به نام خدا

دانشگاه تهران

دانشکده مهندسی برق و کامپیوتر

ماشین لرنینگ

تمرین ۱

نایو بیز و مرز تصمیم گیری

نام و نام خانوادگی	محمد مشرقی
شماره دانشجویی	810199492
تاریخ ارسال گزارش	

سوال ۱

آ

(الف)

$$P(x|w_i) = \frac{1}{\pi b} \frac{1}{1+\left(\frac{x-a_i}{b}\right)^2} \qquad i=1,2 \qquad a_2 > a_1$$

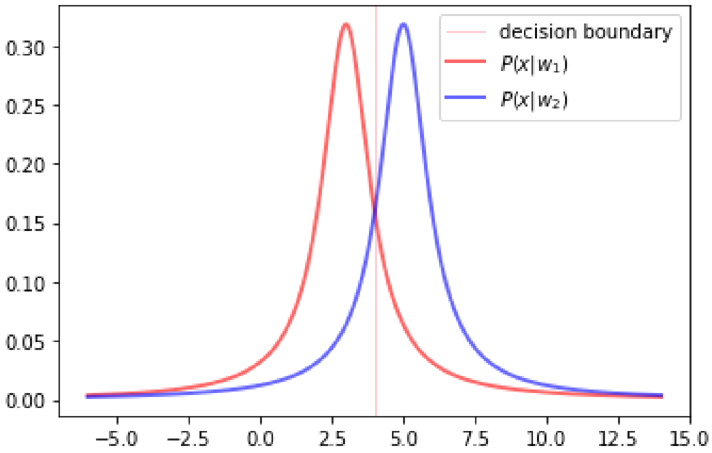
$$\text{if } P(w_1)=P(w_2) \begin{cases} P(w_1|x) = \frac{P(x|w_1)\cancel{P(w_1)}}{\cancel{P(x)}} \\ P(w_2|x) = \frac{P(x|w_2)\cancel{P(w_2)}}{\cancel{P(x)}} \end{cases} \Rightarrow \boxed{\frac{P(w_1|x)}{P(w_2|x)} = \frac{P(x|w_1)}{P(x|w_2)}} \ \star$$

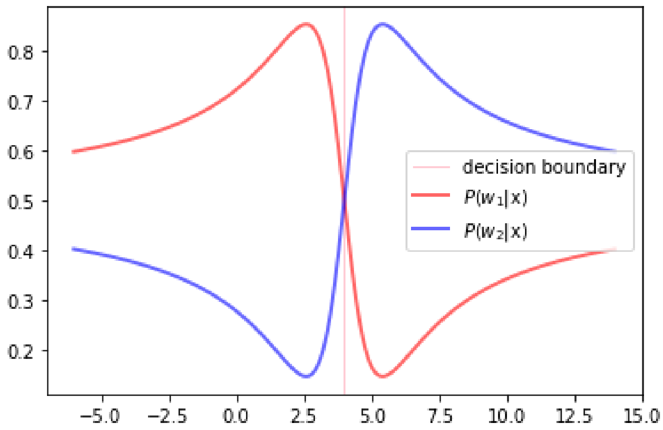
حال اگر فرض کنیم $P(w_1|x) = P(w_2|x)$ باشد داریم:

$$\star \Rightarrow P(x|w_2) = P(x|w_1) =$$

$$\Rightarrow \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2} = \frac{1}{1+\left(\frac{x-a_1}{b}\right)^2} \Rightarrow \left|\frac{x-a_2}{b}\right| = \left|\frac{x-a_1}{b}\right|$$

$$\xrightarrow{b>0} |x-a_2| = |x-a_1| \begin{cases} a_1 = a_2 \ \times \\ x-a_2 = -x+a_1 \Rightarrow \boxed{x = \frac{a_1+a_2}{2}} \end{cases}$$





ب

$$P_{error} = P(x \in R_2|w_1)P(w_1) + P(x \in R_1|w_2)P(w_2)$$

$$= \underbrace{P(w_1)}_{\frac{1}{2}}\int_{R_2} P(x|w_1)dx + \underbrace{P(w_2)}_{\frac{1}{2}}\int_{R_1} P(x|w_2)dx$$

$$= \frac{1}{2}\int_{-\infty}^{\frac{a_1+a_2}{2}} \frac{1}{\pi b_1}\frac{1}{1+\left(\frac{x-a_2}{b}\right)^2} + \frac{1}{2}\int_{\frac{a_1+a_2}{2}}^{+\infty} \frac{1}{\pi b_1}\frac{1}{1+\left(\frac{x-a_1}{b}\right)^2}$$

$$= \frac{1}{2\pi b}\left(\int_{-\infty}^{\frac{a_1+a_2}{2}} \frac{1}{1+\left(\frac{x-a_2}{b}\right)^2} + \int_{\frac{a_1+a_2}{2}}^{+\infty} \frac{1}{1+\left(\frac{x-a_1}{b}\right)^2}\right) = \frac{1}{\pi b}\int_{-\infty}^{\frac{a_1-a_2}{2}} \frac{1}{1+\left(\frac{x-a_1}{b}\right)^2} \qquad y = \frac{x-a_1}{b},\ dy = \frac{dx}{b}$$

