Homework 2

Hao Xu
hx2208

Problem 1.

Answer:

(a) $\mu_{y,j} = \frac{\sum_{i=1}^{n} \mathbb{1}\{y_i=y\} x_{i,j}}{\sum_{i=1}^{n} \mathbb{1}\{y_i=y\}}$

(b) Training error rate is: 0.2163. Test error rate is: 0.3760

(c) Training error rate is: 0.05779. Test error rate is: 0.1313.

(d) Most positive word list: $\left[firearms, occupied, israelis, serdar, argic, appressian, ohanus, melkonian, sahak, villages, cramer, armenia, cpr, sdqa, handgun, optilink, palestine, firearm, budget, arabs\right]$

Most negative word list: $\left[athos, atheism, atheists, clh, teachings, revelation, testament, livesey, atheist, solntze, wpd, scriptures, theology, believer, ksand, alink, benedikt, jesus, prophet, mozumder\right]$

Problem 2.

Answer:

(a) According to the definition: $f^* = min\left\{\frac{2}{3}c\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \frac{1}{3}\frac{1}{\sqrt{\frac{1}{2}\pi}}e^{-2(x-2)^2}\right\}$

So f^* would predict 1 when $0 < \frac{1}{3}\frac{1}{\sqrt{\frac{1}{2}\pi}}e^{-2(x-2)^2} < \frac{2}{3}c\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

Let $\frac{1}{3}\frac{1}{\sqrt{\frac{1}{2}\pi}}e^{-2(x-2)^2} = \frac{2}{3}c\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

get $ce^{-\frac{x^2}{2}} = e^{-2(x-2)^2}$

Take log on both side, get: $2lnc - x^2 = -4(x-2)^2$

$3x^2 - 16x + 16 + 2lnc = 0$

In order to have a root, $\Delta = 256 - 12(16 + 2lnc) > 0$

Thus, $8 - 3lnc > 0$

get $c < e^{\frac{8}{3}} = 14.39$

Therefore, while $0 \leq c \leq 14$, for $\frac{8-\sqrt{16-6lnc}}{3} \leq x \leq \frac{8+\sqrt{16-6lnc}}{3}$,$f^*$ would predict 1.

(b) According to Question (a), when $c > e^{\frac{8}{3}}$, there would be no root for $\frac{1}{3}\frac{1}{\sqrt{\frac{1}{2}\pi}}e^{-2(x-2)^2} = \frac{2}{3}c\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, thus, there would be no 1 in prediction.

Therefore, whatever x value we take, f^* would predict the class as 0.

Problem 3.

Answer:

(a) the eigenvalues of $\Sigma + \sigma^2 I$ are: $\lambda_1 + \sigma^2$, $\lambda_2 + \sigma^2$,..., $\lambda_d + \sigma^2$

(b) the eigenvalues of $(\Sigma + \sigma^2 I)^{-2}$ are $\frac{1}{\lambda^2}$ (λ are: $\lambda_1 + \sigma^2$, $\lambda_2 + \sigma^2$,..., $\lambda_d + \sigma^2$)