بنام خدا

الگوشناسی آماری (CE-725)
دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف
تمرینات سری سوم – بهار ۱۳۸۹

به نکات زیر توجه فرمائید:

۱. زمان تحویل تمرینات در سایت درس مشخص شده است. دقت نمائید که زمانبندی‌های تعیین شده قابل تغییر نیستند.

۲. تمرینات را با عنوان SPR-HWx-8xxxxxxx (مثلا SPR-HW3-88300785) و در یک فایل فشرده با همین نام به آدرس Muhammadi@ce.sharif.edu ایمیل زده و *در اولین جلسه بعد از زمان تحویل، بصورت پرینت شده تحویل استاد درس دهید.*

۳. گزارش شما باید مختصر و مفید باشد. برای تمرینات پیاده‌سازی که با لوگوی مشخص شده‌اند باید کد مطلب نوشته شده ضمیمه گزارش شده و تمامی خروجی‌های برنامه‌ها در گزارش شما ذکر شوند.

سوال ۱) با یک مجموعه داده یک کلاسه‌بند را دو بار آموزش می‌دهیم. بار اول داده‌ها را به صورت تصادفی به نسبت ۵۰–۵۰ به دو دسته آموزش و تست تقسیم کرده و عملیات آموزش کلاسه‌بند را انجام می‌دهیم و بار دوم داده‌ها را به صورت تصادفی به نسبت ۲۰–۸۰ به دو دسته آموزش و تست تقسیم می‌کنیم (مجموعه آموزشی، بزرگتر از مجموعه تست) و عملیات آموزش کلاسه‌بند را انجام می‌دهیم. کارایی کلاسه‌بند اول بر روی مجموعه آموزشی‌اش ۸۰ درصد و کارایی کلاسه‌بند دوم بر روی مجموعه آموزشی‌اش ۹۰ درصد بوده است. آیا برای یک مجموعه داده جدید کلاسه‌بند دوم بهتر از کلاسه‌بند اول عمل می‌کند؟ نظر خود را توضیح دهید.

سوال ۲) یک مساله کلاسه‌بندی دو کلاسه چند بعدی را در نظر بگیرید که هر کدام از کلاس‌ها از یک توزیع نرمال تبعیت می‌کنند. کوواریانس این دو توزیع نیز با هم برابر است.

الف) نشان دهید که لگاریتم نسبت تشابه (likelihood ratio) نسبت به بردار ویژگی‌ها خطی است.

ب) اگر کوواریانس کلاس اول a برابر کوواریانس کلاس دوم شود (a یک عدد اسکالر می‌باشد)، مرز این دو کلاس را پیدا کرده و به ساده‌ترین فرم ممکن بنویسید.

پ) شرایطی را بر حسب احتمال‌های پیشین این دو توزیع پیدا کنید که مرز تصمیم‌گیری بدست آمده بوسیله بیز، بین دو میانگین واقع نشود.

سوال ۳) دو کلاس با توزیع‌های گاوسی و با پارامترهای زیر در نظر بگیرید:

$$p_1 = p_2 = 0.5,\ \mu_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \mu_2 = \begin{bmatrix} +1 \\ 0 \end{bmatrix},\ \Sigma_1 = \Sigma_2 = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$$

الف) میزان خطای کمینه را برای این مساله کلاسه‌بندی بیابید. خطای تشخیص کلاس اول و دوم را جداگانه نیز ارائه دهید.

ب) اگر بخواهیم خطای تشخیص کلاس اول، از نصف حالت قبل بیشتر نشود، مرز کلاسه بندی به چه صورت تغییر خواهد کرد (از روش Neyman-Pearson استفاده کنید). در این صورت خطای تشخیص کلاس دوم چقدر خواهد شد و میزان کل خطا نسبت به حالت قبل به چه صورت تغییر می‌کند؟

پ) مرز تصمیم‌گیری را برای این مساله در صورتی که بخواهیم ریسک را کمینه کنیم، در حالتی که 0-1 loss داشته باشیم با حالتی که ماتریس loss بصورت زیر باشد، را بدست آورده و بر روی یک نمودار هر دوی آنها را نمایش دهید:

$$L = \begin{bmatrix} 0 & 0.5 \\ 1 & 0 \end{bmatrix}$$

ت) با فرض منفی شدن درایه‌های قطر فرعی ماتریس کوواریانس کلاس دوم، یک تابع جداساز خطی را بیابیدکه تابع هدف فیشر را، با بدست آوردن w مناسب و همچنین میزان خطا را، با بدست آوردن یک حد آستانه (w_0) مناسب کمینه کند.