1

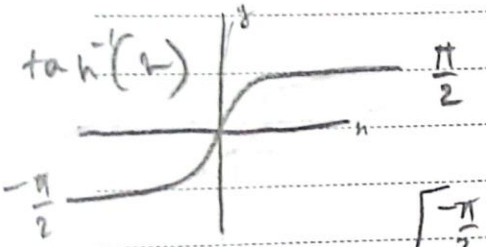
$$= \frac{\cancel{b}}{\cancel{b}\pi}\int_{-\infty}^{\frac{a_1-a_2}{2b}} \frac{dy}{1+y^2} = \left[\frac{\tan^{-1} y}{\pi}\right]_{-\infty}^{\frac{a_1-a_2}{2b}} = \frac{\tan\left|\frac{a_1-a_2}{2b}\right| - \left(-\frac{\pi}{2}\right)}{\pi}$$

$$= \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left|\frac{a_2-a_1}{2b}\right|$$

پ

برای محاسبه $\max(P(error))$ با توجه به قسمت (ب) داریم

$$P(error) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left|\frac{a_1-a_2}{2b}\right|$$



حال برای max گرفتن داریم

با توجه به نمودار $\tan^{-1}(x)$ محدوده $\tan^{-1}(x)$ بین $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

پس محدوده $P(error) \in \left[\underset{x=\pm\infty \text{ برای}}{0}, \underset{x=0 \text{ برای}}{\frac{1}{2}}\right]$

حال برای $\max(P(error))$ باید عبارت داخل $\tan^{-1}\left|\frac{a_1-a_2}{2b}\right|$ به صفر میل کند

داریم

$$\boxed{\frac{a_1-a_2}{2b}=0 \begin{cases} \to a_1=a_2 \\ \text{or} \\ \to b\to\pm\infty \end{cases}}$$

ت) طبق قانون بیز داریم:

$$P(w_1|x) \gtrless P(w_2|x) \Rightarrow \frac{P(x|w_1)P(w_1)}{\cancel{P(x)}} \gtrless \frac{P(x|w_2)P(w_2)}{\cancel{P(x)}} \Rightarrow$$

$$\frac{P(x|w_1)}{P(x|w_2)} \gtrless \frac{P(w_2)}{P(w_1)} \xrightarrow{\text{if } \frac{P(w_2)}{P(w_1)}=1} \frac{P(x|w_1)}{P(x|w_2)} \gtrless 1$$

$$\Rightarrow P(x|w_1) \gtrless P(x|w_2) \Rightarrow \cancel{1}+\left(\frac{x-a_2}{b}\right)^2 \gtrless \cancel{1}+\left(\frac{x-a_1}{b}\right)^2$$

$$\Rightarrow \left|\frac{x-a_2}{\cancel{b}}\right| \geq \left|\frac{x-a_1}{\cancel{b}}\right| \Rightarrow \boxed{x=\frac{a_1+a_2}{2}}$$

میزان خطا طبق قسمت ب بدست آمد

$$P(er) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}\left|\frac{a_1-a_2}{2b}\right|$$

ث

ث)

$$\begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$$

$$\lambda_{21} P(w_1|x) \gtrless \lambda_{12} P(w_2|x) \xRightarrow{\text{if } P(w_1)=P(w_2)} \frac{\lambda_{21}}{\lambda_{12}} = 2 \gtrless \frac{P(x|w_2)}{P(x|w_1)}$$

$$\Rightarrow \frac{1+\left(\frac{x-a_1}{b}\right)^2}{1+\left(\frac{x-a_2}{b}\right)^2} \gtrless 2 \Rightarrow \left(\frac{x-a_1}{b}\right)^2 \gtrless 1+2\left(\frac{x-a_2}{b}\right)^2$$

$$x = 2a_2 - a_1 \pm \sqrt{2a_1^2 + 2a_2^2 - 4a_1a_2 - b^2}$$

$$x \in \left[2a_2 - a_1 - \sqrt{2a_1^2 + 2a_2^2 - 4a_1a_2 - b^2} \,,\, 2a_2 - a_1 + \sqrt{2a_1^2 + 2a_2^2 - 4a_1a_2 - b^2}\right]$$

اگر x در بازه بالایی باشد متعلق به دسته w_1 و اگر در بازه بالایی نباشد برای w_2 است.

