

# Metric based classification

ограничение по времени на тест: 20 секунд  
ограничение по памяти на тест: 512 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Требуется построить метрический классификатор на обучающем наборе данных с известными классами и классифицировать с его помощью тестовый набор данных с неизвестными классами. Ваш классификатор должен содержать несколько метрик, сглаживающих ядер и стратегий выбора ширины окна ядра (числа ближайших соседей). Требуется выбрать оптимальную комбинацию гипер-параметров для каждого отдельного набора данных.

## Входные данные

Первая строка содержит целое число  $M$  ( $5 \leq M \leq 200$ ) — число признаков у объектов исключая класс.

Вторая строка содержит целое число  $K$  ( $2 \leq K \leq 25$ ) — число классов.

Третья строка содержит целое число  $N$  ( $50 \leq N \leq 400$ ) — число объектов в обучающем множестве.

Следующие  $N$  строк содержат описание объектов.  $i$ -тая из этих строк содержит описание  $i$ -того объекта  $M+1$  целых чисел: первые  $M$  чисел  $A_{ij}$  ( $A_{ij} \leq 10^9$ ) — признаки объекта, последнее число  $C_i$  ( $1 \leq C_i \leq K$ ) — его класс.

Следующая строка содержит целое число  $Q$  ( $50 \leq Q \leq 400$ ) — число объектов в тестовом множестве.

Следующие  $Q$  строк содержат описание объектов.  $t$ -тая из этих строк содержит описание  $t$ -того объекта:  $M$  целых чисел  $A_{tj}$  ( $|A_{tj}| \leq 10^9$ ) — признаки объекта.

## Выходные данные

Выведите  $Q$  строк. Каждая  $t$ -тая строка из них должна содержать результат классификации  $t$ -того объекта из тестового множества: целое число  $S_t$  ( $1 \leq S_t \leq 20$ ) — число соседей классифицируемого объекта, затем следует  $S_t$  пар чисел  $i$  и  $w$  ( $1 \leq i \leq N$ ,  $0 \leq w \leq 10^6$ ), где  $i$  — целое число, индекс объекта из тренировочного множества, а  $w$  — вещественное число с плавающей точкой, вес с которым учитывается этот объект. Числа  $S_t$  не обязательно должны быть одинаковыми для всех объектов.

# Система оценки

Для каждого объекта  $t$  будет рассчитан его предсказанный класс  $p_t = \operatorname{argmax}(\sum w[C_i = c])$  исходя из соответствующего множества пар  $(i, w)$ , в случае неопределённости класс выбирается псевдослучайно. Далее на основании предсказанных и реальных классов вычисляется усреднённая по классам микро F мера. Тест считается пройденным, если эта F мера будет выше определённого порога, рассчитанного с 5% запасом с помощью базового метрического классификатора.

## Пример

Входные данные	Выходные данные
2	3 1 0.75 3 0.21 4 0.21
2	3 2 0.75 3 0.65 4 0.65
4	
2 2 1	
6 2 1	
4 1 2	
4 3 2	
2	
1 2	
5 2	

