

Семинар 02. Задачи по байесовскому подходу к классификации

1. Пусть $p(x|y = -1) \sim N(0, 1)$, $p(x|y = 1) \sim U(0, 5)$, $p(y = 1) = 0.5$.
Найти оптимальный байесовский классификатор.
2. Пусть $p(x|y = -1) \sim N(0, 1)$, $p(x|y = 1) \sim U(0, 5)$. $p(y = 1) = 0.5$.
Посчитать средний риск.
3. Пусть в задаче двухклассовой классификации $Y = \{-1, +1\}$ и $\eta(x) = p(y = +1|x)$. Выразить оптимальный средний риск R^* .
4. Пусть дана обучающая выборка $(x_1, y_1), \dots, (x_n, y_n)$, где пара (x_i, y_i) — объект и правильный ответ, причем $x_i \in \mathbb{R}^2$, $y_i \in \{0, 1\}$. Описать процедуру обучения наивного гауссовского классификатора. Выписать решающее правило.
5. Доказать, что наивный байесовский классификатор в случае бинарных признаков $f_i \in \{0, 1\}$ является линейным разделителем:

$$a(x) = a(f_1, \dots, f_n) = [a_0 + a_1 f_1 + \dots + a_n f_n > 0].$$

Выведите формулы для коэффициентов a_i .