

Стажировка осень-зима 2024 | машинное обучение

🕒 20 ноя 2024, 00:51:32

старт: 19 ноя 2024, 18:57:33

финиш: 20 ноя 2024, 00:57:33

до финиша: 00:05:50

...

Объявления жюри

Завершить

Задачи Посылки Сообщения

A. Lord of the Game

Ограничение времени	1 секунда
Ограничение памяти	64.0 Мб
Ввод	стандартный ввод или input.txt
Вывод	стандартный вывод или output.txt

- ☒ A. Lord of the Game
- ☐ B. Вась и котик
- ☒ C. Единичная матрица
- ☐ D. Active Learning
- ☐ E. Учимся ездить
- ☐ F. Произведения электронного искусства

Стажеры Эмси и Райд играют в любимую игру. Перед ними в ряд расположены N кучек с камнями, пронумерованными от 1 до N , в каждой кучке лежит положительное число камней. Игра заключается в следующем:

- 1) Пусть A - минимальное число камней среди всех N кучек. Эмси выбирает любую кучку, число камней в которой равно A .
- 2) Пусть B - максимальное число камней среди всех N кучек. Райд выбирает любую кучку, число камней в которой равно B .
- 3) Результат игры S - суммарное количество камней на отрезке между кучкой, выбранной Эмси, и кучкой, выбранной Райдом (включая концы).

Мальчикам интересно, какой максимальный результат игры они могут получить, если будут играть оптимально. Пожалуйста, помогите им.

Формат ввода

В первой строке содержится одно число n ($1 \leq n \leq 10^5$)

Во второй строке содержится n положительных чисел a_1, a_2, \dots, a_n ($1 \leq a_i \leq 10^9$) - количества камней в 1, 2, ..., n кучках соответственно.

Формат вывода

Для каждого набора входных данных выведите одно целое число — максимальный результат игры, который могут получить мальчики, если будут играть оптимально.

Формат ввода

В первой строке содержится одно число n ($1 \leq n \leq 10^5$)

Во второй строке содержится n положительных чисел a_1, a_2, \dots, a_n ($1 \leq a_i \leq 10^9$) - количества камней в $1, 2, \dots, n$ кучках соответственно.

Формат вывода

Для каждого набора входных данных выведите одно целое число — максимальный результат игры, который могут получить мальчики, если будут играть оптимально.

Пример 1

Ввод 	Вывод 
5 5 2 1 4 1	13

Пример 2

Ввод 	Вывод 
1 10	10

Примечания

В первом наборе входных данных минимальное число камней в кучке равно 1, максимальное число камней в кучке равно 5. Райд может выбрать только кучку под номером 1, а Эмси - кучку под номером 3 или под номером 5. Если Эмси выберет кучку 3, то результат игры $S = 5 + 2 + 1 = 8$, а если кучку 5, то $S = 5 + 2 + 1 + 4 + 1 = 13$. Максимальный результат среди всех возможных вариантов равен 13.

Во втором наборе входных данных $A = B = 10$. Эмси и Райд могут выбрать только кучку под номером 1, поэтому ответ на задачу равен 10.

Стажировка осень-зима 2024 | машинное обучение

🕒 20 ноя 2024, 00:51:53

старт: 19 ноя 2024, 18:57:33

финиш: 20 ноя 2024, 00:57:33

до финиша: 00:05:35

...

Объявления жюри

Завершить

Задачи | Посылки | Сообщения

В. Вася и котики

Ограничение времени	1 секунда
Ограничение памяти	64.0 Мб
Ввод	стандартный ввод или input.txt
Вывод	стандартный вывод или output.txt

- ☒ A. Lord of the Game
- ☒ B. Вася и котики
- ☒ C. Единичная матрица
- ☒ D. Active Learning
- ☒ E. Учимся ездить
- ☒ F. Произведения электронного искусства

Вася очень любит животных, поэтому он решил завести сразу n котов. К сожалению, его новые питомцы не сдружились друг с другом, и перед Васей встала серьезная задача рассадки котиков по лежанкам у него дома.