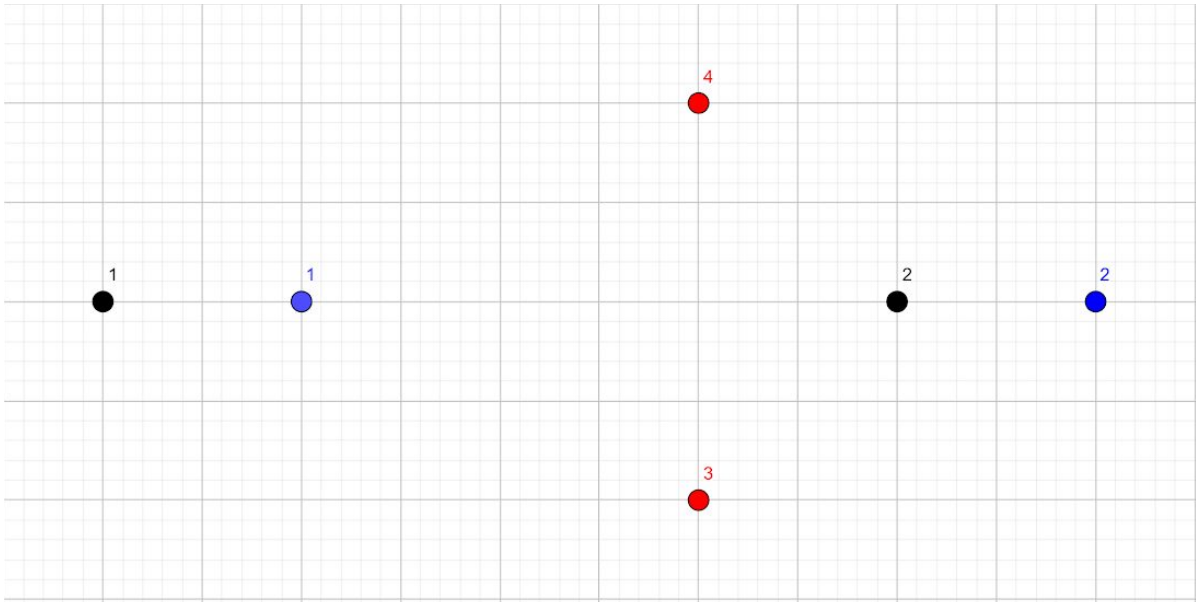


Иллюстрация примера

## Пояснение

Обратите внимание, что экземпляр задачи из данного примера не подходит под нижние ограничения на число объектов и признаков!

В данном примере в качестве метрики используется Евклидово расстояние, а в качестве взвешивающей функции Треугольное ядро. Используется три ближайших соседа с шириной окна 4.

Для первого запроса:

Номер объекта:	1	3	4
Расстояние:	1	$\sqrt{10}$	$\sqrt{10}$
Полученный вес:	0.75	0.21	0.21
Класс:	1	2	2

Суммарный вес первого класса **0.75**, второго 0.42.

Для второго запроса:

Номер объекта:	2	3	4
Расстояние:	1	$\sqrt{2}$	$\sqrt{2}$
Полученный вес:	0.75	0.65	0.65
Класс:	1	2	2

Суммарный вес первого класса 0.75, второго **1.3**.

# SVM

ограничение по времени на тест: 4 секунды  
ограничение по памяти на тест: 512 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Дан набор данных для бинарной классификации. Требуется построить разделяющие правило опирающиеся на объекты из заданного набора.

## Входные данные

Первая строка содержит целое число  $M$  ( $2 \leq M \leq 10$ ) — число признаков у объектов исключая класс.

Вторая строка содержит целое число  $N$  ( $M \leq N \leq 100$ ) — число объектов в обучающем множестве.

Следующие  $N$  строк содержат описание объектов.  $i$ -тая из этих строк содержит описание  $i$ -того объекта:  $M$  целых чисел  $D_{i,j}$  ( $|D_{i,j}| \leq 10^5$ ) — признаки объекта и его класс ('+' если объект положительный  $\text{class}(D_i) = +1$  и '-' если он отрицательный  $\text{class}(D_i) = -1$ ).

## Выходные данные

Выведите решающие правило формата:

$$\text{class}(Q) = \text{sign}((\sum \lambda_i \cdot \text{class}(D_i) \cdot f(D_i, Q)) - b)$$

В первой строке выведите симметричную функцию ядра  $f: \mathbb{R}^M \times \mathbb{R}^M \rightarrow \mathbb{R}$ , состоящую из не более чем 1000 символов. Данная функция должна удовлетворять грамматике:

- $E \rightarrow \text{pow}(E, E)$ , где  $\text{pow}(a, b) = a^b$
- $E \rightarrow \text{sub}(E, E)$ , где  $\text{sub}(a, b) = a - b$
- $E \rightarrow \text{sum}(E, E, \dots, E)$ , где  $\text{sum}(a, b, \dots, z) = a + b + \dots + z$
- $E \rightarrow \text{prod}(E, E, \dots, E)$ , где  $\text{prod}(a, b, \dots, z) = a \times b \times \dots \times z$
- $E \rightarrow A0 \mid A1 \mid \dots \mid A9$ , где  $A_i$  значение  $i$ -той координаты первого вектора ( $0 \leq i < M$ , координаты нумеруются с нуля).
- $E \rightarrow B0 \mid B1 \mid \dots \mid B9$ , где  $B_i$  значение  $i$ -той координаты второго вектора.
- $E \rightarrow$  число с плавающей точкой.

