

آزمایشگاه مدل‌سازی و پردازش تصاویر پزشکی

بسمه تعالی

بازشناسی آماری الگو



دانشگاه صنعتی خواجه نصیرالدین طوسی

۱۴۰۴/۸/۱۰

تاریخ تحویل:

تمرین سری سوم

۱- (اختیاری) در یک مساله دسته‌بندی دو کلاسه ω_1 و ω_2 دارای توزیع یکنواخت بترتیب در فواصل $[0, 2]$ و $[1, 4]$ و احتمالات $P(\omega_1)=P(\omega_2)=0.5$ هستند.

الف- مرز دسته‌بندی کننده بیز را برای رسیدن به حداقل خطا بدست آورید و خطای بیز را محاسبه کنید.

ب- حد بالای چرنوف و مقدار بهینه S متناظر آن را بیابید.

ج- حد بالای باتاچاریا (حد بالای چرنوف بازای $s=0.5$) را بدست آورید.

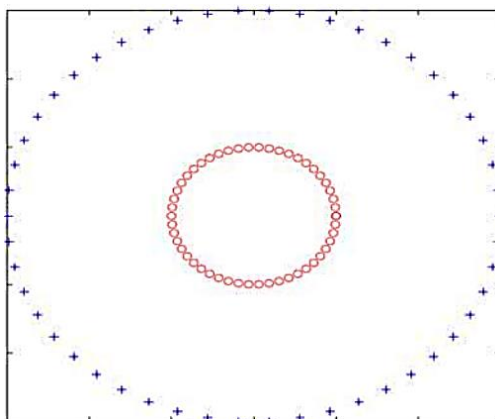
۲- (اختیاری) داده‌های زیر را از دو کلاس قرمز و آبی در نظر بگیرید. احتمالات شرطی از توزیع گاوسی پیروی می‌کنند. کواریانس کلاس قرمز I و کواریانس کلاس آبی kI است ($k > 1$).

الف- مرز دسته‌بندی کننده بیز در حالت کلی به چه شکلی است اگر احتمال اولیه کلاس آبی را P_b و احتمال اولیه کلاس قرمز را P_r در نظر بگیریم؟

ب- در صورتی که احتمال اولیه کلاس‌ها یکی باشد این مرز حدوداً در کجا قرار می‌گیرد؟

ج- در چه صورت دسته‌بندی کننده بیز به کل فضا برچسب کلاس آبی را اختصاص می‌دهد؟

د- فرض کنید احتمالات اولیه به شکلی نسبت داده شده‌اند که مرز به کلاس آبی نزدیک شده است. چگونه می‌توان با در نظر گرفتن هزینه مرز را از این کلاس دور نمود؟



۳- (اختیاری) سه کلاس با توزیع‌های زیر و احتمالات اولیه یکسان را در نظر بگیرید:

$$p(\mathbf{x}|\omega_1) \sim N(\mathbf{0}, I)$$

$$p(\mathbf{x}|\omega_2) \sim N([1, 1]^t, I)$$

$$p(\mathbf{x}|\omega_3) \sim \frac{1}{2} N([0.5, 0.5]^t, I) + \frac{1}{2} N([-0.5, 0.5]^t, I)$$

الف. نقطه $\mathbf{x} = [x_1, x_2] = [0.3, 0.3]^t$ بر اساس دسته‌بندی با کمترین احتمال خطا بر چسب کدام کلاس را می‌گیرد؟

ب- خطای شرطی دسته‌بندی بیز بازای نمونه فوق چقدر است؟

- ج. فرض کنید ویژگی اول را نداشته باشیم. در این صورت $x = [*, 0.3]^t$ به کدام دسته تعلق می‌یابد؟
 (راهنمایی: $p(\omega_i|x_2) = \frac{\int p(x_1, x_2|\omega_i)p(\omega_i)dx_1}{p(x_2)}$ و نیازی به محاسبه $p(x_2)$ نیست چون در مقایسه تأثیری ندارد).
 د. حال اگر ویژگی دوم را نداشته باشیم، بردار $x = [0.3, *]^t$ به کدام کلاس نسبت داده می‌شود؟

۴- (اختیاری) دو توزیع نرمال با پارامترهای زیر در اختیار است:

$$m_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad m_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \Sigma_1 = \Sigma_2 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

- الف- مرز تصمیم‌گیری بیز را برای حداقل کردن احتمال خطای دسته‌بندی رسم کنید.
 ب- مرز تصمیم‌گیری بیز را برای حداقل کردن هزینه دسته‌بندی بازای $\lambda_{11}=\lambda_{22}=0$ و $\lambda_{12}=2\lambda_{21}$ رسم کنید.
 ج- هزینه بیز را محاسبه کنید.

