تمرین شماره دوم درس یادگیری ماشین

پریا پاسه‌ورز

شماره دانشجویی: 810101393

سوال سوم)

برای یک مسئله طبقه‌بندی دو کلاسه با در نظر گرفتن متغیر ورودی تک بعدی X،

$$\boldsymbol{P(x|y=0) = N(0, \sigma^2)}$$

$$\boldsymbol{P(x|y=1) = N(2, \sigma^2)}$$

و فرض برابر بودن احتمال پیشین دو کلاس، آستانه به حداقل رسیدن ریسک با فرض اینکه،

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

به دست آوردید و تاثیر افزایش a را در آستانه به دست آمده تفسیر کنید.

طبق رابطه conditional risks داریم:

$$R(\alpha_1|x) = \lambda_{11}P(\omega_1|x) + \lambda_{12}P(\omega_2|x)$$

$$R(\alpha_2|x) = \lambda_{21}P(\omega_1|x) + \lambda_{22}P(\omega_2|x)$$

حال اگر مقادیر لاندا را جایگذاری کنیم خواهیم داشت:

$$R(\alpha_1|x) = a\ P(\omega_2|x)$$

$$R(\alpha_2|x) = a^2 P(\omega_1|x)$$

برای به دست آوردن آستانه به حداقل رسیدن ریسک، باید نقطه‌ای را پیدا کنیم که:

$$R(\alpha_1|x) \ = R(\alpha_2|x)$$

$$a\ P(\omega_2|x) \ = a^2 P(\omega_1|x)$$

طبق رابطه بیز:

$$aP(\omega_2|x) \ = a^2 P(\omega_1|x)$$

$$a\frac{p(x|\omega_2)p(\omega_2)}{p(x)} = a^2\frac{p(x|\omega_1)p(\omega_1)}{p(x)}$$

با توجه به برابر بودن احتمال پیشین دو کلاس خواهیم داشت:

$$a\ p(x|\omega_2) = a^2 p(x|\omega_1)$$

از طرف دیگر، با توجه اینکه توزیع گاوسی داریم:

$$\mathrm{P}(\mathrm{x}|\mathrm{y}=0)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\mathrm{x}^2}{2\sigma^2}\right)$$

$$\mathrm{P}(\mathrm{x}|\mathrm{y}=1)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(\mathrm{x}-2)^2}{2\sigma^2}\right)$$

پس:

$$a\ \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-2)^2}{2\sigma^2}\right)=a^2\,\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right)$$

$$a\ \exp\left(-\frac{(x-2)^2}{2\sigma^2}\right)=a^2\exp\left(-\frac{x^2}{2\sigma^2}\right)$$

از دو طرف رابطه لگاریتم می‌گیریم:

$$-\frac{(\mathrm{x}-2)^2}{2\sigma^2}=\ln(a)-\frac{\mathrm{x}^2}{2\sigma^2}$$

$$\ln(\mathrm{a})=\frac{\mathrm{x}^2-(\mathrm{x}-2)^2}{2\sigma^2}$$

$$\ln(\mathrm{a})=\frac{\mathrm{x}^2-(\mathrm{x}^2-4\mathrm{x}+4)}{2\sigma^2}$$

$$\ln(a)=\frac{4\mathrm{x}-4}{2\sigma^2}$$

$$\ln(a)=\frac{2(\mathrm{x}-1)}{\sigma^2}$$

$$\mathrm{x}=1+\frac{\sigma^2\ln(a)}{2}$$

افزایش مقدار a باعث می‌شود که آستانه ریسک افزایش یابد و به مقدار $P(x|y=1)$ نزدیک شود، پس دیتاهای بیشتری به عنوان y = 1 دسته‌بندی می‌شوند. از طرف دیگر مدل نسبت به پیش‌بینی y = 0 حساس‌تر می‌شود، چون $a\,P(y\ =\ 0|x)$ سریعتر رشد می‌کند و برای $P(y\ =1|x)$ برقرار ماندن نامساوی سخت‌تر می‌شود، پس مدل نمونه‌های کمتری را به عنوان y = 0 پیش‌بینی می‌کند.