

Максимальное количество баллов за олимпиаду — 112

Задание 1. Пусть $k \geq 2$ — целое число. Будем говорить, что две точки на числовой прямой дружат, если разность их координат делится на k . Непустое множество точек назовём кластером, если любые две точки в этом множестве дружат, и к нему нельзя добавить ни одной точки, чтобы это свойство сохранилось.

При каких целых $k \geq 2$ точки числовой прямой с координатами

$$1, 7, 21, 22, 28, 42, 43, 49, 63$$

можно разбить на два кластера?

Задание 2. Петя загадал два натуральных числа a и b . Вася пытается их отгадать, задавая вопросы. Вопрос Васи состоит в том, что он называет два натуральных числа x и y . В ответ на вопрос Петя сообщает два числа $z < t$, одно из которых равно $ax + by$ (какое именно — не говорит).

Вася задал два вопроса: для $x = 4$ и $y = 1$ Петя ответил 14 и 15; для $x = 1$ и $y = 4$ Петя ответил 9 и 11. Какие числа загадал Петя?

Задание 3. В заповеднике планируют ввести автоматический мониторинг редких *розовых лемуров*. Для тестирования системы была выбрана группа из 100 животных, в которой 10% составляют розовые лемуры (класс 1), а оставшиеся 90% — обычные лемуры (класс 0). Будем называть эту группу *тестовой выборкой*.

Модели ИИ для каждой особи из тестовой выборки сделали предсказание, к какому классу (0 или 1) относится данный лемур. Для оценки качества распознавания введём следующие стандартные обозначения:

- TP . Предсказан класс (1), в действительности — класс (1).
- FP . Предсказан класс (1), в действительности — класс (0).
- TN . Предсказан класс (0), в действительности — класс (0).
- FN . Предсказан класс (0), в действительности — класс (1).

Всего в тестовой выборке $P = TP + FN$ объектов (лемуров) класса 1 и $N = TN + FP$ объектов класса 0.

Четыре модели ИИ (A, B, C и D) распознавали лемуров из тестовой выборки. Модель A всех лемуров отнесла к классу 0, модель B всех лемуров отнесла к классу 1. Данные о работе моделей C и D приведены в таблице ниже.

Модель	TP	FP	TN	FN
C	7	8	82	3
D	8	14	76	2

Для моделей A, B, C, D вычислим следующие пять *метрик* M_1, M_2, M_3, M_4, M_5 (если для некоторой модели возникает деление на ноль, соответствующая метрика для неё не вычисляется).

$$\begin{aligned}
 M_1 &= \frac{TP + TN}{P + N} && \text{(общая точность),} \\
 M_2 &= \frac{TP}{TP + FN} && \text{(верные на классе 1),} \\
 M_3 &= \frac{TN}{TN + FP} && \text{(верные на классе 0),} \\
 M_4 &= \frac{1}{2} (M_2 + M_3) && \text{(сбалансированная точность),} \\
 M_5 &= \frac{TP}{TP + FP} && \text{(точность на предсказаниях 1).}
 \end{aligned}$$

Для каждой из метрик $M_1, M_2, \dots, M_5$ определите, для какой из моделей достигается её максимальное значение.

Задание 4. Вася тестировал модель с действительными параметрами x и y . Он выяснил, что функция потерь задаётся формулой

$$\mathcal{L}(x, y) = x^4 + y^2 + 2x^2y + 4y + 6x^2 - 4x + 14.$$

Помогите Васе определить значения параметров (x^*, y^*) , для которых значение функции потерь наименьшее. В ответ укажите x^* , y^* и значение $\mathcal{L}(x^*, y^*)$.

Задание 5. Есть набор примеров для обучения: всего N примеров, из них ровно два — положительные, остальные — отрицательные. Случайно выбирают группу из 4 примеров. Известно, что вероятность того, что в группе окажутся оба положительных, в 2 раза больше вероятности того, что не окажется ни одного положительного. Найдите все возможные значения N .

Задание 6.

Отчёт об обучении нейросетистандартный вводстандартный вывод1 секунда256 мегабайт

Обучение нейронных сетей обычно делится на эпохи. За одну эпоху, в рамках обучения, модель один раз проходит по обучающему датасету.

Дима долгое время работал над новой разработкой. В процессе её обучения было целых n эпох. После каждой эпохи, Дима записывал текущее время в формате $HH : MM$ (например, 13 : 03). Теперь Диме интересно, какое минимальное количество времени могло уйти на обучение модели.

Дима сейчас очень загружен рабочими задачами, поэтому обратился за помощью к Вам. Посчитайте минимальное количество времени, которое могло уйти на обучение модели.

Формат входных данных

В первой строке входных данных даётся одно целое число n ($2 \leq n \leq 10^4$).