Дом Васи представляет собой прямую, на которой есть m лежанок с целочисленными координатами. Счастье каждого котика выражается целым числом, равным расстоянию до ближайшей лежанки, на которой есть другой котик. Также некоторые питомцы очень привередливые и хотят спать на определенной лежанке. Вася хочет рассадить котиков так, чтобы минимальное счастье среди всех питомцев было максимальным, и просит вас помочь ему найти значение этой величины при оптимальной рассадке.

Формат ввода

В первой строке содержатся числа n, m, k ($2 \leq n \leq m \leq 10^5, 0 \leq k \leq n$), количество питомцев, лежанок и уже занятых лежанок соответственно.

Во второй строке содержатся m положительных целых чисел x_1, x_2, \dots, x_m ($1 \leq x_i \leq 10^9$), соответствующих координатам лежанок. Координаты всех лежанок различны

В третьей строке содержатся k положительных целых чисел y_1, y_2, \dots, y_k ($1 \leq y_i \leq 10^9$), соответствующих координатам уже занятых лежанок. Причем $\{y_i\}_{i=1}^k \subset \{x_i\}_{i=1}^m$

Формат вывода

Выведите одно целое число, равное максимально возможной величине минимального счастья

Формат ввода

В первой строке содержатся числа n, m, k ($2 \leq n \leq m \leq 10^5, 0 \leq k \leq n$), количество питомцев, лежанок и уже занятых лежанок соответственно.

Во второй строке содержатся m положительных целых чисел x_1, x_2, \dots, x_m ($1 \leq x_i \leq 10^9$), соответствующих координатам лежанок. Координаты всех лежанок различны

В третьей строке содержатся k положительных целых чисел y_1, y_2, \dots, y_k ($1 \leq y_i \leq 10^9$), соответствующих координатам уже занятых лежанок. Причем $\{y_i\}_{i=1}^k \subset \{x_i\}_{i=1}^m$.

Формат вывода

Выведите одно целое число, равное максимально возможной величине минимального счастья среди всех питомцев.

Пример 1

Ввод 

```
2 4 1
1 5 6 9
6
```

Вывод 

```
5
```

Пример 2

Ввод 

```
3 5 0
1 7 3 4 2
```

Вывод 

```
3
```

Примечания

- 1) Координаты лежанок могут быть даны в произвольном порядке.
- 2) Обратите внимание, что k может равняться нулю

Язык Python 3.12.3 

Набрать здесь

Отправить файл

Стажировка осень-зима 2024 | машинное обучение

🕒 20 ноя 2024, 00:52:05

старт: 19 ноя 2024, 18:57:33

финиш: 20 ноя 2024, 00:57:33

до финиша: 00:05:21

...

Объявления жюри

Завершить

Задачи Посылки Сообщения

С. Единичечная матрица

Ограничение времени	1 секунда
Ограничение памяти	64.0 Мб
Ввод	стандартный ввод или input.txt
Вывод	стандартный вывод или output.txt

- ✓ A. Lord of the Game
- ✗ B. Вася и котики
- ✓ C. Единичечная матрица
- ✗ D. Active Learning
- E. Учимся ездить
- ✗ F. Произведения электронного искусства

Назовём квадратную матрицу *единичечной*, если выполняются три условия:

- сумма квадратов её элементов равна 1
- её след равен 1
- её ранг равен 1

Кирилл хорошо подумал и вывел общий вид таких матриц.

Повторите подвиг Кирилла и восстановите единичечную матрицу по её первой строке.

Формат ввода

В первой строке входа дано число n ($1 \leq n \leq 100$).

Во второй строке даны n вещественных чисел ($1 \leq n \leq 100$), соответствующих первой строке некоторой единичечной матрицы.

Гарантируется, что существует единственная единичечная матрица с такой первой строкой.

Формат вывода

Выведите сумму элементов восстановленной матрицы с точностью 10^{-3} .

Пример

Назовём квадратную матрицу *единичной*, если выполняются три условия:

- сумма квадратов её элементов равна 1
- её след равен 1
- её ранг равен 1

Кирилл хорошо подумал и вывел общий вид таких матриц.

Повторите подвиг Кирилла и восстановите единичную матрицу по её первой строке.

Формат ввода

В первой строке входа дано число n ($1 \leq n \leq 100$).