۵- (اختیاری) سه توزیع نرمال با پارامترهای زیر داده شده است:

$$m_1 = m_2 = m_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \Sigma_1 = \begin{bmatrix} \sigma_n^2 & 0 \\ 0 & \sigma_n^2 \end{bmatrix}, \quad \Sigma_2 = \begin{bmatrix} \sigma_n^2 + \sigma_s^2 & 0 \\ 0 & \sigma_n^2 \end{bmatrix}, \quad \Sigma_3 = \begin{bmatrix} \sigma_n^2 & 0 \\ 0 & \sigma_n^2 + \sigma_s^2 \end{bmatrix}$$

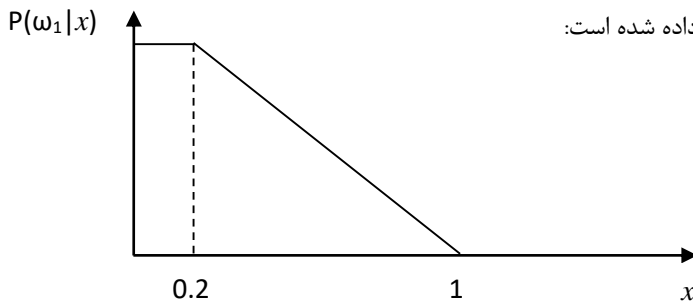
ماتریس هزینه دسته‌بندی بصورت زیر است:

$$L = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & a \\ 1 & a & 0 \end{bmatrix}$$

و داریم: $0 \leq a < 1$ و $P(\omega_1)=P(\omega_2)=p$

- الف- مرز دسته‌بندی بیز را بیابید و رسم کنید.
 ب- رابطه محاسبه احتمال خطا را بنویسید (نیازی به محاسبه انتگرال‌ها نمی‌باشد).

۶- (اختیاری) (سوال امتحان میان ترم بازشناسی آماری الگو سال ۱۳۸۷)



در یک مساله دوکلاسه احتمال پسین کلاس ω_1 ، مطابق نمودار شکل زیر داده شده است:

الف- مطلوب است رسم منحنی احتمال پسین کلاس ω_2 .

- ب- محدوده‌های تصمیم‌گیری R_1 و R_2 را برای یک دسته‌بندی کننده حداقل (احتمال) خطا بدست آورید.
 ج- چنانچه آستانه امتناع (Reject)، $t=0.25$ انتخاب شده باشد، محدوده ناحیه امتناع (R_0) را مشخص نمایید.

د- فرض کنید بخواهیم یک دسته‌بندی کننده حداقل هزینه با امکان امتناع طراحی نماییم. با فرض اینکه هزینه دسته‌بندی غلط برابر ۱ و هزینه دسته‌بندی صحیح برابر صفر و هزینه امتناع λ باشد معیار تصمیم‌گیری برای رسیدن به حداقل هزینه دسته‌بندی را بدست آورید.

ه- نشان دهید که در صورت انتخاب مناسب λ روش‌های بیان شده در بندهای ج و د معادل هستند.

۷. (اختیاری) فرض کنید $p(x|\omega_1)$ و $p(x|\omega_2)$ به صورت زیر تعریف شده باشند:

$$p(x|\omega_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \forall x$$

$$p(x|\omega_2) = \frac{1}{4}, \quad -2 < x < 2$$

الف. مرز کلاس‌بندی با حداقل خطا را با فرض $P(\omega_1) = P(\omega_2)$ بیابید.

ب. مقدار α را طوری بیابید که اگر $P(\omega_1) > \alpha$ نتیجه کلاس‌بندی با حداقل خطا، بدون در نظر گرفتن مقدار x ، همواره ω_1 باشد.

