Машинное обучение. Теоретическое задание 2.

Алёна Егорова, 494 группа

Задача 1.

Введем обозначения: y_{real}, y_{answer} – реальное значение и ответ.
Тогда в первом случае: $y^1_{answer} = \dfrac{1}{n}\sum\limits_{i=0}^{n} y_i$. Во втором случае: $y^2_{answer} = y_k$.
Распишем MSE для каждого случая:

$$E(\frac{1}{n}\sum_{i=0}^{n} y_i - y_{real})^2 = E(\frac{1}{n}\sum_{i=0}^{n} y_i)^2 - \frac{2}{n}E(\sum_{i=0}^{n} y_i \cdot y_{real}) + Ey^2_{real}$$

$$E(y_k - y_{real})^2 = Ey^2_k - 2E(y_k \cdot y_{real}) + Ey^2_{real}$$

Заметим, что 2 последних слагаемых одинаковы. Сравним тогда теперь $E(\dfrac{1}{n}\sum\limits_{i=0}^{n} y_i)^2$ и Ey^2_k. Раскрыв скобки в первом выражении, получим: $\dfrac{1}{n^2}\sum_{i=0}^{n} Ey^2_i + \dfrac{2}{n^2}\sum_{i\neq j}^{n} Ey_iy_j \geq \dfrac{n}{n^2}Ey^2_k + \dfrac{2n(n-1)}{2n^2}Ey^2_k \geq \dfrac{1}{n}Ey^2_k + \dfrac{(n-1)}{n}Ey^2_k$.
Тогда получим, что $E(\dfrac{1}{n}\sum\limits_{i=0}^{n} y_i)^2 \geq Ey^2_k = Ey^2_k$. Переходя к исходным выражениям, получаем: $E(\dfrac{1}{n}\sum\limits_{i=0}^{n} y_i - y_{real})^2 \geq E(y_k - y_{real})^2$. Ответ средними дает меньшую ошибку.

Задача 3.

Известно, что плотность многомерного нормального распределения:

$$p(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{-\frac{(x-a)^T \Sigma^{-1} (x-a)}{2}},$$

где Σ – матрица ковариаций
a – вектор средних

Энтропия: $H(f) = -\int\limits_{\mathbb{R}^n}\dots\int f(x)\ln(f(x))\,dx$

Преобразуем этот многомерный интеграл:

$$\int\limits_{\mathbb{R}^n}\dots\int f(x)\left(\frac{1}{2}(x-a)^T\Sigma^{-1}(x-a) + \ln\left(\sqrt{(2\pi)^n|\Sigma|}\right)\right)dx$$

$$= \frac{1}{2}E\left(\sum_{i,j}(x_i-a_i)(\Sigma^{-1})_{ij}(x_j-a_j)\right) + \frac{1}{2}\ln((2\pi)^n|\Sigma|) =$$

$$= \frac{1}{2}\overbrace{\sum_{i,j}}^{\text{дв-ая сумма}} \Sigma_{ij}(\Sigma)^{-1}_{ij} + \frac{1}{2}\ln((2\pi)^n|\Sigma|) =$$

$$= \frac{1}{2}\sum_{i} \overset{\text{единички}}{I_{ii}} + \frac{1}{2}\ln((2\pi)^n|\Sigma|) =$$

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

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