Во второй строке даны n вещественных чисел ($1 \leq n \leq 100$), соответствующих первой строке некоторой единичной матрицы.

Гарантируется, что существует единственная единичная матрица с такой первой строкой.

Формат вывода

Выведите сумму элементов восстановленной матрицы с точностью 10^{-3} .

Пример

Ввод 

Вывод 

2
0.2 0.4

1.8

Примечания

След матрицы — это сумма элементов главной диагонали матрицы.

Ранг матрицы — это максимальное число линейно независимых строк (столбцов) этой матрицы. Также это наивысший из порядков всевозможных ненулевых миноров и размерность образа линейного оператора, которому соответствует матрица.

Язык Python 3.12.3 

Набрать здесь

Отправить файл

```
1 import math
```


Стажировка осень-зима 2024 | машинное обучение

🕒 20 ноя 2024, 00:52:20

старт: 19 ноя 2024, 18:57:33

финиш: 20 ноя 2024, 00:57:33

до финиша: 00:05:09

...

Объявления жюри

Завершить

Задачи [Посылки](#) [Сообщения](#)

D. Active Learning

Ограничение времени	30 секунд
Ограничение памяти	256.0 Мб
Ввод	стандартный ввод или input.txt
Вывод	стандартный вывод или output.txt

- ✓ A. Lord of the Game
- ✗ B. Вася и котики
- ✓ C. Единичная матрица
- ✗ D. Active Learning
- E. Учимся ездить
- ✗ F. Произведения электронного искусства

Это интерактивная задача.

Успешно пройдя все этапы отбора, Кеша осуществил свою заветную мечту и попал на стажировку в свою любимую IT компанию. Так как Кеша отлично показал себя при прохождении секций, руководитель доверил ему сложную исследовательскую задачу: с помощью машинного обучения решить важную проблему бизнеса - повышение счастья новых пользователей сервиса, который развивает команда Кеши.

Задача действительно является непростой, ведь не так-то и просто формализовать понятие "счастье", а уж тем более придумать количественную меру этого самого счастья. Однако Кеше повезло - руководитель рассказал Кеше про то, что существует некоторый древний сервис, который умеет счастье измерять и оценивать, однако для некоторого фиксированного множества пользователей. Как водится в таких случаях, никто из коллег Кеши ничего не знает о внутреннем устройстве этого сервиса, код его тоже давно утрачен, однако осталась документация, описывающая интерфейс взаимодействия с этим сервисом.

Формально у Кеши есть некоторая выборка пользователей, про каждого из которых известна некоторая информация, которую Кеша намерен использовать как признаки для своей модели (разумеется, знание семантики признаков тоже утрачено). Для каждого пользователя Кеша может сделать некоторое предсказание и получить от сервиса значение производной некоторой неизвестной функции потерь, реализованной на стороне сервиса, для данного пользователя в точке, соответствующей предсказанию Кеши. Задача Кеши состоит в том, чтобы максимально приблизить неизвестные таргеты в смысле используемой функции потерь.

Кеша понимает важность поставленной задачи и хочет как можно скорее ее решить, поэтому он сам поставил себе дедлайн - сделать не более, чем некоторое фиксированное число, обращений к сервису, а дальше показать полученные предсказания руководителю, который оценит насколько предсказания Кеши действительно близки к счастью пользователей (для данных пользователей оценки счастья известны, но так как методология Кеши будет использоваться для оценки счастья новых пользователей, руководитель осознанно не

Это интерактивная задача.

Успешно пройдя все этапы отбора, Кеша осуществил свою заветную мечту и попал на стажировку в свою любимую IT компанию. Так как Кеша отлично показал себя при прохождении секций, руководитель доверил ему сложную исследовательскую задачу: с помощью машинного обучения решить важную проблему бизнеса - повышение счастья новых пользователей сервиса, который развивает команда Кеши.

Задача действительно является непростой, ведь не так-то и просто формализовать понятие "счастье", а уж тем более придумать количественную меру этого самого счастья. Однако Кеше повезло - руководитель рассказал Кеше про то, что существует некоторый древний сервис, который умеет счастье измерять и оценивать, однако для некоторого фиксированного множества пользователей. Как водится в таких случаях, никто из коллег Кеши ничего не знает о внутреннем устройстве этого сервиса, код его тоже давно утрачен, однако осталась документация, описывающая интерфейс взаимодействия с этим сервисом.