Запрещено использовать пробелы и иные символы.  $E, E, \dots, E$  — это непустое перечисление через запятую.

В следующих  $N$  строках выведите  $N$  вещественных чисел  $\lambda_i$  ( $0 \leq \lambda_i \leq 10^6$ ,  $\sum \lambda_i \cdot \text{class}(D_i) = 0$ ) — веса объектов в порядке их перечисления во входных данных.

В последней строке выведите одно вещественное число  $b$  — коэффициент сдвига.

## Система оценки

Решение будет проверено на секретном наборе данных. На основании предсказанных и реальных классов вычисляется усреднённая по классам микро F-мера. Тест считается пройденным, если эта F-мера будет выше определённого порога, рассчитанного с 5% запасом с помощью базового решения.

## Пример

Входные данные	Выходные данные
2	<code>pow(sum(0.0,prod(sub(A0,2.5),sub(B0,2.5),0.73)</code>
12	<code>,prod(sub(A1,2.5),sub(B1,2.5),0.3)),1.0)</code>
1 2 -	0.3823
1 3 -	0
1 4 -	0
2 1 +	0.3823
2 3 -	1.0
2 4 -	0
3 1 +	0
3 2 +	1.0
3 4 -	0.3823
4 1 +	0
4 2 +	0
4 3 +	0.3823
	0

# Linear regression

ограничение по времени на тест: 4 секунды  
ограничение по памяти на тест: 512 мегабайт

ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Найдите уравнения прямой аппроксимирующей положение объектов из заданного набора данных.

## Входные данные

Первая строка содержит целое число  $M$  ( $1 \leq M \leq 1000$ ) — число признаков у объектов исключая зависимую переменную.

Вторая строка содержит целое число  $N$  ( $M \leq N \leq 10000$ ) — число объектов в обучающем множестве.

Следующие  $N$  строк содержат описание объектов.  $i$ -тая из этих строк содержит описание  $i$ -того объекта,  $M$  целых чисел:  $X_{i,j}$  ( $|X_{i,j}| \leq 10^9$ ) — признаки объекта и  $Y_i$  ( $|Y_i| \leq 10^9$ ) — значение его зависимой переменной.

## Выходные данные

Выведите  $M + 1$  вещественных чисел с плавающей точкой  $A_j$  — коэффициенты прямой из уравнения  $y = a_0 \cdot x_0 + a_1 \cdot x_1 + \dots + a_{M-1} \cdot x_{M-1} + a_M$

## Система оценки

Решение будет проверено на секретном наборе данных. На основании предсказанных значений  $Y'_t$  и реальных  $Y_t$  вычисляется ошибка предсказания — нормированная сумма квадратов  $E = \sum (Y'_t - Y_t)^2 / D(Y)$ , где  $D(Y)$  — это дисперсия зависимой величины. Решение засчитывается если полученная ошибка  $E$  отличается от ошибки полученной базовым решением не более чем на 0.01

Пример

Входные данные	Выходные данные
1 4 1 0 1 2 2 2 2 4	2 -1

# Decision tree

ограничение по времени на тест: 6 секунд  
ограничение по памяти на тест: 512 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Постройте дерево принятия решений.

## Входные данные

Первая строка содержит два целых положительных числа  $M$  ( $1 \leq M \leq 100$ ) и  $K$  ( $2 \leq K \leq 20$ ) — число признаков у объектов (исключая класс) и число классов. Вторая строка содержит целое положительное число  $N$  ( $1 \leq N \leq 4000$ ) — число объектов в обучающей выборке.

Следующие  $N$  строк содержат описания объектов в обучающей выборке. В  $i$ -той из этих  $N$  строк перечислены  $M+1$  целое число: первые  $M$  чисел  $A_{i,j}$  ( $|A_{i,j}| \leq 10^9$ ) — признаки  $i$ -того объекта, последнее число  $C_i$  ( $1 \leq C_i \leq K$ ) — его класс.

## Выходные данные

Выведите построенное дерево принятия решений.

В первой строке выведите целое положительное число  $S$  ( $1 \leq S \leq 2^{11}$ ) — число вершин в дереве.