ج. آیا می‌توان مقدار β را به دست آورد به طوری که اگر $P(\omega_2) > \beta$ نتیجه کلاس‌بندی با حداقل خطا، بدون در نظر گرفتن مقدار x ، همواره ω_2 باشد؟ چرا؟

۸. (مساله امتحانی میان ترم ۱۳۹۶)

یک مساله دسته‌بندی ۲ کلاسه که در آن متغیر تصادفی اسکالر x فقط دو مقدار ۱ و ۲ را اختیار می‌کند در نظر بگیرید. چگالی احتمال شرطی کلاسها بازای مقادیر مختلف x در جدول زیر داده شده است:

x	$p(x/\omega_1)$	$p(x/\omega_2)$
1	$\frac{1}{3}$	$\frac{2}{3}$
2	$\frac{2}{3}$	$\frac{1}{3}$

جدول زیر نیز مقادیر عناصر ماتریس تلفات (Loss matrix) را نشان می‌دهد:

$$L = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$$

الف- چنانچه نواحی $R_i, i=1,2$ مشخص‌کننده محدوده کلاسهای دوگانه فوق در فضای ویژگی باشند، هزینه دسته‌بندی متناظر با کلاس k -ام از رابطه زیر بدست می‌آید ($M=2$).

$$r_k = \sum_{i=1}^M \lambda_{ki} \int_{R_i} p(x|\omega_k) dx$$

که در آن λ_{ki} معرف هزینه دسته‌بندی نمونه‌ای از کلاس k -ام در کلاس i -ام است. با در نظر گرفتن حالت‌های مختلف تقسیم‌بندی فضای ویژگی به نواحی دوگانه R_1 و R_2 ، هزینه دسته‌بندی متناظر با هر یک از کلاسهای دوگانه را محاسبه نمایید.

ب- با فرض $P(\omega_1) = \frac{1}{4}$ و $P(\omega_2) = \frac{3}{4}$ هزینه متوسط مربوط به هر قاعده تصمیم‌گیری را مطابق رابطه زیر محاسبه کنید.

$$r = \sum_{k=1}^M r_k P(\omega_k)$$

ج- قاعده دسته‌بندی بهینه بیز (متناظر با هزینه حداقل) کدام است؟

۹. (اختیاری) سوال امتحانی میان ترم ۱۳۹۷

یک مساله دسته‌بندی دو کلاسه با ویژگی تک‌بعدی و باینری x ($x=0, x=1$) را در نظر بگیرید و فرض کنید که داشته باشیم:

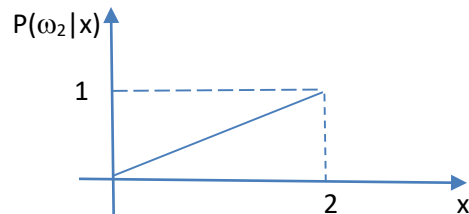
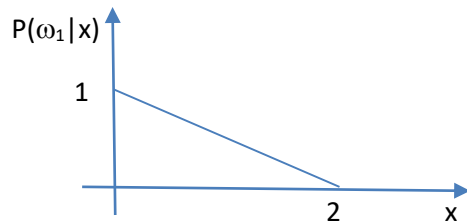
$$P(\omega_1)=0.40$$

الف- چنانچه توزیع نمونه‌ها مطابق جدول زیر باشد، با استفاده از قاعده بیز احتمالات پسین کلاسها $P(\omega_1|x)$ و $P(\omega_2|x)$ را برای $x=0$ و $x=1$ محاسبه نمایید:

	ω_1	ω_2
$x=0$	141	258
$x=1$	259	342

ب- با حفظ احتمال کلاسها توزیع نمونه‌ها را به گونه‌ای تغییر دهید که ابهام در تصمیم‌گیری برای وقوع $x=0$ بوجود آید.

۱۰. (اختیاری) (مساله امتحانی میان ترم ۱۳۹۰) در فضای یک بعدی فرض کنید احتمالات شرطی کلاس‌ها بصورت زیر تعریف شوند:



که در آن کلاس ω_1 معرف آشکارسازی خرابی و کلاس ω_2 معرف عدم آشکارسازی خرابی یک دستگاه CNC می‌باشد. بر اساس آشکارسازی و یا عدم آشکارسازی خرابی سه عملیات زیر امکان پذیر است:

الف- ادامه کار دستگاه با ظرفیت کامل

ب- توقف کامل دستگاه

ج- ادامه کار دستگاه با ظرفیت محدود

جدول زیر هزینه مربوط به هریک از تصمیمات سه گانه فوق را بر اساس آشکارسازی و یا عدم آشکارسازی خرابی دستگاه نشان می‌دهد که در آن ϵ معرف احتمال خطای تشخیص خرابی می‌باشد.