Формально у Кеши есть некоторая выборка пользователей, про каждого из которых известна некоторая информация, которую Кеша намерен использовать как признаки для своей модели (разумеется, знание семантики признаков тоже утрачено). Для каждого пользователя Кеша может сделать некоторое предсказание и получить от сервиса значение производной некоторой неизвестной функции потерь, реализованной на стороне сервиса, для данного пользователя в точке, соответствующей предсказанию Кеши. Задача Кеши состоит в том, чтобы максимально приблизить неизвестные таргеты в смысле используемой функции потерь.

Кеша понимает важность поставленной задачи и хочет как можно скорее ее решить, поэтому он сам поставил себе дедлайн - сделать не более, чем некоторое фиксированное число, обращений к сервису, а дальше показать полученные предсказания руководителю, который оценит насколько предсказания Кеши действительно близки к счастью пользователей (для данных пользователей оценки счастья известны, но так как методология Кеши будет использоваться для оценки счастья новых пользователей, руководитель осознанно не показывает их Кеше, чтобы протестировать, насколько рабочим окажется его подход).

Формат ввода

В первой строке задано три целых числа: $500 \leq N \leq 1000$, $10 \leq k \leq 20$, $10 \leq m \leq 100$ - количество объектов в выборке, число признаков и количество итераций, соответственно.

В каждой из следующих N строк содержится k чисел: i -тая строка соответствует признаковому описанию i -того объекта выборки.

Формат вывода

Должно произойти m итераций вывода предсказаний для каждого объекта выборки. Более формально, вы должны вывести m строк, каждая из которых содержит N чисел, где i -тое число - прогноз модели на i -том объекте.

После вывода каждой из первых $m - 1$ строк, выводится строка, содержащая N чисел, где i -тое число равно значению производной неизвестной вам функции потерь на i -том объекте в

Формат ввода

В первой строке задано три целых числа: $500 \leq N \leq 1000$, $10 \leq k \leq 20$, $10 \leq m \leq 100$ - количество объектов в выборке, число признаков и количество итераций, соответственно.

В каждой из следующих N строк содержится k чисел: i -тая строка соответствует признаковому описанию i -того объекта выборки.

Формат вывода

Должно произойти m итераций вывода предсказаний для каждого объекта выборки. Более формально, вы должны вывести m строк, каждая из которых содержит N чисел, где i -тое число - прогноз модели на i -том объекте.

После вывода каждой из первых $m - 1$ строк, выводится строка, содержащая N чисел, где i -тое число равно значению производной неизвестной вам функции потерь на i -том объекте в точке, соответствующей данному предсказанию.

После вывода последней строки, вы не получаете никакой обратной связи. Эти предсказания используются для того, чтобы замерить значение неизвестной функции потерь. Решение считается верным, если данное значение окажется меньше некоторого порога, то есть ваши предсказания действительно хорошо приближают таргет в смысле данной функции потерь.

В конце каждой строки с предсказаниями не забудьте сделать вывести перевод строки и сбросить буфер вывода. Для сброса буфера используйте:

- `flush(stdout)` или `cout.flush()` в C++;
- `System.out.flush()` в Java;
- `flush(output)` в Pascal;
- `stdout.flush()` в Python;
- смотрите документацию для других языков.

Примечания

Конечно, справиться с задачей для Кеши не составило труда, ведь он готовился к собеседованиям и хорошо знал устройство алгоритма градиентного бустинга, реализация которого удивительным образом похожа на поставленную перед ним задачу (напомним, что в случае градиентного бустинга итоговая модель получается как композиция базовых моделей, каждая из которых приближает антиградиент функционала ошибки в точке, соответствующей сумме предсказаний базовых моделей, полученных на предыдущей итерации).

Стажировка осень-зима 2024 | машинное обучение

🕒 20 ноя 2024, 00:52:48

старт: 19 ноя 2024, 18:57:33

финиш: 20 ноя 2024, 00:57:33

до финиша: 00:04:41

...

