20161190_TheoreticalHW4_leeminjae

Exercise 1

1.1

Probability : To find value about some event occur when given distribution. At countable event, probability value is same with likelihood value. But at uncountable event, probability is not same with likelihood.

Likelihood : To find appropriate distribution when given fixed value. At countable event, likelihood value is same with probability value. And at uncountable event, likelihood value is same with probability density function value.

1.2

Difference between Sampling and Resampling: Sampling is the extraction from a population. Sameple is not perfectly same with population. So there are some noise. Resampling is to extract a subset of samples from the sample that I have already. Therefore, the same sample is used several times to measure performance.

Advantage does Resampling: Resampling can improve accuracy. Using limited data several times so solve the problem than imbalanced data. Also resampling need fewer assumption. For example, there are bootstrapping, k-fold and so on.

Exercise 2

2.1

At our ppt slide, logistic regression

$\frac{odds(X+1)}{odds(X)} = e^{\beta}, \quad odds = \frac{P(Y=1)}{1-P(Y=1)}$

we know the β value -0.0003 so we can calculate

$\frac{odds(X+1)}{odds(X)} = e^{\beta} = e^{-0.0003} = 0.9997$

so we can think that
$Odds(X+1) = 0.9997 \times Odd(X)$

this mean **Virginica** is more probable than **Versicolor**

So our Interpretaion is

$\therefore$ **Species Virginica is 1.0003 times more probable than Species Versicolor when the length of sepals Increase**

2.2

$P(Y=1) = \frac{exp(X\beta)}{1+exp(X\beta)}, \quad 1-P(Y=1) = \frac{1}{1+exp(X\beta)}$

$log(\frac{P(Y=1)}{1-P(Y=1)}) = X\beta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p, \quad (\beta_0 = 42.6)$

$X_1\beta = \beta_0 + (-0.0003 \times 6.4) + (6.7 \times 3.1) + (-9.4 \times 4.3) + (-18.3 \times 1.3) = -0.84192$
$X_2\beta = \beta_0 + (-0.0003 \times 6.9) + (6.7 \times 3.0) + (-9.4 \times 3.9) + (-18.3 \times 1.4) = 0.41793$

$P(Y_1 = 1|x_1) = \frac{e^{-0.84192}}{1+e^{-0.84192}} = 0.30113$
$P(Y_1 = 0|x_1) = 1 - 0.30113 = 0.69887$
$P(Y_2 = 1|x_2) = \frac{e^{0.41793}}{1+e^{0.41793}} = 0.603$
$P(Y_2 = 0|x_2) = 1 - 0.6030 = 0.3970$

Above the result,
$\therefore$ **x1 is Virginica and X2 is Versicolor!**

Exercise 3

3.1

$K(x,z) = (x^T z)^2 = x_1^2 z_1^2 + 2x_1 x_2 z_1 z_2 + x_2^2 z_2^2 \quad x = (x_1, x_2), z = (z_1, z_2)$

$K(x,z) = \phi(x)^T \phi(z)$

So we can know that

$\phi(x) = \begin{bmatrix} x_1^2 \\ \sqrt{2}x_1 x_2 \\ x_2^2 \end{bmatrix}, \quad \phi(z) = \begin{bmatrix} z_1^2 \\ \sqrt{2}z_1 z_2 \\ z_2^2 \end{bmatrix}$

3.2

We get the formaula $\phi(x)$ at 3.1 so input x_1, x_2, x_3 value.

$\phi(x) = \begin{bmatrix} x_1^2 \\ \sqrt{2}x_1 x_2 \\ x_2^2 \end{bmatrix}$

$x_1 = (1,2) \quad \phi(x_1) = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$

$x_2 = (4,3) \quad \phi(x_2) = \begin{bmatrix} 16 \\ 12\sqrt{2} \\ 9 \end{bmatrix}$

$x_3 = (2,0) \quad \phi(x_3) = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$

3.3

$K = \begin{bmatrix} K(x1,x1) & K(x1,x2) & K(x1,x3) \\ K(x2,x1) & K(x2,x2) & K(x2,x3) \\ K(x3,x1) & K(x3,x2) & K(x3,x3) \end{bmatrix}$

We can calculate by $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$. So,

$K(x1,x1) = 25$
$K(x1,x2) = K(x2,x1) = 100$
$K(x1,x3) = K(x3,x1) = 4$
$K(x2,x2) = 625$
$K(x2,x3) = K(x3,x2) = 64$
$K(x3,x3) = 16$

$\therefore K = \begin{bmatrix} 25 & 100 & 4 \\ 100 & 625 & 64 \\ 4 & 64 & 16 \end{bmatrix}$

Exercise 4

4.1

by definition of naive bayes, we can say

$P(y=A|X1=a, X2=b) = \frac{(P(X1=a|y=A) \times P(X2=b|y=A)) \times P(y=A)}{P(X1=a) \times P(X2=b)}$