В следующих n строках задаются моменты времени, когда заканчивалась очередная эпоха в формате $HH : MM$ (24-х часовой формат)

Формат выходных данных

Выведите одно целое число: минимальное количество минут, которое могло уйти на обучение модели. Округлите ответ вниз до ближайшего целого количества минут.

Замечание

В первом тестовом примере, точно прошло 10 полных часов (14 – 24). А также, было 56 минут, начиная с 13 : 03 и до 14 : 00 (не 57, так как 13 : 03 могло быть 13 : 03 : 59). И еще 15 минут прошло с 00 : 00 до 00 : 15. Итого, $10 \cdot 60 + 56 + 15 = 671$

Задание 7.

Специальное дерево принятия решений

Имя входного файла: стандартный ввод

Имя выходного файла: стандартный вывод

Ограничения по времени: 1 секунда

Ограничения по памяти: 256 мегабайт

Дано полное бинарное дерево принятия решений глубины n . От корня до листа делается ровно n шагов. На каждом шаге принимается одно из двух решений: пойти влево или вправо. Изначально оба варианта равновероятны: у каждого ребра вероятность $1/2$.

Кто-то изменил устройство дерева и поменял две вероятности на ребрах на 0:

- первое ребро — это a -е ребро на пути «всегда влево» (ребро между уровнями $a - 1$ и a , если от корня на каждом шаге выбирать влево);
- второе ребро — это b -е ребро на пути «всегда вправо» (ребро между уровнями $b - 1$ и b , если от корня на каждом шаге выбирать вправо).

Все остальные ребра по-прежнему имеют вероятность $1/2$ (кроме тех ребер, которые лишились соседнего ребра, у них вероятность теперь равна единице).

Исходами в этом дереве называются листья (вершины на самом нижнем уровне). Вероятность каждого исхода, это произведение вероятностей на пути до соответствующего листа.

Требуется определить, сколько различных исходов (листов дерева) всё ещё имеют **ненулевую** вероятность.

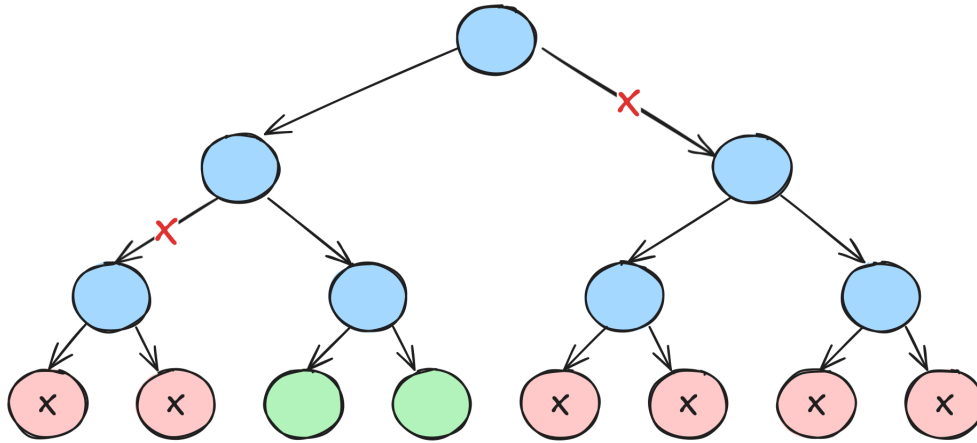
Формат входных данных

В единственной строке заданы три целых числа n , a , b ($1 \leq n \leq 60$, $1 \leq a, b \leq n$).

Формат выходных данных

Выведите одно целое число — количество листьев, которые остаются достижимыми (то есть соответствуют путям ненулевой вероятности).

Замечание



В первом тестовом примере, доступными останется лишь 2 исхода.

Задание 8. С помощью трёх моделей машинного обучения a_1, a_2, a_3 мы хотим различать фотографии собачек (класс 0) и кошечек (класс 1). Чтобы оценить качество моделей, мы запустили их на наборе изображений, для которых известен правильный ответ.

В таблице (формат XLSX формат CSV) в каждой строке записаны четыре числа: в столбце y указан правильный класс, а в столбцах a_1, a_2, a_3 — ответы трёх моделей. Все числа равны 0 или 1.

Доверие к моделям различается; веса заданы так:

$$w_1 = 2, \quad w_2 = 1, \quad w_3 = 3.$$

Для каждой строки вычисляется взвешенная сумма

$$S = w_1 \cdot a_1 + w_2 \cdot a_2 + w_3 \cdot a_3.$$

Итоговый ответ равен $\hat{y} = 1$, если $S \geq 3$, и $\hat{y} = 0$ иначе.

Найдите количество строк, в которых итоговый ответ $\hat{y}$ совпадает с правильным ответом y . В ответ укажите только это число.