(ت ادامه)

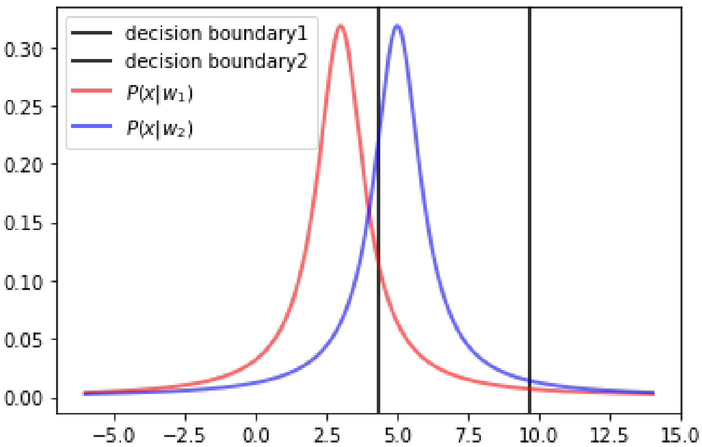
$$P(er) = P(x\in R_2|w_1)P(w_1) + P(x\in R_1|w_2)P(w_2)$$

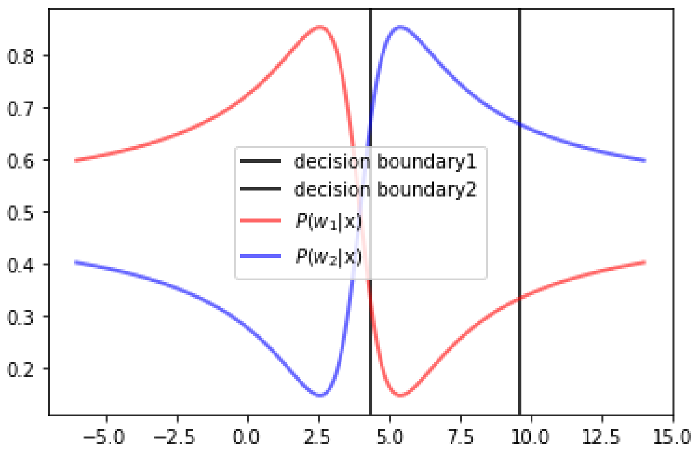
$$= \underbrace{P(w_1)\int_{R_2} P(x|w_1)dx}_{1} + \underbrace{P(w_2)\int_{R_1} P(x|w_2)dx}_{2}$$

$$1: \frac{1}{\pi b}\int_{-\infty}^{2a_2-a-\sqrt{\dots}} \frac{dx}{1+\left(\frac{x-a_1}{b}\right)^2} + \frac{1}{\pi b}\int_{2a_2-a+\sqrt{\dots}}^{+\infty} \frac{dx}{1+\left(\frac{x-a_1}{b}\right)^2}$$

$$2: \frac{1}{\pi b}\int_{2a_2-a-\sqrt{\dots}}^{2a_2-a+\sqrt{\dots}} \frac{dx}{1+\left(\frac{x-a_2}{b}\right)^2} =$$

بعد از محاسبه ① و ② به ترتیب در $P(w_1)$ و $P(w_2)$ ضرب کرده و جمع می‌کنیم و احتمال $P(error)$ بدست می‌آید





2

2)

$$P(x|w_i) = \begin{cases} \frac{x}{\sigma_i^2} \exp\left(-\frac{x^2}{2\sigma_i^2}\right) & x \geq 0 \\ 0 & ow \end{cases}$$

$$\frac{x}{\sigma_1^2} e^{\left(\frac{-x^2}{2\sigma_1^2}\right)} = \frac{x}{\sigma_2^2} e^{\left(\frac{-x^2}{2\sigma_2^2}\right)}$$

باید $f_1(x) = f_2(x)$ پیدا کنیم

$$\Rightarrow x^2 = \frac{4\sigma_1^2 \sigma_2^2}{(\sigma_2^2 - \sigma_1^2)} \ln\left(\frac{\sigma_1^2}{\sigma_2^2}\right)$$

$$\Rightarrow \boxed{x = \text{decision boundary} = \sqrt{\frac{4\sigma_1^2\sigma_2^2}{(\sigma_2^2 - \sigma_1^2)} \ln\left(\frac{\sigma_2}{\sigma_1}\right)}}$$

