$

$$P(Y = 2 \mid (0,0)) = \frac{P(X_1 \mid Y = 2)P(X_2 \mid Y = 2)P(Y = 2)}{0.33}$$

$$= \frac{(0.1 + 0.3 + 0.2) \times (0.1 + 0.4) \times 0.3}{0.33}$$

$$= \frac{0.6 \times 0.5 \times 0.3}{0.33}$$

$$= \frac{0.09}{0.33}$$

$$\approx 0.273.$$

(0,1)

If X = (0, 1), and

$$P(\mathbf{X}) = \sum_{i=0}^{2} P(X_1 \mid Y = i) P(X_2 \mid Y = i) P(Y = i)$$

$$= (0.2 + 0.4 + 0.0) \times (0.4 + 0.2) \times 0.4$$

$$+ (0.6 + 0.1 + 0.1) \times (0.1 + 0.1) \times 0.3$$

$$+ (0.1 + 0.3 + 0.2) \times (0.3 + 0.0) \times 0.3$$

$$= 0.6 \times 0.6 \times 0.4 + 0.8 \times 0.2 \times 0.3 + 0.6 \times 0.3 \times 0.3$$

$$= 0.144 + 0.048 + 0.054$$

$$= 0.246.$$

$$P(Y = 0 | (0,1)) = \frac{P(X_1 = 0 | Y = 0)P(X_2 = 1 | Y = 0)P(Y = 0)}{0.246}$$

$$= \frac{(0.2 + 0.4 + 0.0) \times (0.4 + 0.2) \times 0.4}{0.246}$$

$$= \frac{0.6 \times 0.6 \times 0.4}{0.246}$$

$$= \frac{0.144}{0.246}$$

$$\approx 0.585,$$

$$\begin{split} P(Y=1 \,|\, (0,1)) &= \frac{P(X_1=0 \,|\, Y=1)P(X_2=1 \,|\, Y=1)P(Y=1)}{0.246} \\ &= \frac{(0.6+0.1+0.1)\times(0.1+0.1)\times0.3}{0.246} \\ &= \frac{0.8\times0.2\times0.3}{0.246} \\ &= \frac{0.048}{0.246} \\ &\approx 0.195, \end{split}$$

$$P(Y = 2 \mid (0,1)) = \frac{P(X_1 = 0 \mid Y = 2)P(X_2 = 1 \mid Y = 2)P(Y = 2)}{0.246}$$

$$= \frac{(0.1 + 0.3 + 0.2) \times (0.3 + 0.0) \times 0.3}{0.246}$$

$$= \frac{0.6 \times 0.3 \times 0.3}{0.246}$$

$$= \frac{0.054}{0.246}$$

$$\approx 0.219.$$

(0, 2)

If X = (0, 2), and

$$P(\mathbf{X}) = \sum_{i=0}^{2} P(X_1 \mid Y = i) P(X_2 \mid Y = i) P(Y = i)$$

$$= (0.2 + 0.4 + 0.0) \times (0.0 + 0.1) \times 0.4$$

$$+ (0.6 + 0.1 + 0.1) \times (0.1 + 0.0) \times 0.3$$

$$+ (0.1 + 0.3 + 0.2) \times (0.2 + 0.0) \times 0.3$$

$$= 0.6 \times 0.1 \times 0.4 + 0.8 \times 0.1 \times 0.3 + 0.6 \times 0.2 \times 0.3$$

$$= 0.024 + 0.024 + 0.036$$

$$= 0.084.$$

$$P(Y = 0 | (0, 2)) = \frac{P(X_1 = 0 | Y = 0)P(X_2 = 2 | Y = 0)P(Y = 0)}{0.084}$$

$$= \frac{(0.2 + 0.4 + 0.0) \times (0.0 + 0.1) \times 0.4}{0.084}$$

$$= \frac{0.6 \times 0.1 \times 0.4}{0.084}$$

$$= \frac{0.024}{0.084}$$

$$\approx 0.286,$$

$$\begin{split} P(Y=1 \,|\, (0,2)) &= \frac{P(X_1=0 \,|\, Y=1)P(X_2=2 \,|\, Y=1)P(Y=1)}{0.084} \\ &= \frac{(0.6+0.1+0.1)\times(0.1+0.0)\times0.3}{0.084} \\ &= \frac{0.8\times0.1\times0.3}{0.084} \\ &= \frac{0.024}{0.084} \\ &\approx 0.286, \end{split}$$

$$\begin{split} P(Y=2 \,|\, (0,2)) &= \frac{P(X_1=0 \,|\, Y=2)P(X_2=2 \,|\, Y=2)P(Y=2)}{0.084} \\ &= \frac{(0.1+0.3+0.2)\times(0.2+0.0)\times0.3}{0.084} \\ &= \frac{0.6\times0.2\times0.3}{0.084} \\ &= \frac{0.036}{0.084} \\ &\approx 0.429. \end{split}$$

