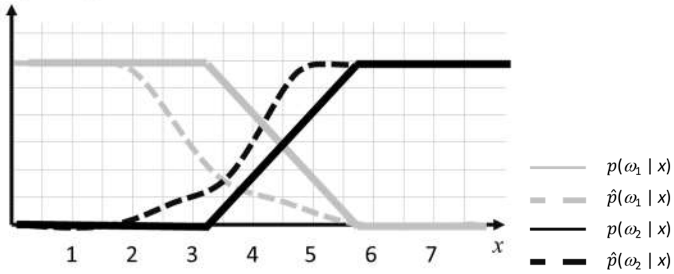
Artificial Intelligence III: Artificial Intelligence and Deep Learning

Ch03 – Supervised Learning

Tutorial Answer

1. Peter estimated the posterior probability of two classes using training samples. The following figure shows the real posterior probability $p(\omega_i \mid x)$ (solid line) and his estimation $\hat{p}(\omega_i \mid x)$ (dashed line), where x means the feature and y_i means class i. Assume that all prior probabilities are the same.



a) According to Bayes rule, which class $x = 4$ should belong to?

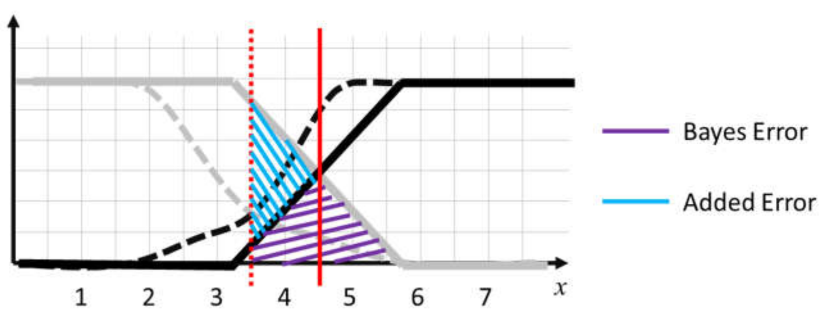
Answer:
Class 2

b) Write down the formulas for the best classifier ($g(x)$) and the estimated classifier ($g'(x)$) according to Peter's estimation, where the formula's output is larger than 0 if x belongs to class 1; otherwise, it is class 2?

Answer:
$g(x) = -x + 4.5$
$g'(x) = -x + 3.5$

c) Indicate $g(x)$ and $g'(x)$, and the regions which represents the Bayes Error and Added Error for $g'(x)$.

Answer:



2. Tom is facing a junk email problem. Averagely, 40% of his emails are junk mail. He wants to develop an automatic system to filter out the junk emails.

a) His friend suggests using the size of an email to decide if an email is junk. Do you think it is a good feature? Please explain why.

Answer:

Size is not a good feature.
This is because the junk email can be large or small.

b) Given a probability density function of the Normal Distribution:
$p(x|\omega_i) = -\frac{1}{2\sigma_i^2}(x-\mu_i)^2 - \frac{1}{2}ln\,2\,\pi - \frac{1}{2}ln\,\sigma_i^2$, where σ_i^2 and μ_i are the variance and mean of the class i.
Tom finds out that the normal and junk emails follow the Normal Distribution and the both variances are 400 KB. The means of the normal email size and the abnormal email size are 120.5 KB and 23.9 KB respectively. Now, someone sends him three emails with 89.4KB, 60.1 KB and 15.0 KB. According to the Bayes Decision Theory, please decide if these emails are junk mail or not.

Answer:

Let ω_1 be the Normal Email:

,

$\hat{\mu}_1$= 120.5 , $\hat{\sigma}_1^2$ = 400

$$p(x|\omega_1)$$
$$= -\frac{1}{800}(x-120.5)^2 - \frac{1}{2}ln\,2\,\pi - \frac{1}{2}ln\,4\,00$$
$$= -\frac{(x-120.5)^2}{800} - 0.9189 - 2.9957$$
$$= -\frac{(x-120.5)^2}{800} - 3.9146$$

Let ω_2 be the Junk Email:

$p(\omega_2)$ = 0.4, $\hat{\mu}_2$= 23.9, $\hat{\sigma}_2^2$ = 400

$$p(x|\omega_2)$$
$$= -\frac{1}{800}(x-23.9)^2 - \frac{1}{2}ln\,2\,\pi - \frac{1}{2}ln\,4\,00$$
$$= -\frac{(x-23.9)^2}{800} - 0.9189 - 2.9957$$
$$= -\frac{(x-23.9)^2}{800} - 3.9146$$

x	x = 89.4	x = 60.1	x = 15.0
$p(x\|\omega_1)$	-5.1237	-8.4749	-17.8275
$p(x\|\omega_2)$	-9.2775	-5.5527	-4.01371

$p(\omega_1) = 0.6$ ln $(p(\omega_1))$ = -0.5108
$p(\omega_2) = 0.4$ ln $(p(\omega_2))$ = -0.9163

x	x = 89.4	x = 60.1	x = 15.0
$p(x\|\omega_1)+p(\omega_1)$	-5.6345	-8.9857	-18.3383
$p(x\|\omega_2)+p(\omega_2)$	-10.1938	-6.469	-4.93
Result	Normal Email	Junk Email	Junk Email

c) If 5% of emails is junk mail, what are the decisions in (b)?

Answer:
$p(\omega_1) = 0.95$ ln $(p(\omega_1))$ = -0.0513
$p(\omega_2) = 0.01$ ln $(p(\omega_2))$ = -2.9957

x	x = 89.4	x = 60.1	x = 15.0
$p(x\|\omega_1)+p(\omega_1)$	-5.6858	-9.0370	-18.3896
$p(x\|\omega_2)+p(\omega_2)$	-13.1895	-9.4647	-7.9257
Result	Normal Email	Normal Email	Junk Email

d) If 90% of emails is junk mail, what are the decisions in (b)?

Answer:
$p(\omega_2) = 0.1$ ln $(p(\omega_2))$ = -2.3026
$p(\omega_1) = 0.9$ ln $(p(\omega_1))$ = -0.1053

x	x = 89.4	x = 60.1	x = 15.0
$p(x\|\omega_1)+p(\omega_1)$	-7.9371	-11.2883	-20.6409
$p(x\|\omega_2)+p(\omega_2)$	-10.2992	-6.5744	-5.0354
Result	Normal Email	Junk Email	Junk Email