آ) ۱) ابتدا احتمال پیشین را با توجه به تعداد داده ها بدست می آوریم

$W_1 =$ نقاط مشکی $\qquad P(W_1) = \frac{10}{19}$

$W_2 =$ نقاط قرمز $\qquad P(W_2) = \frac{9}{19}$

$$g_i(x) = X^T W_i X + w_i^T X + w_{i0} \quad \begin{cases} w_1 \\ w_2 \end{cases}$$

$$X = \begin{bmatrix} x \\ y \end{bmatrix} \begin{cases} \to w_1 \Rightarrow \mu_1 = \begin{bmatrix} -0.15 \\ -0.15 \end{bmatrix} \\ \to w_2 \Rightarrow \mu_2 = \begin{bmatrix} 1.33 \\ 1.61 \end{bmatrix} \end{cases}$$

مقادیر [illegible] پایتون

از [illegible]

$$W_i = -0.5 \Sigma_i^{-1} \qquad W_1 = \begin{bmatrix} -0.322 & 0.0016 \\ +0.0016 & -0.995 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} -0.97087 & 0.18347 \\ 0.18347 & -0.54487 \end{bmatrix}$$

$$w_i = \Sigma_i^{-1} \mu_i \qquad w_1 = \begin{bmatrix} -0.3463 \\ 0.2185 \end{bmatrix} \qquad w_2 = \begin{bmatrix} 1.9975 \\ 1.265 \end{bmatrix}$$

$$w_{i0} = -0.5 \times \mu_i^T \Sigma_i^{-1} \mu_i \qquad w_{10} = -1.02 \qquad w_{20} = -3.063$$

$$g_i(x) = X^T W_i X + w_i^T X + w_{i0}$$

$$g_1(x) = -0.322x^2 + 0.0032xy - 0.995y^2 - 0.346x + 0.2185y - 1.02$$

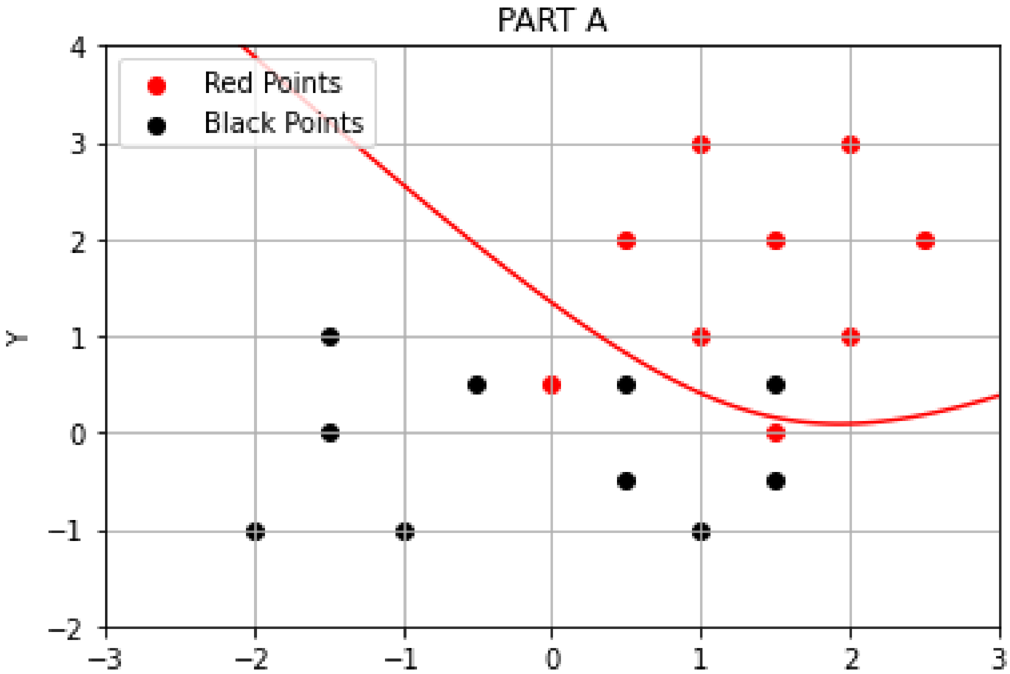
$$g_2(x) = -0.97x^2 + 0.3769xy - 0.544y^2 + 1.99x + 1.265y - 3.063$$

$$\frac{P(w_2)}{P(w_1)} = \frac{\frac{9}{19}}{\frac{10}{19}} = \frac{9}{10} = \frac{g_1(x)}{g_2(x)}$$