(1,0)

If X = (1, 0), and

$$P(\mathbf{X}) = \sum_{i=0}^{2} P(X_1 \mid Y = i) P(X_2 \mid Y = i) P(Y = i)$$

$$= (0.1 + 0.2 + 0.1) \times (0.2 + 0.1) \times 0.4$$

$$+ (0.1 + 0.1 + 0.0) \times (0.6 + 0.1) \times 0.3$$

$$+ (0.4 + 0.0 + 0.0) \times (0.1 + 0.4) \times 0.3$$

$$= 0.4 \times 0.3 \times 0.4 + 0.2 \times 0.7 \times 0.3 + 0.4 \times 0.5 \times 0.3$$

$$= 0.048 + 0.042 + 0.06$$

$$= 0.15.$$

$$\begin{split} P(Y=0 \,|\, (1,0)) &= \frac{P(X_1=1 \,|\, Y=0)P(X_2=0 \,|\, Y=0)P(Y=0)}{0.15} \\ &= \frac{(0.1+0.2+0.1)\times(0.2+0.1)\times0.4}{0.15} \\ &= \frac{0.4\times0.3\times0.4}{0.15} \\ &= \frac{0.048}{0.15} \\ &= 0.32, \end{split}$$

$$P(Y = 1 | (1,0)) = \frac{P(X_1 = 1 | Y = 1)P(X_2 = 0 | Y = 1)P(Y = 1)}{0.15}$$

$$= \frac{(0.1 + 0.1 + 0.0) \times (0.6 + 0.1) \times 0.3}{0.15}$$

$$= \frac{0.2 \times 0.7 \times 0.3}{0.15}$$

$$= \frac{0.042}{0.15}$$

$$= 0.28,$$

$$P(Y = 2 | (1,0)) = \frac{P(X_1 = 1 | Y = 2)P(X_2 = 0 | Y = 2)P(Y = 2)}{0.15}$$

$$= \frac{(0.4 + 0.0 + 0.0) \times (0.1 + 0.4) \times 0.3}{0.15}$$

$$= \frac{0.4 \times 0.5 \times 0.3}{0.15}$$

$$= \frac{0.06}{0.15}$$

$$= 0.4.$$

(1, 1)

If X = (1, 1), and

$$P(\mathbf{X}) = \sum_{i=0}^{2} P(X_1 \mid Y = i) P(X_2 \mid Y = i) P(Y = i)$$

$$= (0.1 + 0.2 + 0.1) \times (0.4 + 0.2) \times 0.4$$

$$+ (0.1 + 0.1 + 0.0) \times (0.1 + 0.1) \times 0.3$$

$$+ (0.4 + 0.0 + 0.0) \times (0.3 + 0.0) \times 0.3$$

$$= 0.4 \times 0.6 \times 0.4 + 0.2 \times 0.2 \times 0.3 + 0.4 \times 0.3 \times 0.3$$

$$= 0.096 + 0.012 + 0.036$$

$$= 0.144.$$

$$P(Y = 0 | (1,1)) = \frac{P(X_1 = 1 | Y = 0)P(X_2 = 1 | Y = 0)P(Y = 0)}{0.144}$$

$$= \frac{(0.1 + 0.2 + 0.1) \times (0.4 + 0.2) \times 0.4}{0.144}$$

$$= \frac{0.4 \times 0.6 \times 0.4}{0.144}$$

$$= \frac{0.096}{0.144}$$

$$\approx 0.667,$$

$$\begin{split} P(Y=1 \,|\, (1,1)) &= \frac{P(X_1=1 \,|\, Y=1)P(X_2=1 \,|\, Y=1)P(Y=1)}{0.144} \\ &= \frac{(0.1+0.1+0.0)\times(0.1+0.1)\times0.3}{0.144} \\ &= \frac{0.2\times0.2\times0.3}{0.144} \\ &= \frac{0.012}{0.144} \\ &\approx 0.083, \end{split}$$