$P(X1=a|y=A) = \frac{1}{2}, \quad P(X1=a) = \frac{3}{7}$
$P(X2=b|y=A) = \frac{1}{2}, \quad P(X2=b) = \frac{4}{7}$
$P(y=A) = \frac{4}{7}$

$\therefore P(y=A|X1=a, X2=b) = \frac{7}{12}$

4.2

To determine whether A or B, I calculate for each case. y = A and y = B

$P(y=A|X1=a, X2=b, x3=b) = \frac{(P(X1=a|y=A) \times P(X2=b|y=A) \times P(X3=b|y=A)) \times P(y=A)}{P(X1=a) \times P(X2=b) \times P(X3=b)}$

$P(X1=a|y=A) = \frac{1}{2}, \quad P(X1=a) = \frac{3}{7}$
$P(X2=b|y=A) = \frac{1}{2}, \quad P(X2=b) = \frac{4}{7}$
$P(X3=b|y=A) = \frac{1}{2}, \quad P(X3=b) = \frac{3}{7}$
$P(y=A) = \frac{4}{7}$

$\therefore P(y=A|X1=a, X2=b, x3=b) = \frac{49}{72}$

$P(y=B|X1=a, X2=b, x3=b) = \frac{(P(X1=a|y=B) \times P(X2=b|y=B) \times P(X3=b|y=B)) \times P(y=B)}{P(X1=a) \times P(X2=b) \times P(X3=b)}$

$P(X1=a|y=B) = \frac{1}{3}, \quad P(X1=a) = \frac{3}{7}$
$P(X2=b|y=B) = \frac{2}{3}, \quad P(X2=b) = \frac{4}{7}$
$P(X3=b|y=B) = \frac{1}{3}, \quad P(X3=b) = \frac{3}{7}$
$P(y=B) = \frac{3}{7}$

$\therefore P(y=B|X1=a, X2=b, x3=b) = \frac{49}{162}$

Because P(y=A|X1 = a,X2= b,x3= b) is higher than P(y=B|X1 = a,X2= b,x3= b),

$\therefore$ **The new data being assigned to Class B.**

Exercise 5

Because This problem is 10 points on final, I prepared a conceptual question that can be solved quickly.
By using given dataset find the values! -ve point A(-1,1), B(1,3) / +ve C(-4,2).
Exercise 5.1 Find kernel matrix K (2pt)
Exercise 5.2 Find $\alpha_1, \alpha_2, \alpha_3$ and b (2pt)
Exercise 5.3 Find w and hyperplane equation. (3pt)
Exercise 5.4 Prove margin and find margin (3pt)

5.1

$K = \begin{bmatrix} K(A,A) & K(A,B) & K(A,C) \\ K(B,A) & K(B,B) & K(B,C) \\ K(C,A) & K(C,B) & K(C,C) \end{bmatrix}$

We can calculate by $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$. So,

$K(A,A) = 2$
$K(A,B) = K(B,A) = 2$
$K(A,C) = K(C,A) = 6$
$K(B,B) = 10$
$K(B,C) = K(C,B) = -2$
$K(C,C) = 20$

$\therefore K = \begin{bmatrix} 2 & 2 & 6 \\ 2 & 10 & -2 \\ -6 & -2 & 20 \end{bmatrix}$

5.2

$\Sigma \alpha_i y_i = 0$

$\Sigma \alpha_i y_i K(x_i, x) + b = 1$

$\Sigma \alpha_i y_i K(x_i, x) + b = -1$

so we can set this!

$-\alpha_1 - \alpha_2 + \alpha_3 = 0$

$-2\alpha_1 - 2\alpha_2 - 6\alpha_3 + b = -1$

$-2\alpha_1 - 10\alpha_2 - 2\alpha_3 + b = -1$

$6\alpha_1 + 2\alpha_2 + 20\alpha_3 + b = 1$

solving system of equation
$\therefore \alpha_1 = \frac{1}{32} \quad \alpha_2 = \frac{1}{32} \quad \alpha_3 = \frac{1}{16} \quad b = -\frac{1}{2}$

5.3

$W = \Sigma \alpha_i y_i x_i = -\frac{1}{32} \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \frac{1}{32} \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \frac{1}{16} \begin{bmatrix} 4 \\ -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \end{bmatrix}$

$Wx + b = 0$

$\therefore$ hyperplane equation is $x_1 - x_2 - 2 = 0, \quad W = \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \end{bmatrix}$

5.4

we know, $x^+ = x^- + \lambda w$

$w^T x^+ + b = w^T(x^- + \lambda w) + +b = 1$

$w^T x^- + b + \lambda w^T w = 1$

$\lambda = \frac{2}{w^T}$

we know, Margin $= ||x^+ - x^-||_2$

$= ||x^- + \lambda w + -x^-||_2 = ||\lambda w||_2$

$= \lambda \sqrt{w^T w} = \frac{2}{\sqrt{w^T w}} = \frac{2}{||w||_2}$

$\therefore$ Margin $= \frac{2}{||w||_2}$

our $W = \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \end{bmatrix}$

so, Margin = $4\sqrt{2}$