مرز تصمیم گیری

$$9g_2(x) - 10g_1(x) = 0$$

$$+5.51x^2 - 3.36xy - 21x - 5y^2 - 9.2y + 21.45 = 0$$



ب

یکشنبه

ب)

$$M_1 = \frac{1}{10}\begin{bmatrix} \sum_{i=1}^{9} x_i \\ \sum_{i=1}^{9} y_i \end{bmatrix} = \frac{-15}{10} = \begin{bmatrix} -1.5 \\ -1.5 \end{bmatrix} \qquad M_2 = \frac{1}{9}\begin{bmatrix} \sum_{i=1}^{8} x_i \\ \sum_{i=1}^{8} y_i \end{bmatrix} = \begin{bmatrix} \frac{+12}{9} \\ \frac{14.5}{9} \end{bmatrix}$$

$$\Sigma_1 = \begin{bmatrix} var\,x & cov\,xy \\ cov\,yx & var\,y \end{bmatrix} \Rightarrow \begin{cases} var\,x = \frac{1}{10}\sum_{i=1}^{10}(x - M_{1x})^2 = \\ var\,y = \frac{1}{10}\sum_{i=1}^{10}(y - M_{1y})^2 = \\ cov\,xy = cov\,yx = \frac{1}{10}\sum_{i=1}^{10}\frac{(y-M_y)(x-M_x)}{N} = \end{cases}$$

$$\Sigma_2 = \begin{bmatrix} \; & \; \\ \; & \; \end{bmatrix} \Rightarrow \begin{cases} var\,x = \frac{1}{9}\sum_{i=1}^{9}(x - M_{2x})^2 = 0.55 \\ var\,y = \frac{1}{9}\sum_{i=1}^{9}(y - M_{2y})^2 = 0.98 \\ cov\,xy = cov\,yx = \frac{1}{9}\sum_{i=1}^{9}\frac{(y-M_y)(x-M_x)}{9} = 0.1852 \end{cases}$$

$$\Sigma_1 = \begin{bmatrix} 1.5525 & 0.0025 \\ 0.0025 & 0.5025 \end{bmatrix} \qquad \Sigma_2 = \begin{bmatrix} 0.55 & 0.1852 \\ 0.1852 & 0.98 \end{bmatrix}$$

ب)

