

## Тестовый вариант

ВСОШ по ИИ • школьный этап • 9–11 класс

### Задание №1

В лаборатории «КотоПёс-АИ» идёт финальный тур шоу «Кот или Пёс?». Два опытных судьи-бота  $h_1$  и  $h_2$  уже выставили оценки для 12 снимков, а правильные ответы  $y$  известны из экспертной разметки. Для повышения рейтинга продюсеры разрешили подключить третьего стажёра-бота  $h_3$ : для каждого снимка его голос можно задать отдельно (независимо от других снимков). Окончательное решение по каждому снимку принимается большинством голосов трёх ботов.

Даны правильные ответы  $y$  и предсказания двух программ  $h_1, h_2$  для 12 объектов:

$i$	1	2	3	4	5	6	7	8	9	10	11	12
$y$	1	0	1	0	1	0	1	0	1	0	1	0
$h_1$	1	1	1	0	0	1	1	0	0	0	1	1
$h_2$	1	1	0	0	0	0	1	1	0	0	0	0

Добавим третью программу  $h_3$ . Итоговый ответ определяется большинством голосов среди  $h_1, h_2, h_3$ . Для какого максимального и какого минимального числа объектов итоговое большинство может совпасть с  $y$ , если  $h_3$  можно выбирать независимо по объектам?

### Задание №2

В цехе фабрики шоколада встречается редкий дефект глазури: из всех плиток лишь 0,2% оказываются бракованными. Автоматический сканер контроля качества работает так: если плитка *бракована*, он пишет «Брак» с вероятностью 95%; если плитка *хорошая*, он ошибочно пишет «Брак» с вероятностью 2%. Случайно выбранная плитка отмечена сканером как «Брак». Какова вероятность, что плитка действительно бракована?

### Задание №3

На числовой прямой даны 10 точек со следующими координатами: синие — 1, 3, 7, 10, 11, 15; красные — 4, 6, 8, 13. Классификатор с порогом  $T$ , который не видит цвета, действует по следующему правилу: точки с координатой  $< T$  считает синими, а точки с координатой  $\geq T$  — красными. Ошибка первого рода — доля синих точек, ошибочно отнесённых к красным; ошибка второго рода — доля красных точек, ошибочно отнесённых к синим. Найдите наименьшее целое  $T$ , при котором сумма двух ошибок минимальна.

### Задание №4

Есть три независимых детектора (№1, №2, №3). Для случайной фразы каждый срабатывает с вероятностью  $1/2$ . Система включает режим слушания, если сработали хотя бы два детектора. Костя упростил правило: «включится» тогда и только тогда, когда сработал детектор №1. Какова вероятность, что предсказание Кости совпадёт с решением системы?

## Задание №5

Дан обучающий текст — строка из маленьких латинских букв без пробелов. Для каждой буквы запроса  $s$  нужно посмотреть все вхождения  $s$  в текст и вывести, какая буква чаще всего стоит сразу после неё; при равенстве берётся первая по алфавитному порядку. Если у  $s$  нет ни одного следующего символа (например,  $s$  встречается только в конце строки или не встречается вовсе), выведите символ ?.

**Формат входных данных.**

Первая строка: обучающий текст  $s$  (только буквы  $a..z$ ).

Вторая строка: целое  $q$  ( $1 \leq q \leq 26$ ).

Третья строка:  $q$  букв  $a..z$  через пробел.

**Формат выходных данных.**

$q$  символов через пробел — ответы для запросов по порядку.

**Ограничения.**

$1 \leq |s| \leq 10^6$ .

	Ввод	Вывод
Пример.	abacaba	b a a
	3	
	a b c	

## Задание №6

В горной долине работает ИИ-метеостанция. Каждую минуту она измеряет температуру воздуха, но отдельные измерения могут быть ненадёжными, поэтому для стабильности показателей используют «скользящее среднее»: для каждого положения окна длины  $k$  берут среднее  $k$  подряд идущих значений. Формально, для целых  $a_1, \dots, a_n$  и целого  $k$  ( $1 \leq k \leq n$ ):

$$b_i = \frac{a_i + a_{i+1} + \dots + a_{i+k-1}}{k} \quad (i = 1, 2, \dots, n - k + 1).$$

Нужно вывести точное значение каждого  $b_i$ : если  $b_i$  — целое, вывести его как целое; иначе — несократимую дробь  $p/q$ .

**Формат входных данных.**

Первая строка:  $n, k$  ( $1 \leq k \leq n \leq 1000, k \leq 10$ ).

Вторая строка:  $a_1, \dots, a_n, |a_i| \leq 1000$  (целые).

**Формат выходных данных.**

$n - k + 1$  значений  $b_i$  в одной строке через пробел — целым числом или несократимой дробью  $p/q$ .

	Ввод	Вывод
Пример.	7 3	3 8/3 3 7/3 2
	3 0 6 2 1 4 1	

## Задание №7

В почтовом ящике  $N = 1000$  писем, из них  $P = 300$  — спам, остальные 700 — обычные письма. В почтовом сервисе также установлен фильтр спама, который не всегда срабатывает корректно. Каждое письмо фильтр либо помечает как «СПАМ», либо никак не помечает. Определим:

$TP$  — число писем-спама, правильно помеченных как «СПАМ»;

$FP$  — число обычных писем, ошибочно помеченных как «СПАМ»;  
 $FN$  — число писем-спама, ошибочно оставленных обычными;  
 $TN$  — число обычных писем, правильно оставленных обычными.

Даны формулы:

Полнота (Recall):  $R = TP/P$ .

Доля верных ответов (Accuracy):  $A = (TP + TN)/N$ .

Известно, что по результатам работы фильтра  $R = 0.7$ .

Найдите минимальное и максимальное возможные значения Accuracy.

### Задание №8

Дан табличный файл в формате csv. В таблице даны три столбца с натуральными числами:  $f_1, f_2, f_3$  (по одной строке на объект). Задан прототип  $v = (2, 1, 2)$ , для которого сумма равна  $S_v = 2 + 1 + 2 = 5$ . Для каждой строки вычислите сумму  $S = f_1 + f_2 + f_3$ . Строка считается похожей на прототип, если  $|S - 5| \leq 2$ . Требуется подсчитать количество похожих на прототип строк в таблице и вывести одно это число.