В следующих  $S$  строках выведите описание вершин дерева. В  $v$ -той из этих строк выведите описание  $v$ -той вершины:

- Если  $v$ -тая вершина **узел**, выведите через пробел: заглавную латинскую букву 'Q', целое положительное число  $f_v$  ( $1 \leq f_v \leq M$ ) — индекс признака по которому происходит проверка в данном узле, вещественное число с плавающей точкой  $b_v$  — константа с которой происходит сравнения для проверки, два целых положительных числа  $l_v$  и  $r_v$  ( $v < l_v, r_v \leq S$ ) — индекс вершины дерева в которую следует перейти, если выполняется условие  $A_i[f_v] < b_v$ , и индекс вершины дерева в которую следует перейти, если условие не выполняется.
- Если  $v$ -тая вершина **лист**, выведите через пробел: заглавную латинскую букву 'C' и целое положительное число  $D_v$  ( $1 \leq D_v \leq K$ ) — класс объекта попавшего в данный лист.



Вершины нумеруются с единицы. Корнем дерева считается первая вершина.  
Глубина дерева не должна превышать 11-ти вершин.

## Система оценки

Решение будет проверено на секретном наборе данных. На основании предсказанных и реальных классов вычисляется усреднённая по классам микро F-мера. Тест считается пройденным, если эта F-мера будет выше определённого порога, рассчитанного с 5% запасом с помощью базового решения.

## Пример

Входные данные	Выходные данные
2 4 8 1 2 1 2 1 1 3 1 2 4 2 2 3 4 3 4 3 3 1 3 4 2 4 4	7 Q 1 2.5 2 5 Q 2 2.5 3 4 C 1 C 4 Q 2 2.5 6 7 C 2 C 3

# Naive Bayes

ограничение по времени на тест: 2 секунды  
ограничение по памяти на тест: 512 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Постройте спам фильтр на основе Наивного Байесовского классификатора. Ваш классификатор должен быть строго смещён в сторону класса легитимных писем (не спама). Это значит, что он не должен отправить ни одного легитимного письма в спам.

## Входные данные

Первая строка содержит целое число  $N$  ( $2 \leq N \leq 10^5$ ) — число писем в обучающей выборке. Далее перечислены  $N$  писем из обучающей выборки в  $2 \cdot N$  строках (по две строки на письмо). Для каждого письма:

- Первая строка содержит целое число  $K$  ( $1 \leq K \leq 3 \cdot 10^5$ ) и заглавную латинскую букву — длину письма и его класс ('L' — легитимное и 'S' — спам).
- Вторая строка содержит  $K$  целых чисел  $w_j$  ( $1 \leq w_j \leq 10^6$ ) — содержание письма.

Далее следует целое число  $T$  ( $2 \leq T \leq 10^5$ ) — число писем в тестовой выборке. Далее перечислены  $T$  писем из тестовой выборки в  $2 \cdot T$  строках (по две строки на письмо). Для каждого письма:

- Первая строка содержит целое число  $K$  ( $1 \leq K \leq 3 \cdot 10^5$ ) — длину письма.
- Вторая строка содержит  $K$  целых чисел  $w_j$  ( $1 \leq w_j \leq 10^6$ ) — содержание письма.

Сумма длин всех писем из обучающей и тестовой выборки не превышает  $3 \cdot 10^6$

## Выходные данные

Выведите  $T$  заглавных латинских букв: для каждого письма из тестовой выборки в соответствующем порядке выведите его класс: 'L' — если письмо легитимное и 'S' — если оно спам.

# Система оценки

На основании предсказанных и реальных классов вычисляется точность. Тест считается пройденным, если эта точность будет выше определённого порога, рассчитанного с 5% запасом с помощью базового решения и **ни одно легитимное письмо не было помечено как спам.**

## Пример

Входные данные
2 10 L 10 11 12 13 14 15 16 17 18 19 10 S 20 21 22 23 24 25 26 27 28 29 4 10 10 11 12 13 14 15 16 17 18 19 10 10 11 12 13 14 15 26 27 28 29 10 20 21 22 23 24 25 16 17 18 19 10 20 21 22 23 24 25 26 27 28 29
Выходные данные
L L L S

# Deep Neural Network

ограничение по времени на тест: 10 секунд  
ограничение по памяти на тест: 512 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Обучите нейронную сеть с заданой архитектурой.

