

Задачи по байесовскому подходу к классификации

Инструкция по получению и сдаче задания Получить задание можно в боте @ml2021sharebot, выполнив команду /get seminar02 В качестве решения необходимо отправить в бот, указав caption seminar02, следующие файлы:

- seminar02_solution.pdf — файл с описанием вашего решения.
 - seminar02.py — реализация функций $a1(x)$, $R()$, $a2(x)$ на языке Python.
 - файлы seminar02_task1.txt, seminar02_task2.txt, seminar02_task3.txt, созданные при запуске python3 seminar02.py.
1. Пусть функция потерь $L(a(x), y) = \lambda_y[a(x) \neq y]$. Для задачи двухклассовой классификации ($Y = \{-1, +1\}$) найти оптимальный байесовский классификатор для заданных $p(x|y = -1)$, $p(x|y = +1)$, $p(y = -1)$, λ_+ , λ_- . Ваши параметры вы сможете найти в файле task1.txt. В качестве решения необходимо прислать описание решения и код функции $a1(x)$ в файле с именем seminar02.py.

Замечание. Список распределений с параметризацией:

- Нормальное, $N(\mu, \sigma^2)$, $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$.
 - Равномерное, $U(a, b)$, $p(x) = \frac{1}{b-a} I_{[a,b]}$.
 - Экспоненциальное, $Exp(\lambda)$, $p(x) = \lambda e^{-\lambda x} I_{[0,+\infty)}$.
 - Лапласа, $Laplace(\alpha, \beta)$, $p(x) = \frac{\alpha}{2} \exp(-\alpha|x - \beta|)$.
 - Коши, $Cauchy(x_0, \gamma)$, $p(x) = \frac{1}{\pi} \left(\frac{\gamma}{\gamma^2 + (x-x_0)^2} \right)$
2. В условиях предыдущей задачи посчитать средний риск для оптимального байесовского классификатора. В качестве решения необходимо прислать описание решения и код функции $R()$ в файле с именем seminar02.py.
3. Найти решающее правило классификации для данных из файлов task2.csv, используя наивный байесовский классификатор. Функция потерь $L(a(x), y) = [a(x) \neq y]$. Предположение о виде распределения необходимо принять самостоятельно при анализе данных. В качестве решения необходимо прислать описание решения и код функции $a2(x)$ в файле с именем seminar02.py.