سوال ٤) یک مساله دو کلاسه با احتمال‌های پیشین مساوی را در نظر بگیرید. نمونه‌های زیر از این مساله موجود می‌باشند:

x_1	x_2	x_3	$p_1(x)$	$p_2(x)$
-1	-1	-1	1/3	0
+1	-1	-1	1/24	1/8
-1	+1	-1	1/24	1/8
+1	+1	-1	0	1/3
-1	-1	+1	1/3	0
+1	-1	+1	1/24	1/8
-1	+1	+1	1/24	1/8
+1	+1	+1	0	1/3

الف) یک تابع جداساز خطی برای این مساله با کمینه‌سازی خطای mean-square ارائه دهید.

ب) شش نمونه زیر را در نظر بگیرید:

w_1: (1, 2), (2, -4), (-3, -1)

w_2: (2, 4), (-1, 3), (5, 0)

آیا این نمونه‌ها بصورت خطی جدا پذیر هستند؟ در صورت مثبت بودن جواب یک جداساز خطی برای این مساله با روش Minimum Squared Error ارائه دهید.

سوال ٥) نشان دهید اگر $f_i(x),\ i=1, 2, ..., M$ توابع جداساز یک مساله M کلاسه باشند، می‌توانیم M-1 تابع جدید با توجه به آنها بسازیم، که همان عملکرد قبلی را داشته باشند (نکته: از تفاضل توابع جداساز می‌توانید کمک بگیرید).

سوال ٦) برای یک مساله دو کلاسه یک بعدی قانون تصمیم‌گیری زیر را در نظر بگیرید:

یک نمونه را به کلاس w_1 نسبت بده اگر $x>\theta$ در غیر این صورت نمونه را به w_2 نسبت بده.

الف) نشان دهید که احتمال خطا برای این قانون مطابق رابطه زیر می‌باشد:

$$p(error) = P(w_1)\int_{-\infty}^{\theta} p(x \mid w_1)dx + P(w_2)\int_{\theta}^{+\infty} p(x \mid w_2)dx$$

ب) نشان دهید که شرط لازم برای کمینه شدن احتمال خطا این است که $p(\theta \mid w_1)P(w_1) = p(\theta \mid w_2)P(w_2)$ باشد.

سوال ۷) در بسیاری از مسایل کلاسه‌بندی یک نمونه جدید یا به یکی از c کلاس نسبت داده می‌شود یا بدلیل عدم تشخیص رد (reject) می‌شود. اگر هزینه (cost) رد کردن یک نمونه زیاد نباشد، ممکن است رد کردن، یک عمل دلخواه کلاسه‌بند شود. با توجه به این توضیحات تابع loss زیر را در نظر بگیرید:

$$\lambda(\alpha_i | w_j) = \begin{cases} 0 & i = j \\ \lambda_r & i = c+1 \\ \lambda_s & otherwise \end{cases}$$

الف) نشان دهید که کمترین ریسک زمانی رخ می‌دهد که داشته باشیم:

«یک نمونه جدید را به کلاس w_i نسبت بده اگر به ازای تمامی jهای دیگر $P(w_i | x) \geq P(w_j | x)$ بوده و همچنین $P(w_i | x) \geq 1 - \lambda_r / \lambda_s$ باشد، در غیر اینصورت بدلیل عدم تشخیص نمونه را رد کن.»

ب) اگر λ_r صفر باشد چه اتفاقی می‌افتد؟

پ) اگر $\lambda_r > \lambda_s$ باشد چه اتفاقی می‌افتد؟

ت) با افزایش λ_r / λ_s از صفر به یک چه اتفاقی می‌افتد؟

ث) با توجه به نتایج قسمت‌های قبلی نشان دهید که تابع‌های جداساز زیر برای این مساله بهینه هستند:

$$g_i(x) = \begin{cases} p(x | w_i) P(w_i) & i = 1, \ldots, c \\ \dfrac{\lambda_s - \lambda_r}{\lambda_s} \sum_{j=1}^{c} p(x | w_j) P(w_j) & i = c+1 \end{cases}$$

سوال ۸) نمودار ROC را برای حالتی که دو کلاس دارای دو توزیع گاوسی $P(x, w_1) \sim N(0,1)$ و $P(x, w_2) \sim N(1,2)$ باشند، رسم کنید.

سوال ۹) الف) آیا نمونه‌های جداپذیر خطی کامل (totally linearly separable)، جداپذیر خطی (linearly separable) نیز هستند؟ عکس این حالت چطور؟

ب) آیا نمونه‌های جداپذیر خطی دو به دو (Pairwise linearly separable)، جداپذیر خطی نیز هستند؟ عکس این حالت چطور؟