



پردیس دانشکده های فنی

به نام خدا
دانشکده‌ی مهندسی برق و کامپیوتر
تمرین سری اول یادگیری ماشین



دانشگاه تهران

سلام بر دانشجویان عزیز، چند نکته مهم:

1. حجم گزارش به هیچ عنوان معیار نمره دهی نیست، در حد نیاز توضیح دهید.
2. نکته‌ی مهم در گزارش نویسی روشن بودن پاسخ‌ها می‌باشد، اگر فرضی برای حل سوال استفاده می‌کنید حتماً آن را ذکر کنید، اگر جواب نهایی عددی است به صورت واضح آن را بیان کنید.
3. کدهای ارسال شده بدون گزارش فاقد نمره می‌باشند.
4. برای سوالات شبیه سازی، فقط از دیتاست داده شده استفاده کنید.
5. فایل نهایی خود را در یک فایل زیپ شامل، pdf گزارش و فایل کدها آپلود کنید. نام فایل زیپ ارسالی الگوی `ML_HW#_StudentNumber` داشته باشد.
6. از بین سوالات **شبیه سازی** حتماً به هر دو مورد پاسخ داده شود.
7. نمره تمرین ۱۰۰ نمره می‌باشد و حداکثر تا نمره ۱۱۰ (**۱۰ نمره امتیازی**) می‌توانید کسب کنید.
8. هرگونه شباهت در گزارش و کد مربوط به شبیه سازی، به منزله تقلب می‌باشد و کل تمرین برای طرفین **صفر** خواهد شد.
9. در صورت داشتن سوال، از طریق ایمیل `taheriarmin60@gmail.com` سوال خود را مطرح کنید.

سوال ۱: (۲۰ نمره)

تابع توزیع کوشی^۱ را برای یک مسئله طبقه‌بندی دو کلاسه و یک بعدی در نظر بگیرید:

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

(آ) با فرض $P(\omega_1) = P(\omega_2)$ ، نشان دهید $P(\omega_1|x) = P(\omega_2|x)$ اگر $x = \frac{a_1 + a_2}{2}$ به کمک متلب یا پایتون

$P(\omega_1|x)$ و $P(\omega_2|x)$ روی یک axis رسم کنید. ($a_1 = 3, a_2 = 5, b = 1$)

(ب) نشان دهید که حداقل احتمال خطا برابر است با:

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

(پ) بیش‌ترین مقدار $p(error)$ چیست و تحت چه شرایطی اتفاق می‌افتد؟

(ت) یک طبقه‌بند بهینه بیزی طراحی کنید بر اساس a_i, b اگر $P(\omega_1) = P(\omega_2)$. مرز تصمیم^۲ را برای این

حالت رسم کنید و میزان احتمال خطا را گزارش کنید.

(ث) یک طبقه‌بند بیزی برای کمینه کردن ریسک با وزن‌های زیر طراحی کنید:

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

مرز تصمیم را رسم کنید و احتمال خطا را گزارش کنید. نتایج حاصل از این بخش را با نتایج بخش قبل مقایسه

کنید.

¹ Cauchy distribution

² Decision boundary

سوال ۲: (۱۰ نمره)

یک مسئله دو کلاسه یک بعدی با توزیع رایلی^۳ برای هر دو کلاس را در نظر بگیرید:

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

با فرض یکسان بودن توزیع پیشین^۴ برای هر دو کلاس، مرز تصمیم را محاسبه نمایید.

³ Rayleigh distribution

⁴ Prior

سوال ۳: (۲۰ نمره)

آ) مرز تصمیم را برای یک مسئله دو کلاسه مطابق شکل زیر محاسبه کنید. کلاس یک را مجموعه نقاط مشکی و کلاس دو را مجموعه نقاط قرمز در نظر بگیرید.

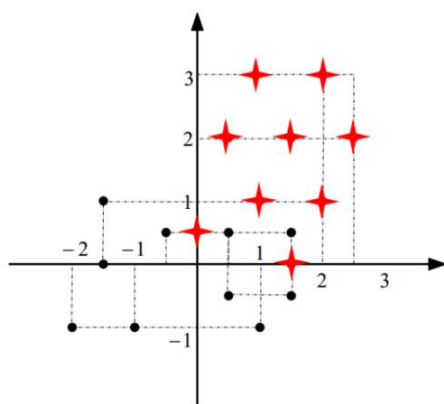


Figure 1

ب) میانگین و کواریانس را برای توزیع گوسی هر دو کلاس بیابید.