Объявления жюри

Завершить

Задачи Посылки Сообщения

Е. Учимся ездить

Николай всегда хотел водить автомобиль, однако, сдать на права у него никак не получается, то погода дождливая, то инструктор очень строгий. Наш герой подумал-подумал и решил, что можно ведь научить машину ездить без водителя! Николай побродил по просторам интернета, нашел небольшой датасет и решил попробовать свои силы на нем.

По сути датасет содержит следующие данные: сцены с набором векторов единичной длины, которые описывают собой дорожное полотно (полосы движения). Сцена - по сути "замороженная" дорожная ситуация, для которой мы знаем, как нужно ее проехать, если нас вдруг "разморозят". Одна строка данных соответствует одной сцене.

Рассмотрим его компоненты:

1. x, y - координаты исходной точки вектора дорожного полотна.
2. dir_x, dir_y - компоненты единичного вектора, прибавив их к исходным координатам можно получить вторую точку вектора.
3. $types$ - типы каждой точки. 15 - вектор границы полосы, 16 - вектор центра полосы.
4. $target_point$ - содержит координату точки маршрута, которая примерно указывает направление, в которое мы должны приехать.
5. $target_x, target_y$ - непосредственно точки, отображающие траекторию, по которой надо проехать данный участок. Первый элемент траектории соответствует текущему положению автономного автомобиля, далее каждая следующая точка отображает положение через 0.1 секунду в будущем. Это то, что мы должны научиться предсказывать. Все вышеуказанные координаты приведены в метрах.

Чтобы замерять качество построенной траектории, Николай выбрал следующую метрику (формула для одной сцены):

$$ADE(target_x, target_y, pred_x, pred_y) = 1/T * \sum_{i=0}^{T-1} \sqrt{(target_x_i - pred_x_i)^2 + (target_y_i - pred_y_i)^2}, \text{ где } T \text{ количество точек}$$

траектории, а $pred_x, pred_y$ - наше предсказание. Данная метрика усредняется по всем сценам.

Сможет ли Николай обучить такую модель, чтобы ее ADE было не больше 3 на тестовой выборке?

- ✓ A. Lord of the Game
- ✗ B. Вась и котик
- ✓ C. Единичная матрица
- ✗ D. Active Learning
- ✓ E. Учимся ездить
- ✗ F. Произведения электронного искусства

$$\sum_{i=0}^{T-1} \sqrt{(target_x_i - pred_x_i)^2 + (target_y_i - pred_y_i)^2}$$
, где T количество точек траектории, а pred_x, pred_y - наше предсказание. Данная метрика усредняется по всем сценам.

Сможет ли Николай обучить такую модель, чтобы ее ADE было не больше 3 на тестовой выборке?

Формат ввода

В качестве входных данных подается [архив](#), который содержит два файла: train.csv, test.csv.

Оба файла содержат данные в следующем формате:

1. x, y, dir_x, dir_y, types - каждая строка это список длины k, который может отличаться от строки k строки. k - это количество векторов дорожного графа.
2. target_point - каждая строка это список длины 2, содержащий координаты x, y.
3. target_x, target_y (только для train.csv) - каждая строка содержит список длины 46, содержащий соответствующую координату.

Формат вывода

Ответ должен содержать столько же строк сколько в файле test.csv и содержать траектории в таком же формате, как они были представлены в train.csv, только вместо [target_x, target_y] колонки pred_x, pred_y.

Пример:

```
"[3320.6877 3320.77375 3320.8598 3320.94585 3321.0319 3321.11795 3321.204
3321.29005 3321.3761 3321.46215 3321.5482 3321.63425 3321.7203 3321.80635
3321.8924 3321.97845 3322.0645 3322.15055 3322.2366 3322.32265 3322.4087
3322.49475 3322.5808 3322.5808 3321.80643182 3321.03206364 3320.25769545
3319.48332727 3318.70895909 3317.93459091 3317.16022273 3316.38585455
3315.61148636 3314.83711818 3314.06275 3313.28838182 3312.51401364 3311.73964545
3310.96527727 3310.19090909 3309.41654091 3308.64217273 3307.86780455
3307.09343636 3306.31906818 3305.5447]"
```


Количество пробелов между координатами может быть произвольным. Количество знаков после запятой также может быть произвольным.