## Входные данные

Первая строка содержит целое положительное число  $D$  ( $2 \leq D \leq 6$ ) — глубину нейронной сети.

Следующая строка содержит  $D$  целых положительных чисел  $n_i$  ( $1 \leq n_i \leq 100$ ) — число нейронов в  $i$ -том слое. Первый слой — вход сети, его размер равен числу признаков у объекта  $M = n_1$ . Последний слой — выход сети, он задаёт бинарный класс объекта в зависимости от знака числа полученного на нём, его размер всегда равен единице  $n_D = 1$ . Гарантируется, что  $\sum (n_i + 1) \cdot (n_{i+1}) \leq 1000$ .

Следующая строка содержит целое положительное число  $T$  ( $1 \leq T \leq 1024$ ) — число объектов в тестовой выборке.

Следующие  $T$  строк содержат описания соответствующих объектов. Каждый объект задаётся  $M+1$  целым числом: первые  $M$  из этих чисел  $x_j$  ( $|x_j| \leq 10^9$ ) — признаки объекта, последнее число  $c$  — его класс ( $-1$  или  $1$ ).

## Выходные данные

Выведите  $D-1$  переход между соответствующими слоями.

Переход между слоями  $i$  и  $i+1$  описывается  $n_{i+1}$  строкой, описанием соответствующих нейронов на  $i+1$  слое. Каждый нейрон описывается строкой состоящей из  $n_i$  вещественного числа с плавающей точкой  $w_j$  и одного вещественного числа  $b$  — описание линейной зависимости текущего нейрона от выходов предыдущего  $i$ -того слоя. Линейная зависимость задается по формуле:  $y = \sum w_j \cdot x_j + b$ . Предполагается, что после каждого суммирования к его результату применяется гиперболический тангенс.

## Система оценки

Решение будет проверено на секретном наборе данных. На основании предсказанных и реальных классов вычисляется усреднённая по классам

микро F-мера. Тест считается пройденным, если эта F-мера будет выше определённого порога, рассчитанного с 5% запасом с помощью базового решения.

## Пример

Входные данные	Выходные данные
3	1.322 1.950 -0.883
2 2 1	3.301 2.078 -1.230
4	1.371 1.860 0.107
0 0 -1	
1 0 1	
0 1 1	
0 0 -1	

# Logic Expression

ограничение по времени на тест: 2 секунды  
ограничение по памяти на тест: 512 мегабайт

ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Постройте нейронную сеть по таблице истинности.

## Входные данные

Дана логическая функция  $f$ , заданная таблицей истинности.

Первая строка содержит целое число  $M$  ( $1 \leq M \leq 10$ ) — число переменных в  $f$ .

Следующие  $2^M$  строк содержат значения  $f$  в таблице истинности (0 — ложь, 1 — правда). Строки в таблице истинности последовательно отсортированы по аргументам функции от первого к последнему. Например:

$M = 1$

$f(0)$   
 $f(1)$

$M = 2$

$f(0,0)$   
 $f(1,0)$   
 $f(0,1)$   
 $f(1,1)$

$M = 3$

$f(0,0,0)$   
 $f(1,0,0)$   
 $f(0,1,0)$   
 $f(1,1,0)$   
 $f(0,0,1)$   
 $f(1,0,1)$   
 $f(0,1,1)$   
 $f(1,1,1)$

## Выходные данные

В первой строке выведите целое положительное число  $D$  ( $2 \leq D \leq 3$ ) — глубину нейронной сети. Далее выведите  $D$  целых положительных чисел  $n_i$  ( $1 \leq n_i \leq 2000$ ) — число нейронов на  $i$ -том слое. Первый слой (входной) должен содержать ровно  $M$  нейронов, а последний (выходной) — ровно один.

Далее выведите описание  $D - 1$  перехода между соответствующими слоями. Переход между слоями  $i$  и  $i+1$  описывается  $n_{i+1}$  строкой, описанием соответствующих нейронов на  $i+1$  слое. Каждый нейрон описывается строкой состоящей из  $n_i$  вещественного числа с плавающей точкой  $w_j$  и одного вещественного числа  $b$  — описание линейной зависимости текущего нейрона от выходов предыдущего  $i$ -того слоя. Линейная зависимость задается по формуле:  $y = \sum w_j x_j + b$ . Предполагается, что после каждого суммирования к его результату применяется функция ступенчатой активации, график функции

представлен на Рисунке 1. Обратите внимание, что в нуле данная функция не определена, и если в ходе вычисления вашей сети будет вызвана активация от нуля, вы получите ошибку.