پ) فرض کنید احتمال پسین^۵ برای هر دو کلاس برابر با 0.5 باشد. مرز تصمیم را بیابید و رسم کنید. خطای آموزش تجربی را روی این داده‌ها محاسبه کنید. (مثلا درصد نقاطی که اشتباه طبقه‌بندی شده‌اند).

ت) مرز تصمیم را برای یک طبقه‌بند بیزی برای کاهش ریسک با مقادیر زیر بیابید و رسم کنید.

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

ث) برای مقادیر $P(\omega_1) = \frac{1}{3}$, $P(\omega_2) = \frac{2}{3}$ ، بخش ب را تکرار کنید.

⁵ Posterior

سوال ۴: (۲۰ نمره)

متغیر تصادفی X را با توزیع احتمال پواسون^۶ با پارامتر λ در نظر بگیرید:

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

فرض کنید یک مجموعه داده شامل n نمونه $D = \{X_1, \dots, X_n\}$ از این متغیر تصادفی در اختیار داریم.

(آ) تابع لگاریتم درست‌نمایی^۷ را تشکیل داده و تخمین‌گر بیشینه درست‌نمایی^۸ را برای پارامتر λ به دست آورید.

(ب) توزیع احتمال پیشین زیر را برای پارامتر λ در نظر بگیرید:

$$P(\lambda) = \text{Gamma}(\lambda|\alpha, \beta) = c\lambda^{\alpha-1}e^{-\beta\lambda}$$

که در رابطه بالا، c یک ضریب ثابت است و α و β پارامترهای توزیع گاما هستند. توزیع احتمال پسین را برای پارامتر λ به دست آورید.

$$P(\lambda|D) = ?$$

(پ) آیا توزیع احتمال پیشین فوق، برای پارامتر λ یک *conjugate prior* است؟ توضیح دهید.

(ت) با استفاده از توزیع احتمال پیشین فوق، تخمین‌گر *MAP* برای پارامتر λ چیست؟ (راهنمایی: مقدار بیشینه توزیع

گاما در نقطه $\lambda = \frac{\alpha-1}{\beta}$ رخ می‌دهد).

(ث) آیا اگر تعداد داده‌ها به بی‌نهایت میل کند، تخمین‌گر *MAP* به *MLE* میل می‌کند؟ توضیح دهید.

(ج) توضیح دهید در چه شرایطی استفاده از هر کدام از این دو روش تخمین بر دیگری برتری دارد.

^۶ poisson

^۷ Log likelihood

^۸ Maximum likelihood

سوال 5: (شبیه سازی، ۲۰ نمره)

هدف از این سوال آشنایی و پیاده سازی طبقه‌بند naïve bayes است.

آ) در ابتدا در مورد طبقه‌بند naïve bayes توضیح دهید و تفاوت ساختاری آن را با یک طبقه‌بند بیزی بیان کنید. توضیح دهید که چرا به جای طبقه‌بند بیز از این طبقه‌بند استفاده می‌کنیم، هزینه‌ای که می‌دهیم چیست و در چه زمان‌هایی استفاده از این طبقه‌بند کاری منطقی است.

مجموعه داده penguins به پیوست ارسال شده. در ابتدا در صورت نیاز روی داده‌ها پیش‌پردازش انجام دهید. (هر پیش‌پردازی که روی داده‌ها انجام می‌دهید را باید با ذکر دلیل توضیح دهید).

ب) این مجموعه داده شامل سه کلاس است. یک طبقه‌بند naïve bayes را از پایه و بدون استفاده از کتابخانه پیاده‌سازی کنید. و برای هر کلاس به صورت 1 vs all از طبقه‌بندی که طراحی کردید استفاده کنید. دقت، Recall، precision و ماتریس آشفتگی را بررسی و تحلیل نمایید.

پ) مورد ب را به کمک کتابخانه SKLEARN انجام دهید. نتایج دو بخش را مقایسه کنید.

سوال ۶: (شبیه سازی، ۲۰ نمره)

یک طبقه‌بند دو کلاسه برای تشخیص تصاویر مربوط به دریا و جنگل در مجموعه داده image طراحی کنید. الگوریتم پیاده‌سازی را روی داده‌ها تست نمایید و دقت، ماتریس آشفستگی، Precision و Recall را گزارش کنید. (راهنمایی: برای طبقه‌بندی نیازی به استفاده از طبقه‌بند معروفی نیست صرفاً از ویژگی‌های داده مانند رنگ برای جداسازی استفاده نمایید).

داده‌هایی که به اشتباه جداسازی شدند را معرفی کنید و بیان کنید با توجه به ویژگی‌ای که بر طبق آن جداسازی انجام شده آیا این اشتباهات منطقی است یا خیر.