تشخیص \ تصمیم	عدم وجود خرابی	وجود خرابی
ادامه کار با ظرفیت کامل	0	2ϵ
توقف کامل	4ϵ	0
ادامه کار با ظرفیت محدود	2ϵ	ϵ

الف- احتمال خطای تشخیص خرابی (ϵ) را محاسبه نمایید.

ب- قاعده تصمیم‌گیری را برای یک کلاسیفایر حداقل هزینه بیان کنید و محدوده مقادیر x را برای اخذ هر یک از تصمیمات مشخص نمایید.

۱۱. (اختیاری) سوال امتحانی میان ترم ۱۳۹۸

یک مساله دسته‌بندی دو کلاسه با ویژگی دوبعدی (X_1, X_2) که در آن X_1 و X_2 مقادیر $(-1, 0, 1)$ را اختیار می‌کنند در نظر بگیرید و فرض کنید که داشته باشیم: $P(C_1) = P(C_2) = 0.50$. چنانچه چگالی شرطی توزیع نمونه‌ها بر حسب هریک از دو ویژگی مطابق جداول زیر باشد:

$X_1 =$ \ Class	C_1	C_2
-1	0.2	0.3
0	0.4	0.6
1	0.4	0.1

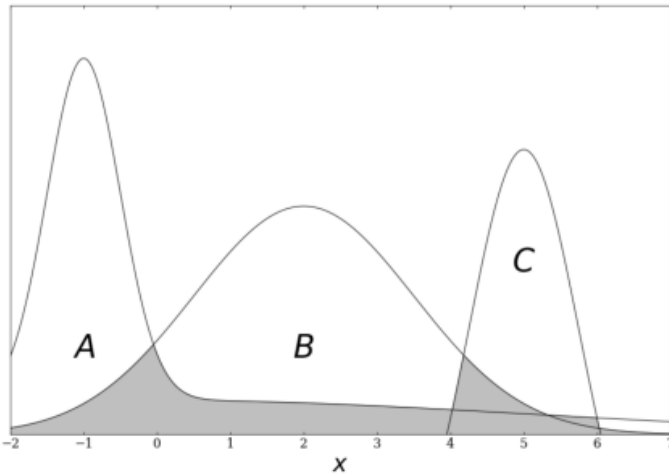
$X_2 =$ \ Class	C_1	C_2
-1	0.4	0.1
0	0.5	0.3
1	0.1	0.6

الف- مطلوب است دسته‌بندی نمونه‌ی $(-1, -1)$ با استفاده از قاعده بیز. به عبارت دیگر مطلوب است محاسبه احتمالات پسین زیر:

$$P(C_1 | X_1 = -1, X_2 = -1) = ?$$

(راهنمایی: در صورت نیاز متغیرهای X_1 و X_2 را مستقل فرض کنید.)

ب- خطای شرطی تصمیم‌گیری در این حالت چقدر است؟



۱۲. (اختیاری) نمودار زیر توزیع شرطی $p(x|y)$ سه کلاس A ، B و C را نشان می‌دهد. احتمال اولیه هر سه کلاس بیش از صفر است و نمودار تاثیر احتمال اولیه را نشان نمی‌دهد. توضیح دهید که:

الف) آیا خطای Bayes، ناحیه هاشورخورده است؟

ب) آیا امکان دارد در شرایطی دسته‌بندی کننده Bayes

برچسب کلاس B را به تمام نواحی x نسبت دهد؟

ج) آیا امکان دارد در شرایطی دسته‌بندی کننده Bayes

برچسب کلاس C را به تمام نواحی x نسبت دهد؟

د) آیا امکان دارد خطای Bayes صفر شود؟

۱۳. (اختیاری) رباتی را در نظر بگیرید که با استفاده از یک دسته‌بندی کننده Bayes سگ‌ها و گربه‌ها را از هم تمیز می‌کند. این ربات با

مثال‌های نویزدار زیر آموزش می‌بیند. فرض کنید ویژگی‌ها به شرط داشتن کلاس از هم مستقل هستند $(p(F_1, F_2 | C) =$

$$p(F_1 | C)p(F_2 | C))$$

Example	Sound	Fur	Color	Class
Example #1	Meow	Coarse	Brown	Dog
Example #2	Bark	Fine	Brown	Dog
Example #3	Bark	Coarse	Black	Dog
Example #4	Bark	Coarse	Black	Dog
Example #5	Meow	Fine	Brown	Cat
Example #6	Meow	Coarse	Black	Cat
Example #7	Bark	Fine	Black	Cat
Example #8	Meow	Fine	Brown	Cat