$$P(Y = 2 | (1,1)) = \frac{P(X_1 = 1 | Y = 2)P(X_2 = 1 | Y = 2)P(Y = 2)}{0.144}$$

$$= \frac{(0.4 + 0.0 + 0.0) \times (0.3 + 0.0) \times 0.3}{0.144}$$

$$= \frac{0.4 \times 0.3 \times 0.3}{0.144}$$

$$= \frac{0.036}{0.144}$$

$$= 0.25.$$

(1, 2)

If X = (1, 2), and

$$P(\mathbf{X}) = \sum_{i=0}^{2} P(X_1 \mid Y = i) P(X_2 \mid Y = i) P(Y = i)$$

$$= (0.1 + 0.2 + 0.1) \times (0.0 + 0.1) \times 0.4$$

$$+ (0.1 + 0.1 + 0.0) \times (0.1 + 0.0) \times 0.3$$

$$+ (0.4 + 0.0 + 0.0) \times (0.2 + 0.0) \times 0.3$$

$$= 0.4 \times 0.1 \times 0.4 + 0.2 \times 0.1 \times 0.3 + 0.4 \times 0.2 \times 0.3$$

$$= 0.016 + 0.006 + 0.024$$

$$= 0.046.$$

$$P(Y = 0 | (1,2)) = \frac{P(X_1 = 1 | Y = 0)P(X_2 = 2 | Y = 0)P(Y = 0)}{0.046}$$

$$= \frac{(0.1 + 0.2 + 0.1) \times (0.0 + 0.1) \times 0.4}{0.046}$$

$$= \frac{0.4 \times 0.1 \times 0.4}{0.046}$$

$$= \frac{0.016}{0.046}$$

$$\approx 0.348,$$

$$\begin{split} P(Y=1 \,|\, (1,1)) &= \frac{P(X_1=1 \,|\, Y=1)P(X_2=2 \,|\, Y=1)P(Y=1)}{0.046} \\ &= \frac{(0.1+0.1+0.0)\times(0.1+0.0)\times0.3}{0.046} \\ &= \frac{0.2\times0.1\times0.3}{0.046} \\ &= \frac{0.006}{0.046} \\ &\approx 0.130, \end{split}$$

$$P(Y = 2 | (1,1)) = \frac{P(X_1 = 1 | Y = 2)P(X_2 = 2 | Y = 2)P(Y = 2)}{0.046}$$

$$= \frac{(0.4 + 0.0 + 0.0) \times (0.2 + 0.0) \times 0.3}{0.046}$$

$$= \frac{0.4 \times 0.2 \times 0.3}{0.046}$$

$$= \frac{0.024}{0.046}$$

$$= 0.522.$$

Summary

	Y = 0	Y = 1	Y = 2
(0,0)	0.218	0.509	0.273
(0,1)	0.585	0.195	0.219
(0,2)	0.286	0.286	0.429
(1,0)	0.32	0.28	0.4
(1,1)	0.667	0.083	0.25
(1,2)	0.348	0.130	0.522

The Naïve Bayes classifier is

$$Y(\mathbf{X}) = \begin{cases} 0 & \text{if } \mathbf{X} \in \{(0,1), (1,1)\} \\ 1 & \text{if } \mathbf{X} = (0,0) \\ 2 & \text{if } \mathbf{X} \in \{(0,2), (1,0), (1,2)\}. \end{cases}$$

(d)

X	$E(\mathbf{X})$	$P(\mathbf{X})$	$E(\mathbf{X})P(\mathbf{X})$	
(0,0)	0.491	0.330	0.162	
(0,1)	0.415	0.246	0.102	
(0,2)	0.571	0.084	0.048	
(1,0)	0.600	0.150	0.09	
(1,1)	0.333	0.144	0.048	
(1,2)	0.478	0.046	0.022	
0.090 + 0.048 + 0.022 = 0.472.				

So the error rate is 0.162+0.102+0.048+

(e)

The error rate of naïve Bayes classifier is 31% higher than that of optimal Bayes classifier. And no, I really had to calculate all the stuff to "guess".

Problem 2 (3 points)

Problem 3 (15 points)