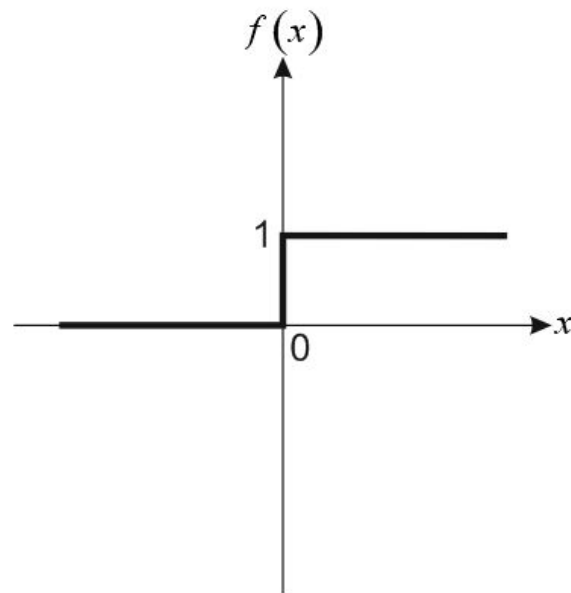


Рисунок 1. Ступенчатая функция активации.

## Примеры

Входные данные	Выходные данные	Полученная нейронная сеть
2 0 1 0 1	2 2 1 1.0 0.0 -0.5	
2 0 1 1 1 0	3 2 2 1 1.0 -1.0 -0.5 -1.0 1.0 -0.5 1 1 -0.5	

$$\chi^2$$

ограничение по времени на тест: 2 секунды  
ограничение по памяти на тест: 256 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Посчитайте зависимость двух категориальных признаков согласно критерию хи-квадрат (критерий согласия Пирсона).

## Входные данные

Первая строка содержит два целых положительных числа  $K_1$  и  $K_2$  ( $1 \leq K_1, K_2 \leq 10^5$ ) — максимальное число различных значений первого и второго признака.

Следующая строка содержит целое положительное число  $N$  ( $1 \leq N \leq 10^5$ ) — число объектов.

Следующие  $N$  строк содержат описания соответствующих объектов. Каждая из этих  $N$  строк содержит описание одного объекта: два целых положительных числа  $x_1$  и  $x_2$  ( $1 \leq x_1 \leq K_1, 1 \leq x_2 \leq K_2$ ) — значения первого и второго признака описываемого объекта.

## Выходные данные

Выведите одно вещественное число с плавающей точкой — критерий хи-квадрат зависимости двух признаков у заданных объектов.

## Пример

Входные данные	Выходные данные
2 3 5 1 2 2 1 1 1 2 2 1 3	0.833333333333333



## Пояснение

В данном примере таблицы числа наблюдений выглядит как:

реальное число наблюдений	1	2	3
1	1	1	1
2	1	1	0

ожидаемое число наблюдений	1	2	3
1	1.2	1.2	0.6
2	0.8	0.8	0.4

# Pearson correlation

ограничение по времени на тест: 2 секунды  
ограничение по памяти на тест: 256 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Посчитайте корреляцию Пирсона двух численных признаков.

## Входные данные

Первая строка содержит целое положительное число  $N$  ( $1 \leq N \leq 10^5$ ) — число объектов.

Следующие  $N$  строк содержат описания соответствующих объектов. Каждая из этих  $N$  строк содержит описание одного объекта: два целых числа  $x_1$  и  $x_2$  ( $-10^9 \leq x_1, x_2 \leq 10^9$ ) — значения первого и второго признака описываемого объекта.

## Выходные данные

Выведите одно вещественное число с плавающей точкой — корреляцию Пирсона двух признаков у заданных объектов.