الف) احتمالات زیر را محاسبه نمایید:

$$P(\text{Sound}=\text{Bark} \mid \text{Class}=\text{Dog})$$

$$P(\text{Fur}=\text{Coarse} \mid \text{Class}=\text{Dog})$$

$$P(\text{Color}=\text{Brown} \mid \text{Class}=\text{Dog})$$

ب) نمونه جدید زیر از راه می‌رسد. ربات چه برچسبی به آن می‌زند؟

$$(\text{Sound}=\text{Bark} \wedge \text{Fur}=\text{Coarse} \wedge \text{Color}=\text{Brown})$$

۱۴. (اختیاری) در یک مساله دسته‌بندی دو کلاسه، بردار ویژگی هر کلاس به صورت نرمال با ماتریس کوواریانس Σ و بردارهای میانگین μ_1 و μ_2 تعریف شده است. نشان دهید که برای دسته بندی کمینه خطای بیز احتمال خطا برابر رابطه زیر است که در آن d_m فاصله ماکسیمیسی بین μ_1 و μ_2 است.

$$P_B = \int_{0.5d_m}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

۱۵. یک مساله دسته بندی دو کلاسه در فضای یک بعدی را در نظر بگیرید. فرض کنید توابع چگالی شرطی کلاس‌ها بصورت زیر داده شده باشد:

$$p(x|\omega_1) = N(2,1)$$

$$p(x|\omega_2) = \begin{cases} \frac{1}{3} & , 0 < x < 3 \\ 0 & otherwise \end{cases}$$

الف- دسته‌بند پیشینه شباهت را طراحی و نواحی کلاس‌ها را مشخص نمایید.

ب- چنانچه احتمال کلاس ۱ دو برابر احتمال کلاس ۲ باشد، دسته‌بند پیشینه احتمال پسین را طراحی و نواحی کلاس‌ها را مشخص نمایید.

ج- با فرض برابر بودن احتمال کلاس‌ها و ماتریس هزینه‌های زیر، دسته‌بند حداقل هزینه بیز را طراحی و نواحی کلاس‌ها را مشخص نمایید.

$$C = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

۱۶. (میان ترم ۱۴۰۲) یک مساله دسته بندی سه کلاسه در فضای ۴ بعدی در نظر بگیرید. چنانچه احتمال کلاس‌ها برابر، توابع چگالی احتمال کلاس‌ها نرمال با ماتریس کواریانس برابر بصورت زیر:

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 20 & 0 \\ 0 & 0 & 0 & 30 \end{bmatrix}$$

و میانگین کلاسها بصورت زیر باشد:

$$\mu_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 3 \end{bmatrix}, \quad \mu_3 = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix}$$

مطلوب است دسته بندی نمونه $x = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$.

۱۷. (میان ترم ۱۴۰۳) جدول زیر دآوری ۵ فیلم را در دو دسته + و - بر اساس دارا بودن هر یک سه ویژگی great, fine و terrible نشان می دهد. مثلاً فیلمی که دارای ویژگی great=1 ، fine=0 و terrible=0 بوده در دسته + دسته بندی شده است.

great	fine	terrible	class
1	0	0	+
0	1	1	--
0	1	1	--
0	0	0	+
1	0	1	--

چنانچه یک فیلم فقط دارای ویژگی terrible باشد، با استفاده از دسته‌بند Naïve Bayes در کدام دسته (+ یا -) قرار می‌گیرد؟ با محاسبه نشان دهید یا برای پاسخ خود دلیل ذکر نمایید.

راهنمایی: دسته‌بند Naïve Bayes حالت خاصی از دسته‌بند Bayes است که در آن فرض می‌شود مولفه‌های ویژگی در (چگالی) احتمال شرطی کلاسها دارای خاصیت استقلال هستند. یعنی:

$$p(\mathbf{x}|\omega_i) = p(x_1, x_2, \dots, x_d|\omega_i) \cong p(x_1|\omega_i)p(x_2|\omega_i) \dots p(x_d|\omega_i)$$

موفق باشید