

Вариант 1

1. Рассматривается задача классификации объектов на два класса по одному признаку. Предполагается, что значение признака x для объектов из классов K_1, K_2 распределено по закону Рэлея:

$$p(x|K_j) = \beta_j x \exp\left(-\frac{\beta_j}{2} x^2\right), \quad x \geq 0, j = 1, 2.$$

Пусть $\beta_1 = 7.3, \beta_2 = 1.3$. Требуется найти области значений признака x , соответствующие отнесению объектов в каждый из двух классов байесовским классификатором, если априорные вероятности классов равны, соответственно, 0.3 и 0.7.

2. Имеется задача распознавания с 4-мя классами и одним признаком. Предполагается, что с использованием метода «Линейная машина» для каждого класса найдены следующие линейные разделяющие функции:

$$\begin{aligned} f_1(x) &= 4.8 - 2.3x, & f_3(x) &= 4.5 - 2.3x, \\ f_2(x) &= -4.6 - 2.6x, & f_4(x) &= 4.2 - 0.4x. \end{aligned}$$

Требуется изобразить на графике области, соответствующие отнесению к каждому из четырех классов.

3. Предполагается, что линейный дискриминант Фишера используется для распознавания объектов из двух классов по паре признаков x_1 и x_2 . Требуется вычислить вектор, задающий направление перпендикуляра к прямой, разделяющей объекты двух классов:

	Класс 1		Класс 2
x_1	2.3 3.0 3.2 3.1		-0.9 -0.1 0.1
x_2	1.8 1.9 2.1 0.6		-3.6 -3.8 -2.8

4. При проведении выборов на ряде избирательных участков производятся фальсификации результатов голосования. Посылка наблюдателя на такой участок предотвращает фальсификации. Пусть известно несколько точек ROC-кривой для метода идентификации «грязных» участков. Требуется определить оптимальную стратегию распределения наблюдателей по участкам и максимальный выигрыш относительно стратегии равномерного распределения по участкам, если всего участков 1000, наблюдателей — 200 и доля «грязных» участков — 30%. При этом под оптимальностью понимается максимизация количества честных участков.

Чувствительность	Ложная тревога
0.86	0.11
0.90	0.31
0.92	0.32

5. Задана таблица совместных значений прогнозируемой переменной Y и объясняющей переменной X . Требуется вычислить ковариацию между Y и X , коэффициент корреляции между Y и X , коэффициенты одномерной линейной регрессии.

Y	5.9 4.0 2.4 1.7
X	8.3 7.6 3.0 2.3

6. Заданы таблицы значений бинарных признаков для классов K_1 и K_2 . Требуется найти все тупиковые тесты минимальной длины, а также указать для каждого класса по одному представительному набору, который не совпадает по признакам с тупиковым тестом.

	Класс 1					Класс 2			
	X1	X2	X3	X4		X1	X2	X3	X4
0	0	0	0	0		1	0	1	1
1	1	0	1	0		1	1	0	0
0	0	0	0	1		1	1	0	0