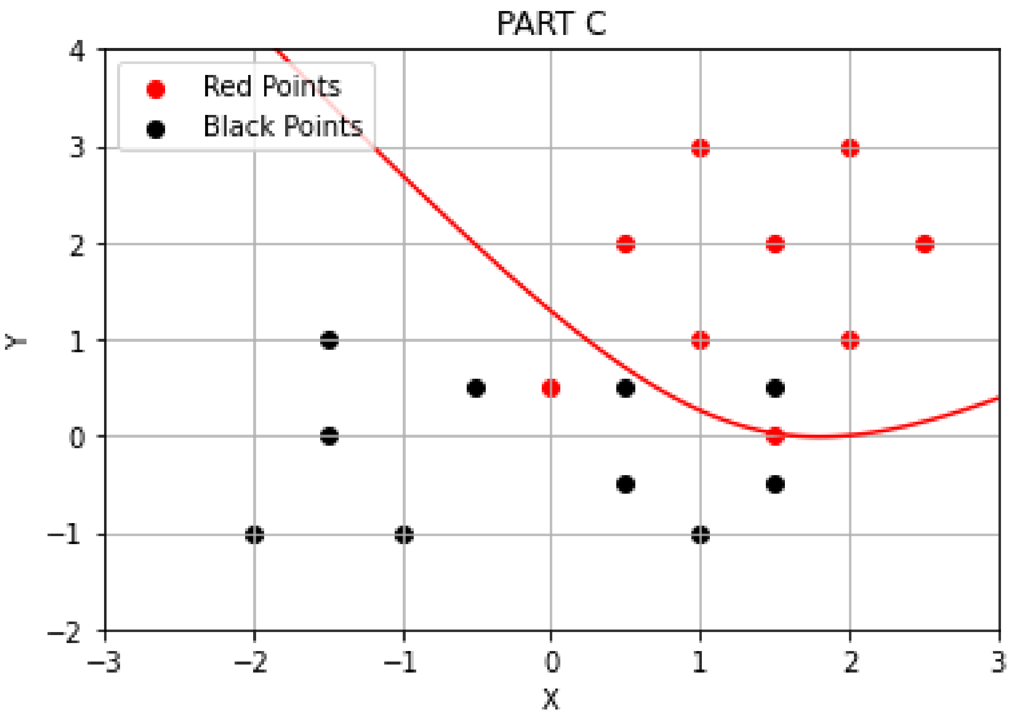
$$P(w_1) = P(w_2) = 0.5$$

پ

$$\frac{g_1(x)}{g_2(x)} = 1 \Rightarrow g_2(x) - g_1(x) = 0$$

$$0.6488x^2 - 0.373xy - 2.343x - 0.45y^2 - 1.04y + 2.043 = 0$$

$$\text{accuracy} = \frac{16}{19} \xrightarrow{\text{نرخ خطا}} \boxed{\frac{3}{19}}$$



ت

$$R(\alpha_i|x)=\sum_{j=1}^{c}\lambda(\alpha_i|w_j)P(w_j|x)$$

$$\left.\begin{aligned}R(\alpha_1|x)&=\lambda_{12}P(w_2|x)\\R(\alpha_2|x)&=\lambda_{21}P(w_1|x)\end{aligned}\right\}\xrightarrow{\text{Decision boundary}}$$

$$\lambda P(w_2|x)=2\lambda P(w_1|x)$$

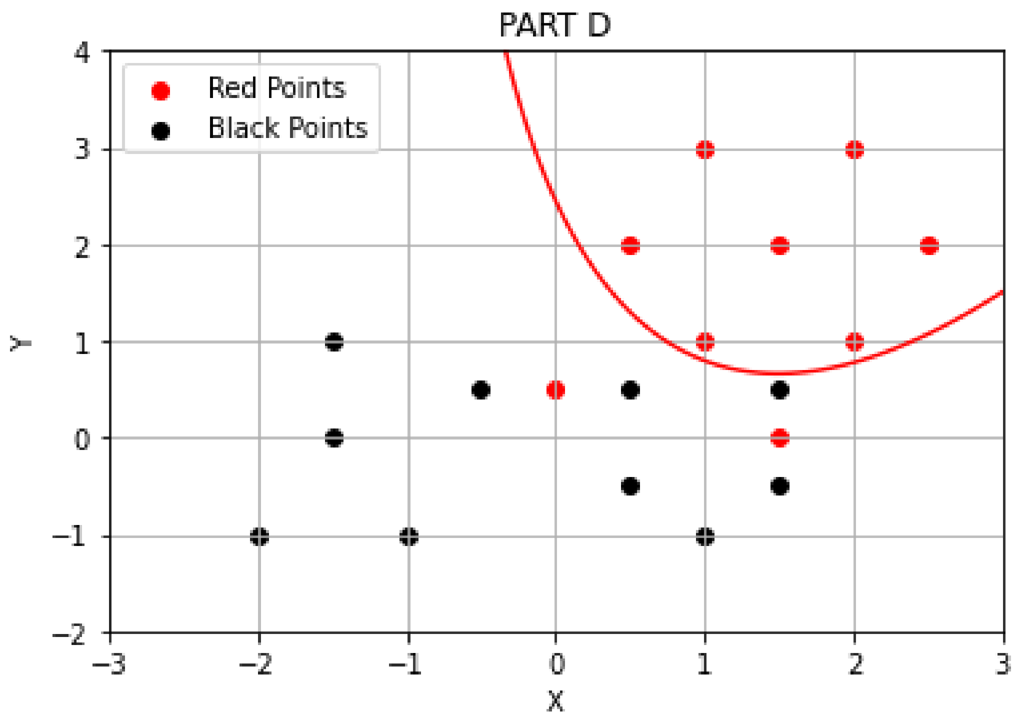
$$\frac{P(x|w_2)}{P(x|w_1)}=\frac{\lambda_{21}}{\lambda_{12}}\frac{P(w_1)}{P(w_2)}\xrightarrow[P(w_1)=P(w_2)]{}\frac{g_2(x)}{g_1(x)}=0.5$$

