# Задание 1

Шишкин Евгений

27 февраля 2017 г.

**I** При равенстве априорных плотностей класса мы выбираем тот, у которого наибольшая плотность вероятности на данных признаках. Так как классификатор наивный баесовский, то

$$\begin{aligned} P(x|y) = \prod_{k=1}^{n} P(x^{(k)}|y) = \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x^k - \mu_y^{(k)})^2}{2\sigma^2}} &= \\ = \frac{1}{\sqrt[n/2]{2\pi\sigma^2}} e^{-\frac{\sum_{k=1}^{n}(x^k - \mu_y^{(k)})^2}{2\sigma^2}} &= \\ = \frac{1}{\sqrt[n/2]{2\pi\sigma^2}} e^{-\frac{\rho^2(x,\mu_y)}{2\sigma^2}}&. \end{aligned}$$

Тогда наибольшая плотность будет у того класса $y$, чей центр $\mu_y$ ближе всего находится к признакам $x$.

**II** По условию с вероятностью $p$ классификатор относит обьект к классу 1 и с вероятностью $(1-p)$ к классу 0. Пусть выборка размера $N$, и $N_0$ и $N_1$ - размеры 0 и 1 класса в большой выборке.

$$TN + FP = N_0 \qquad TP + FN = N_1$$

ROC-кривая строится в координатах $TPR = \frac{TP}{N_1}$ $FPR = \frac{FP}{N_0}$. Так как классификатор рандомный, то это будет ломаная, построенная по трем точкам $(0,0)$, $(TPR, FPR)$ и $(1,1)$.

$$S = 1 - \frac{TPR}{2} - \frac{1-FPR}{2} = \frac{1+FPR-TPR}{2},$$

$$2S - 1 = FPR - TPR = \frac{FP}{N_0} - \frac{TP}{N_1} = \frac{FP \cdot N_1 - TP \cdot N_0}{N_0 \cdot N_1},$$

$$\begin{aligned} E[FP \cdot N_1 - TP \cdot N_0] &= E[FP \cdot N_1 - (p \cdot N - FP) \cdot N_0] = \\ &= E[FP \cdot (N_1 + N_0) - p \cdot N \cdot N_0] = N \cdot E[FP - p \cdot N_0] = 0, \end{aligned}$$

тогда $E[2S-1] = 0$ и $E[S] = 0.5$, что и требовалось доказать.

**III** Баесовский классификатор дает нам ответ на объект класса $y$ с признаками $x$

$$a_B(x) = \begin{cases} 0, & P(0|x) > P(1|x) \\ 1, & P(1|x) > P(0|x) \end{cases},$$

а 1NN классификатор

$$a_n(x) = y_n,$$

где $y_n$ - класс ближайшего соседа. Тогда математическое ожидание того, что ошибется баесовкий классификатор

$$error_B(x) = E_B[y \neq a(x)] = min(P(0|x), P(1|x)),$$

а для 1NN классификатора

$$\begin{aligned} error_n = E_n[y \neq a(x)] &= P(y \neq y_n) = \\ &= P(y = 0, y_n = 1) + P(y = 1, y_n = 0) = \\ &= P(0|x) \cdot P(1|x_n) + P(1|x) \cdot P(0|x_n) = \\ &= |\text{по ассимптотическому приближениею } x_n \to x| = \\ &= 2 \cdot P(0|x) \cdot P(1|x) = \\ &= 2 \cdot error_B(x) \cdot (1 - error_B(x)) = \\ &= 2 \cdot error_B(x) - 2 \cdot error_B^2(x) \geq \\ &\geq 2 \cdot error_B(x), \end{aligned}$$

что и требовалось доказать.