Машинное обучение. Задание 2.

Гаркавый Андрей 494

1.

1.1. Ответы в листьях регрессионного дерева.

Пусть $X_1, \dots X_n$ — выборка (точнее таргеты объектов обучающей выборки в листе)

$\overline{X} = \frac{X_1+\dots+X_n}{n}$ — средний таргет

$\hat{X} = \begin{cases} X_1, & \frac{1}{n} \\ \vdots \\ X_n, & \frac{1}{n} \end{cases}$ — таргет случайного объекта.

$$MSE(\overline{X}) = \frac{1}{n}\sum_{i=1}^{n}(\overline{X}-X_i)^2 = \frac{n\overline{X}^2 - 2\overline{X}\sum_{i=1}^{n}X_i + \sum_{i=1}^{n}X_i^2}{n}$$

$$EMSE(\hat{X}) = \frac{1}{n}E\sum_{i=1}^{n}(\hat{X}-X_i)^2 = nE(\hat{X}^2) - 2\sum X_i E(\hat{X}) + \sum_{i=1}^{n}X_i^2$$

$$EMSE(\overline{X}) - EMSE(\hat{X}) = \frac{n(\overline{X}^2 - E\hat{X}^2) - 2\sum_{i=1}^{n}X_i(\overline{X}-E(\hat{X}))}{n} =$$

$$= \frac{n\left(\overline{X}^2 - \sum_{i=1}^{n}X_i^2\cdot\frac{1}{n}\right) - 2\sum_{i=1}^{n}(\overline{X}-\overline{X})}{n} = \overline{X}^2 - \overline{X^2} = -S^2 \le 0$$

$\Rightarrow EMSE(\overline{X}) \le EMSE(\hat{X})$

$\Rightarrow$ $\overline{X}$ — средний таргет приводит к меньшей ошибке в среднем по MSE

1.3. $f(x) = \frac{e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}}$

$$\Rightarrow H(x) = -\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x) \ln f(x) \, dx_1 \dots dx_n =$$

$$= -\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \left(-\frac{1}{2} \ln\left((2\pi)^n |\Sigma|\right) - \frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) \right) \cdot f(x) \, dx_1 \dots x_n =$$

$$= \frac{1}{2} \ln\left((2\pi)^n |\Sigma|\right) \cdot \underbrace{\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x) \, dx_1 \dots dx_n}_{=1} + \frac{1}{2} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \underbrace{(x-\mu)^T \Sigma^{-1} (x-\mu)}_{\text{число}} f(x) \, dx_1 \dots dx_n$$

$$= \frac{1}{2} \ln\left((2\pi)^n |\Sigma|\right) + \frac{1}{2} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \mathrm{Tr}\left((x-\mu)^T \Sigma^{-1} (x-\mu)\right) f(x) \, dx_1 \dots dx_n =$$

$$= \frac{1}{2} \ln\left((2\pi)^n |\Sigma|\right) + \frac{1}{2} \mathrm{Tr}\left(\Sigma^{-1} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} (x-\mu)(x-\mu)^T f(x) \, dx_1 \dots dx_n\right) =$$

$$= \frac{1}{2} \ln\left((2\pi)^n |\Sigma|\right) + \frac{1}{2} \mathrm{Tr}\, \Sigma^{-1} \Sigma = \frac{1}{2} \ln (2\pi)^n |\Sigma| + \frac{1}{2} \cdot n =$$

$$= \frac{1}{2} \ln\left((2\pi e)^n |\Sigma|\right)$$