$$g_1(x)-2g_2(x)=0$$

$$=1.619x^2-0.75xy-4.34x+0.0947y^2$$

$$-2.3115y+5.106$$

$$\text{accuracy}=\frac{17}{19}\Rightarrow\boxed{\text{خطا}=\frac{2}{19}}$$



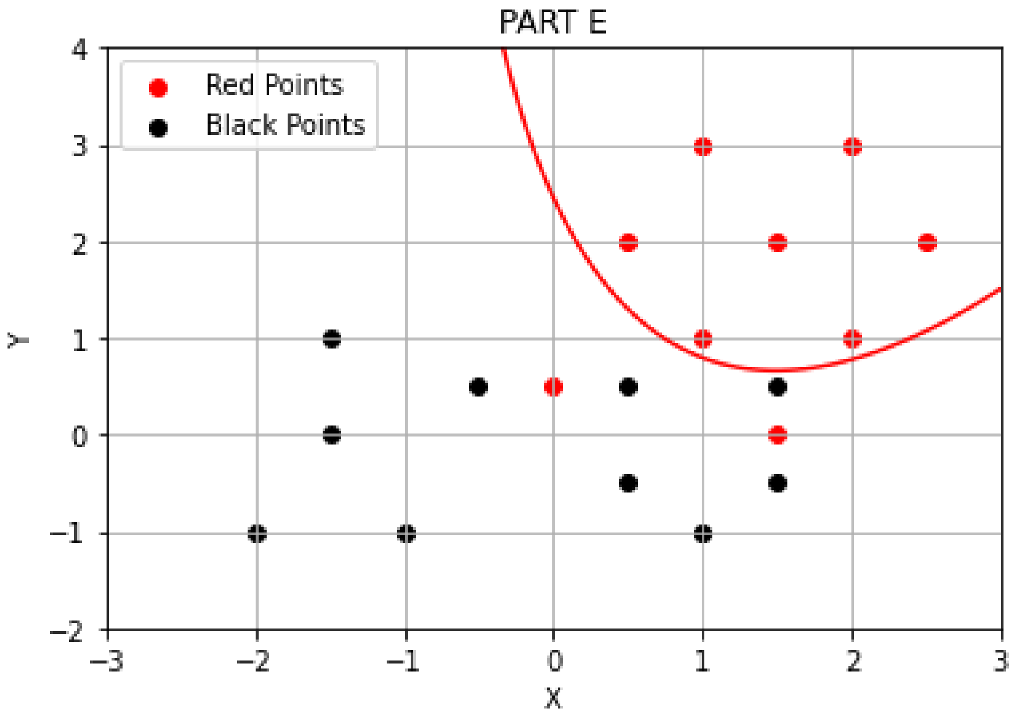
$$\frac{g_2(x)}{g_1(x)} = \frac{P(w_1)}{P(w_2)} = \frac{\frac{1}{3}}{\frac{2}{3}} = 0.5$$

ث

عبارت ساده شده قسمت ت به صورت زیر

$$\boxed{1.62x^2 - 0.75xy - 4.34x + 0.1y^2 - 2.3y + 5.1 = 0}$$

$$\text{accurency} = \frac{17}{19} \rightarrow \boxed{\text{خطا} = \frac{2}{19}}$$



4

آ

4) $P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$

$$L(\lambda; D) = \sum_{i=1}^{n} \log\left(P(X = x_i; \lambda)\right) = \sum_{i=1}^{n} \left[-\lambda + x_i \log(\lambda) - \log(x_i!)\right]$$

$$= -n\lambda + \log(\lambda) \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \log(x_i!)$$

حال برای تخمین گرفتن باید مشتق گرفته و برابر صفر قرار دهیم

$$\frac{dL(\lambda; D)}{d\lambda} = 0 \Rightarrow \left[-n + \frac{1}{\lambda} \sum_{i=1}^{n} x_i - 0\right] = 0$$

$$\Rightarrow \hat{\lambda}_{MLE} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

ب پ ت

ب) 4

$$P(\lambda) = Gamma(\lambda|a,\beta) = C\lambda^{a-1} e^{-\beta\lambda}$$

اصل بیسین

$$P(\lambda|D) = \frac{P(D|\lambda)\,P(\lambda)}{P(D)} = Gamma(\lambda|a,\beta)\prod_{i=1}^{n}\frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$P(D) = \int P(D|\lambda)P(\lambda)\,d\lambda$$

$$= \underbrace{C\lambda^{a-1} e^{-\beta\lambda} \times \frac{\lambda^{\sum_{i=1}^{n} x_i}\, e^{-\lambda n}}{\prod_{i=1}^{n} x_i!}}_{C'} = C'\lambda^{(a-1+\sum_{i=1}^{n} x_i)}\, e^{(-\beta-n)\lambda}$$

میانگین

$$= C'\lambda^{(n\bar{x}+a-1)}\, e^{-\lambda(\beta+n)} = G(\lambda \mid a+n\bar{x},\ \beta+n) \quad *$$