## Пример

Входные данные	Выходные данные
5 1 4 2 5 3 1 4 2 5 3	-0.5

# Distance

ограничение по времени на тест: 2 секунды  
ограничение по памяти на тест: 256 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Посчитайте зависимость категориального признака  $Y$  от числового  $X$  по внутриклассовому и межклассовому расстоянию:

- Внутриклассовое расстояние =  $\sum |x_i - x_j| \cdot [y_i = y_j]$
- Межклассовое расстояние =  $\sum |x_i - x_j| \cdot [y_i \neq y_j]$

## Входные данные

Первая строка содержит одно целое положительное число  $K$  ( $1 \leq K \leq 10^5$ ) — максимальное число различных значений  $Y$  второго признака. Следующая строка содержит одно целое положительное число  $N$  ( $1 \leq N \leq 10^5$ ) — число объектов.

Следующие  $N$  строк содержат описания соответствующих объектов. Каждая из этих  $N$  строк содержит описание одного объекта: два целых числа  $x$  и  $y$  ( $|x| \leq 10^7$ ,  $1 \leq y \leq K$ ) — значения первого и второго признака описываемого объекта.

## Выходные данные

В первой строке выведите одно целое число — внутриклассовое расстояние. Во второй строке выведите одно целое число — межклассовое расстояние.

## Пример

Входные данные	Выходные данные
2 4 1 1 2 2 3 2 4 1	8 12

# Conditional entropy

ограничение по времени на тест: 2 секунды  
ограничение по памяти на тест: 256 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Посчитайте условную энтропию  $H(Y|X)$ . При расчётах используйте натуральные логарифмы  $\ln(x)$ , либо логарифмы идентичные натуральному  $\log_e(x)$ .

## Входные данные

Первая строка содержит два целых положительных числа  $K_x$  и  $K_y$  ( $1 \leq K_x, K_y \leq 10^5$ ) — максимальное число различных значений признаков  $X$  и  $Y$ .  
Следующая строка содержит целое положительное число  $N$  ( $1 \leq N \leq 10^5$ ) — число объектов.

Следующие  $N$  строк содержат описания соответствующих объектов. Каждая из этих  $N$  строк содержит описание одного объекта: два целых положительных числа  $x$  и  $y$  ( $1 \leq x \leq K_x, 1 \leq y \leq K_y$ ) — значения признаков  $X$  и  $Y$ .

## Выходные данные

Выведите одно вещественное число с плавающей точкой — условную энтропию.

## Пример

Входные данные	Выходные данные
2 3 5 1 2 2 1 1 1 2 2 1 3	0.9364262454248438

# Spearman rank correlation coefficient

ограничение по времени на тест: 2 секунды  
ограничение по памяти на тест: 256 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Посчитайте коэффициент ранговой корреляции Спирмена двух численных признаков.

## Входные данные

Первая строка содержит целое положительное число  $N$  ( $1 \leq N \leq 10^5$ ) — число объектов.

Следующие  $N$  строк содержат описания соответствующих объектов. Каждая из этих  $N$  строк содержит описание одного объекта: два целых числа  $x_1$  и  $x_2$  ( $-10^9 \leq x_1, x_2 \leq 10^9$ ) — значения первого и второго признака описываемого объекта. Гарантируется, что все значения каждого признака различны.

## Выходные данные

Выведите одно вещественное число с плавающей точкой — коэффициент ранговой корреляции Спирмена у заданных объектов.

## Пример

Входные данные	Выходные данные
5 1 16 2 25 3 1 4 4 5 9	-0.5

# Conditional variance

ограничение по времени на тест: 2 секунды  
ограничение по памяти на тест: 256 мегабайт  
ввод: стандартный ввод  
вывод: стандартный вывод

## Задача

Посчитайте условную дисперсию  $D(Y|X)$ . При расчётах используйте смещённую оценку дисперсии случайной величины.

## Входные данные

Первая строка содержит одно целое положительное число  $K$  ( $1 \leq K \leq 10^5$ ) — максимальное число различных значений признака  $X$ .

Следующая строка содержит целое положительное число  $N$  ( $1 \leq N \leq 10^5$ ) — число объектов.

Следующие  $N$  строк содержат описания соответствующих объектов. Каждая из этих  $N$  строк содержит описание одного объекта: два целых положительных числа  $x$  и  $y$  ( $1 \leq x \leq K, |y| \leq 10^9$ ) — значения признаков  $X$  и  $Y$ .

## Выходные данные

Выведите одно вещественное число с плавающей точкой — условную дисперсию.

## Пример

Входные данные	Выходные данные
2 4 1 1 2 2 2 3 1 4	1.25