

Тестовый вариант

ВСОШ по ИИ • школьный этап • 7–8 класс

Задание №1

Текст считается «**длинным**» тогда и только тогда, когда его длина ≥ 60 символов. Простая (но неверная) программа для оценки длины текста сначала удаляет каждый третий символ исходной строки (3-й, 6-й, 9-й и т.д.), а затем сравнивает длину очищенного текста с порогом 60. Программа выводит «**длинный**», если длина очищенного текста ≥ 60 . Для скольких исходных длин текста n от 1 до 200 (включительно) программа выдаёт неверный ответ?

Задание №2

В задаче обработки изображений из квадратного снимка размера $n \times n$ выбирают все окошки 3×3 с шагом 2 пикселя по горизонтали и вертикали. Это означает, что окно 3×3 всегда целиком лежит внутри снимка, а его верхний левый угол принимает все положения с нечётными координатами, для которых окно не выходит за границы. При этом окошки могут пересекаться. Известно, что всего получилось ровно 100 окошек. Какое наименьшее значение может принимать n ?

Задание №3

В очереди обработки модели компьютерного зрения 2025 картинок (нумеруются с 1). Их переставляют так: сначала идут все картинки с нечётным номером в порядке возрастания, затем все картинки с чётным номером в порядке убывания. Сколько картинок останется на своих местах?

Задание №4

ИИ удаляет из трёхзначного числа все нечётные цифры и выводит получившееся число без ведущих нулей (если все цифры удалены, выводится 0). Для скольких трёхзначных чисел результат — двузначное число?

Задание №5

Языковая модель дополняет напечатанные слова ещё до того, как они были написаны. Для этого ей нужно уметь искать слово по кратчайшему уникальному префиксу — самой короткой начальной части слова, которая не совпадает с началом ни одного другого слова из словаря. Если такого префикса короче самого слова нет, берут слово целиком. Для словаря

alpha, alps, alpine, beta, bet, better, cat, catalog, dog

найдите сумму длин кратчайших уникальных префиксов по всем словам.

Задание №6

Для каждой фотографии модель возвращает целое число s_i : чем больше s_i , тем увереннее модель, что на фото **кот**; чем меньше — тем вероятнее **собака**. Истинная метка $y_i \in \{0, 1\}$: 1 — кот, 0 — собака.

Модель не знает, y_i поэтому нам нужно выбрать один общий порог T . Модель будет предсказывать по порогу следующим образом:

$$\hat{y}_i = \begin{cases} 1, & s_i \geq T, \\ 0, & s_i < T. \end{cases}$$

Найдите минимально возможное число ошибок классификации на всех картинках и любой подходящий порог T ; при равенстве выберите *наименьший* T .

Формат входных данных.

Первая строка: целое n ($1 \leq n \leq 100$).

Следующие n строк: по два целых s_i, y_i ($|s_i| \leq 1000$, $y_i \in \{0, 1\}$). Известно, что s_i идут в порядке возрастания.

Формат выходных данных.

Два числа: минимальное число ошибок и соответствующий наименьший порог T .

Ввод

5
1 0
2 1
3 0
4 1
5 1

Вывод

1 2

Задание №7

Даны два текста: эталон и результат распознавания. Одну из простых метрик качества распознавания речи назовём *дословной точностью*:

$$\text{Точность} = \frac{\text{число совпадающих слов}}{\max\{\text{число слов в эталонном тексте, число слов в распознанном тексте}\}}.$$

Правила сравнения:

- слова в текстах разделяются пробелами;
- сравнение слов не зависит от регистра («PrIvEt» и «privet» считаются одинаковыми);
- сравнение производится по позициям, начиная с первого слова;
- если длины текстов различаются, «лишние» позиции считаются несовпадениями.

Формат входных данных.

Две строки, не более 1000 символов каждая, содержащие маленькие, большие латинские буквы и пробелы: эталонный текст и распознанный текст. Гарантируется, что в строках нет двух подряд идущих пробелов.

Формат выходных данных.

Одно число — значение метрики, округлённое до трёх знаков после запятой.

Ввод 1

Примеры. mama mila ramu
mama milom ramu

Выход 1

0.667