پ) بله چون هر دو از توزیع Gamma هستند و فقط پارامترهای آن ها فرق دارد

ت) $$\hat{\lambda}_{MAP} = \arg\max P(D|\lambda)\,P(\lambda) \xrightarrow[*]{\text{lint } \lambda = \frac{a-1}{\beta}} \begin{cases} \alpha = a + n\bar{x} \\ \beta = \beta + n \end{cases}$$

$$\Rightarrow \boxed{\hat{\lambda}_{MAP} = \frac{a+n\bar{x}-1}{\beta+n}}$$

ث

$$\lim_{n\to\infty} \widehat{lamda}_{MAP} \xrightarrow{a\ \&\ B \text{ بی تاثیر میشن}} \frac{n\overline{x_i}}{n} = \overline{x_i} = \sum_{i=1}^{n} x_i = lamda_{MLE}$$

هنگامی که n به سمت بی نهایت می رود مقدار لاندا به سمت مقدار مشخصی می رود که آن نقطه برابر مقدار لاندا mle هستش.

ج

بستگی به دیتابیس دارد اگه دیتا بیس بزرگ باشد پردازش اطلاعات زمان بر و سخت است برا همین تخمین گر MLE را استفاده می کنیم چون وقتی دیتابیس کوچک باشد محاسبه احتمال پیشین آسان و سریع است برا همین از تخمین گر MAP استفاده می کنیم.

5

آ

طبقه بندی بیزی با احتمال شرطی کار می کند در حالی که نایوبیز یک نوع خاص از طبقه بندی بیزی است که فرض می کند که ویژگی ها مستقل از هم هستند و باعث می شود محاسبات نسبت به طبقه بندی بیزی بسیار راحت تر و کارامد تر شود.

همچنین طبقه بندی بیزی با توجه به توضیحات برا اینکه درست کار کند نیاز به یک دیتابیس بزرگ دارد چون شرطی کار می کند و نیاز به تعداد داده زیادی دارد و هم اینکه وقت گیر است

اما در نهایت دقیق تر است.

اما نایوبیز به دلیل مستقل بودن تعداد داده کمتری نیاز داده و به دلیل مستقل بودن محاسبات ساده تر است.

پیش پردازش :

- در این داده ها تعدادی داده نامعتبر بودش که ابتدا باید حذف می شدن که پایتون اینکار رو انجام دادیم.
- و بعد از آن داده ها را بر می زنیم(shuffle) که سعی شود داده های مختلف از هم داشته باشیم.
- حال که داده ها را داریم آن ها را دسته بندی می کنی برای train و test
- کار دیگه ای که می توان انجام داد تا محاسبات سریعتر شود و خطا کمتر شود میتوان مقیاس داده ها را بین 1- و 1 تعیین کرد .

ب

روش پیاده سازی دستی نایوبیز

```
[[41  0  0]
 [ 1 12  0]
 [ 0  0 31]]
accuracy_score =  0.9882352941176471
precision_score =  0.9920634920634921
recall_score =  0.9743589743589745
```

پ

روش پیاده سازی دستی نایوبیز

```
[[41  0  0]
 [ 1 12  0]
 [ 0  0 31]]
accuracy_score =  0.9882352941176471
precision_score =  0.9920634920634921
recall_score =  0.9743589743589745
```

نتایج دقیقا یکی شد می توان نتیجه گرفت که به جای اینکه این همه وقت بزاریم و از پایه چیزی بنویسیم می تونیم از کتابخونه آن استفاده کنیم و با پارمتر دهی یکسان به نتیجه یکسان برسیم.

۶

در این قسمت کاری که ما انجام دادیم این بودش که هر وقت عکسی را بررسی می کردیم یک ماتریس سه بعدی داشتیم که دو بعد تعداد پیکسل و بعد سوم رنگ آن به صورت RGB هستش برای تشخیص بین درخت و دریا میانگین رنگ های ابی و سبز کل پیکسل های یک عکس را می گیریم.

اگه سبز بیشتر بود جنگل و ابی بیشتر بود دریا می شود.

نتیجه:

فقط سه عکس اشتباه تشخیص داده شدند به نام های

```
['j45', 'j44']
['s24']
```

عکس پایین عکس دریاست اما به دلیل جلبک های سبز رنگ اشتباه جداسازی شده.

در در این عکس ها که عکس های جنگل هستند به دلیل وجود اسمان و برف و نبود رنگ سبز اشتباه تشخیص داده شده و دریا در نظر گرفته شده اند

با توجه به بررسی ها اشتباهات منطقی هستند.

```
[[40  2]
 [ 1 39]]
accuracy_score =  0.9634146341463414
precision_score =  0.9512195121951219
recall_score =  0.975
```