۴۰۱۲۱۰۹۸۹ تمرین ۱ یادگیری عمیق یاسر مرادی

سوالات تشریحی

(1)

Naive Bayes Classifier $\rightarrow P(\underline{x} \mid \Omega_i) \cong P(x_1 \mid \Omega_i)\, P(x_2 \mid \Omega_i) \cdots P(x_n \mid \Omega_i)$

$\underline{x} = (x_1, x_2, \ldots, x_n)^T$

$\underline{x} = (x_1, x_2) \rightarrow P(\underline{x} \mid Y_i) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma|}} \exp\left(-\frac{1}{2}(\underline{x}-\mu_i)^T \Sigma_i^{-1} (x-\mu_i)\right)$

الف) $x = [50, 0.5]$

$P(\underline{x} \mid Y=1) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_1|}} \exp(-875.0875)$

$P(\underline{x} \mid Y=2) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_2|}} \exp(-953.075)$

$P(\underline{x} \mid Y=3) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_3|}} \exp(-954.075)$

$\boxed{Y=1}$

$P(\underline{x} \mid Y=1) > P(\underline{x} \mid Y=2),\ P(\underline{x} \mid Y=3)$

ب) $x = [0.5, 0.5]$

$P(\underline{x} \mid Y=1) = 0.1591 \qquad P(\underline{x} \mid Y=2) = 0.2945 \qquad P(\underline{x} \mid Y=3) = 0.1494$

$P(\underline{x} \mid Y=2) \rightarrow \max \qquad \boxed{Y=2}$

۵

الف)

$$E(X) = \sum x P(x) \xrightarrow[P(x)\geq 0]{x\geq 0} \sum x P(x) \geqslant \sum_{x\geqslant \alpha} x P(x)$$

با فرض متغیر تصادفی گسسته X

$$\sum x P(x) \geqslant \sum_{x\geqslant\alpha} x P(x) \geqslant \sum_{x\geqslant\alpha} \alpha P(x) = \alpha \sum_{x\geqslant\alpha} P(x) = \alpha P(X\geqslant\alpha)$$

$$\Rightarrow E(X) \geqslant \alpha P(X\geqslant\alpha) \Rightarrow \frac{E(X)}{\alpha} \geqslant P(X\geqslant\alpha)$$

برای متغیر تصادفی پیوسته Σ تبدیل به $\int$ می‌شود

ب)

$$P(|Z-\mu| \geqslant \varepsilon) = P((Z-\mu)^2 \geqslant \varepsilon^2) \underset{\text{نتیجه بخش الف}}{\leqslant} E\left(\frac{|Z-\mu|^2}{\varepsilon^2}\right)$$

$$= \frac{E(|Z-\mu|^2)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

پ) هر نقطه متغیر تصادفی x_i با احتمال $\frac{\pi}{4}$ در داخل دایره واحد و با احتمال $1-\frac{\pi}{4}$ خارج، ۰، دایره واحد قرار می‌گیرد.

$$x_i \sim \text{Bernoulli}\left(\frac{\pi}{4}\right), \quad i = \{1, \dots, N\}$$

$$X = \sum_{i=1}^{N} x_i \sim \text{Binomial}\left(N, \frac{\pi}{4}\right) \Rightarrow \begin{cases} E\left[\sum_{i=1}^{N} x_i\right] = N \times \frac{\pi}{4} \\ Var\left[\sum_{i=1}^{N} x_i\right] = N \times \frac{\pi}{4}\left(1-\frac{\pi}{4}\right) \end{cases}$$

$$\rightarrow X = \frac{\sum_{i=1}^{N} x_i}{N} \Rightarrow \begin{cases} E\{X\} = \frac{\pi}{4} \\ Var\{X\} = \frac{\pi}{4N}\left(1-\frac{\pi}{4}\right) \end{cases}$$

$Z=4X$ تمرین ۵ام

$$Z=4X \rightarrow \begin{cases} \mu_Z = \pi \\ Var(Z) = \frac{\pi}{N}(4-\pi) \end{cases}$$