Примечания

Тест4

Выбрать Файл не выбран

Отправить

 осталось 100 попыток

Стажировка осень-зима 2024 | машинное обучение

🕒 20 ноя 2024, 00:52:59

старт: 19 ноя 2024, 18:57:33

финиш: 20 ноя 2024, 00:57:33

до финиша: 00:04:31

...

Объявления жюри

Завершить

Задачи Посылки Сообщения

Ф. Произведения электронного искусства

Ограничение времени	1 секунда
Ограничение памяти	64.0 Мб
Ввод	test.csv
Вывод	answers.csv

- ☒ A. Lord of the Game
- ☐ B. Вася и котики
- ☒ C. Единичная матрица
- ☐ D. Active Learning
- ☐ E. Учимся ездить
- ☐ F. Произведения электронного искусства

Олег любит создавать произведения электронного искусства на компьютере, а затем их продавать по интернету. Его деяния настолько оказались популярны и востребованы в сети, что за рецептом его успеха стали гоняться многие конкуренты. После хороших выходных на природе, Олег сел за работу и обнаружил, что его алгоритм, создающий его творения — испорчен трояном. Олег — творец и художник, но восстановить данные самостоятельно он не способен, поэтому он обратился за помощью к вам, чтобы решить эту непростую задачу.

Доподлинно известно, что алгоритм Олега состоит из трех частей, каждая из которых применяется по очереди. На вход каждому алгоритму подается картинка, а на выходе получается матрица чисел. Каждая часть алгоритма работает одинаково, но имеет разные коэффициенты обработки. Сам алгоритм работает следующим образом: берется матрица чисел той же размерности, что и коэффициенты обработки, элементы матриц перемножаются и складываются, а затем результат операции помещается на выходную матрицу ровно в то место, где лежит центр взятой матрицы.

К счастью Олега, вирус съел не всё. Известно, что алгоритмов было всего три: два из них остались целыми и невредимыми. А также осталась коллекция, которую Олег ещё не успел продать. Но что очень грустно, так это то, что вирус съел информацию об одном алгоритме и порядке применения алгоритмов. **Ваша задача** — восстановить утерянный алгоритм, а также порядок применения алгоритмов.

Формат ввода

Данные будут предоставлены в архиве:

- Все входные картинки находятся в формате PNG. Гарантируется, что все изображения

К счастью Олега, вирус съел не все. Известно, что алгоритмов было всего три, два из них остались целыми и невредимыми. А также осталась коллекция, которую Олег ещё не успел продать. Но что очень грустно, так это то, что вирус съел информацию об одном алгоритме и порядке применения алгоритмов. **Ваша задача** — восстановить утерянный алгоритм, а также порядок применения алгоритмов.

Формат ввода

Данные будут предоставлены в архиве:

- Все входные картинки находятся в формате PNG. Гарантируется, что все изображения имеют 1 канал.
- Сохраненная коллекция произведений Олега — это текстовые файлы массивов чисел. Для каждой входной картинки — один текстовый файл с соответствующим именем. Файлы были сохранены функцией `numpy.savetxt("", arr)`.

Также в архиве будет лежать файл `algos.csv`, где будут в произвольном порядке лежать данные для алгоритмов: в первой строке для одного, во второй строке — для второго. В каждой строке по 9 чисел — линейное представление матрицы 3×3 .

Датасет можно скачать по [ссылке](#).

Формат вывода

Сохраните два известных и один восстановленный алгоритм в порядке применения в csv файл. Ожидается 3 строки по 9 чисел в каждой, каждое число разделено запятой.

Примечания

В задаче не ожидается точный ответ. Каждый ответ будет оцениваться как среднее между тремя MSE по каждой матрице коэффициентов алгоритмов.

Обратите внимание, что на задачу даётся только 10 попыток.

Язык (make) yandexdataschool

Выбрать Файл не выбран

Отправить 1 осталось 6 попыток

Предыдущая

Время отправки	ID	Задача	Компилятор	Вердикт	Тип отправки	Время	Память	Тест	Баллы
20 ноя 2024, 00:39:57	125684934	F	(make) yandexdataschool	WA	-	449ms	39.88Mb	1	-

[отчёт](#)