نامساوی چبیشف $$P(|Z-\mu_Z| \geq \varepsilon) = \frac{\sigma_Z^2}{\varepsilon^2}$$

$$\rightarrow 1 - P(|Z-\mu_Z| \geq \varepsilon) = 1 - \frac{\sigma_Z^2}{\varepsilon^2}$$

$$\rightarrow P(|Z-\mu_Z| \leq \varepsilon) = 1 - \frac{\sigma_Z^2}{\varepsilon^2}$$

$$\varepsilon = 0.01\pi \rightarrow 1 - \frac{\sigma_Z^2}{\varepsilon^2} = 0.95 \rightarrow \sigma_Z^2 = 0.05\varepsilon^2$$

$$= 0.05(0.01)^2\pi^2$$

$$\Rightarrow \frac{\pi}{N}(4-\pi) = 0.05(0.01)^2\pi^2$$

$$\rightarrow N = \frac{4-\pi}{0.05(0.01)^2\pi} = 54648$$

Scanned with CamScanner

$$A^{-1} = V \Sigma^{-1} U^T$$

الف)

$$\|A^{-1}\|_2 = \sigma_{max}(A^{-1}) = \max_i \operatorname{diag}(\Sigma^{-1})_i$$

$$= \max_i \frac{1}{\sigma_i} = \frac{1}{\min \sigma_i} = \frac{1}{\min \sigma(A)}$$

$$A = (V \Sigma^{-1} U^T)^{-1} = U \Sigma V^T \rightsquigarrow \|A\|_2 = \sigma_{max}(A) = \max \sigma_i$$

$$\sigma_{max}(A^{-1}) = \frac{1}{\sigma_{min}(A)}$$

$$\sigma_{max}(A^{-1})\,\sigma_{max}(A) = \frac{\sigma_{max}(A)}{\sigma_{min}(A)} \geq 1$$

ب)

$$\|A\|_2^2 = \lambda_{max}(A^H A) = \sigma_{max}^{2}$$

$$\|A\|_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2 = \operatorname{trace}(A^H A) = \sum_{i=1}^{n} \lambda_i (A^H A) \geq \lambda_{max}(A^H A)$$

$$= \|A\|_2^2 \quad \Rightarrow \quad \|A\|_F^2 \geq \|A\|_2^2$$

$\operatorname{rank}(A) =$ تعداد مقادیر تکین غیرصفر A $=$ مقادیر ویژه غیرصفر $A^H A$

$$\|A\|_F^2 = \sum_{i=1}^{n} \lambda_i (A^H A) = \sum_{\substack{i=1 \\ \lambda_i \neq 0}}^{n} \lambda_i (A^H A) \leq \lambda_{max}(A^H A) = \operatorname{rank}(A) \times \lambda_{max}(A^H A)$$

$$= \operatorname{rank}(A) \times \|A\|_2^2 \Rightarrow \|A\|_F^2 \leq \operatorname{ran}(A) \cdot \|A\|_2^2$$

$$\|A\|_2^2 \leq \|A_F\|^2 \leq \operatorname{ran}(A)\, \|A\|_2^2$$

$$\Rightarrow \|A\|_2 \leq \|A\|_F \leq \sqrt{\operatorname{rank}(A)}\, \|A\|_2$$

$$\sigma(x) = \frac{1}{1+e^{-x}} = \frac{e^{x}}{1+e^{x}}$$

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} = \frac{e^{2x}-1}{e^{2x}+1}$$

$$\frac{1+\tanh\left(\frac{x}{2}\right)}{2} = \frac{1}{2} + \frac{e^{x}-1}{(e^{x}+1)2} = \frac{2e^{x}}{2(1+e^{x})} = \frac{e^{x}}{1+e^{x}} = \sigma(x)$$

$$y(x,w) = w_0 + \sum_{j=1}^{n}\left[w_j\,\sigma\left(2\,\frac{x-\mu_j}{s}\right)\right]$$

$$= w_0 + \sum_{j=1}^{n}\left[w_j \times \frac{1}{2}\left(1+\tanh\left(\frac{x-\mu_j}{s}\right)\right)\right]$$

$$= \underbrace{w_0 + \sum_{j=1}^{n}\frac{w_j}{2}}_{u_0} + \sum_{j=1}^{n}\left[\underbrace{\frac{w_j}{2}}_{u_j}\tanh\left(\frac{x-\mu_j}{s}\right)\right]$$

$$= u_0 + \sum_{j=1}^{n}\left[u_j \tanh\left(\frac{x-\mu_j}{s}\right)\right]$$

$$u_0 = w_0 + \sum_{j=1}^{n}\frac{w_j}{2}$$

$$u_j = \frac{w_j}{2}$$

$j=1,\ldots,n$

Y