A. Что случилось?

Необходимо восстановить коэффициенты функции f(x), зная её значения на некотором наборе точек.

При этом известно, чтоf(x)=((a+da)ex+(b+db)x3∕4)2+(c+dc)cos2x,

где di – случайные величины, которые принимают значения из отрезка [−0.001,0.001]; a, b, c – неизвестные вещественные константы, которые требуется найти (абсолютная ошибка не должна превышать 10−2).

Формат ввода

В каждой строке файла <https://yadi.sk/d/llKlDyAK982aLQ> записаны два числа x и f(x), разделённые пробелом.

Формат вывода

Выведите через пробел 3 вещественных числа с точностью не менее 2 значащих цифр после десятичной точки, которые соответствуют набору a, b, c.

B. Где все клиенты?

Одна из проблем в сфере услуг – отток клиентов. Многие клиекты покупают абонементы, но перестают посещать заведения, и отменяют их до истечения срока. Аднимистрация фитнес-центра собирается понять, что заставляет пользователей уйти, и выработать стратегию по работе с клиентами. Для борьбы с оттоком собрана информация о клиетах.

Данные можно скачать по ссылке <https://yadi.sk/d/Fr6homfYTsakFw>.

**Данные**

* Row – (int) индекс строки данных
* Registration – (int) уникальный номер клиента
* Zipcode – (int) индекс клиента
* Age – (int) возраст
* Partner\_company – (0/1) продан ли абонемент по партнерской программе
* Friend\_promo – (0/1) продан ли абонемент по рекомендации друга
* Contract\_period – (int) продолжительность подписки в месяцах
* Lifetime – (int) продожительность истории посещений
* Class\_registration\_weekly – (ﬂoat) среднее число занятий в неделю
* Avg\_additional\_charges\_total – (ﬂoat) средняя плата за дополнительные услуги
* Cancellation\_freq – (ﬂoat) (ﬂoat) среднее число отмен и пропусков занятий
* **Exited**– (0/1) анулирован ли абонемент

Изучите профили клиентов и проанализируйте, почему они отказываются от услуг. Составьте прогноз для разных групп клиентов: тех, кто может уйти, и тех, кто останется.

Вам нужно составить прогноз для каждого конкретного клиента.

В вашем распоряжении есть также дополнительный список клиентов, которые выражали такие же недовольства. На основе их данных нужно предсказать, кто из них уйдет, а кто останется. Новые данные для прогнозирования находятся в файле ниже.

Формат ввода

Данные можно скачать по ссылке <https://yadi.sk/d/myTXgFfCsRRqoQ>.

Формат вывода

**gym\_test.csv**с добавленной колонкой Exited, заполненной значениями предсказаний 0 или 1.

**Замечание:**Не изменяйте порядок столбцов в таблице.

Примечания

Вам не обязательно использовать специфические инструменты анализа данных, но вы можете.

**Подсказка:**Попробуйте найти правила по размеченной выборке и используйте эти эвристики для решения.

Решение будет зачтено, если значение F1-меры будет не менее 0.9.

C. Лента рекомендаций

|  |  |
| --- | --- |
| Ограничение времени | 2 секунды |
| Ограничение памяти | 64Mb |
| Ввод | стандартный ввод или input.txt |
| Вывод | стандартный вывод или output.txt |

Разработчик Илья создал социальную сеть с лентой рекомендаций разнородного контента. В ленте перемешаны различные объекты (фото, новости, видео и т.д.). Илья настроил ленту таким образом, что объекты упорядочиваются по релевантности. Чем больше лайков набирает объект, тем он ближе к началу поиска. Илья не предусмотрел одно важное условие, что при таком упорядочивании в списке рекомендаций часто встречается несколько объектов одного типа подряд.

Это не нравится пользователям, потому что рекомендации выглядят однообразными. Необходимо реализовать алгоритм, который по списку рекомендаций составит новый список – лишённый этого недостатка и при этом максимально релевантный.

Пусть задан исходный список рекомендаций a=[a0,a1,…,an−1] длиной n>0. Объект под номером i имеет тип с номером bi∈{0,…,m−1} и релевантность r(ai)=2−i.

Рассмотрим список, который получается из исходного выбором подмножества объектов и их перестановкой: x=[ai0,ai1,…,aik−1] длины k (1≤k≤n). Список называется допустимым, если никакие два подряд идущих объекта в нем не совпадают по типу, т.е. bij≠bij+1 для всех j=0,…,k−2. Релевантность списка вычисляется по формуле ∑k−1j=02−jr(aij). Вам нужно найти максимальный по релевантности список среди всех допустимых.

Формат ввода

В первой строке через пробел записаны числа n и m (1≤n≤100000, 1≤m≤n).

В следующих n строках записаны числа bi для i=0,…,n−1 (0≤bi≤m−1).

Формат вывода

Выпишите через пробел номера объектов итогового списка: i0,i1,…,ik−1.

D. Пандемия и карантин

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Язык | Ограничение времени | Ограничение памяти | Ввод | Вывод |
| Все языки | 2 секунды | 500Kb | стандартный ввод | стандартный вывод |
| python3.6+numpy+pandas | 5 секунд | 9Mb |
| (make) python2.7+numpy | 5 секунд | 9Mb |
| (make) python3.5+numpy | 5 секунд | 9Mb |

На фоне пандемии жизнь миллионов людей кардинально изменилась – многим пришлось отказаться от привычных вещей, меньше общаться и посещать общественные мероприятия. Вместе с изменением уклада жизни поменялись и запросы пользователей в Яндекс.Поиске, которые поступают десятками тысяч в секунду.

Во время карантина в России вырос спрос на ряд товаров и услуг: на онлайн-концерты и онлайн-экскурсии по городам и музеям, на товары для занятия спортом и здоровое питание, на онлайн курсы, на товары для дачи и ремонта, на доставку на дом. Подобные вопросы задавали сотни раз в час.

На фоне отмены и переноса массовых мероприятий все реже стали повторяться запросы про покупку билетов в кинотеатры, обучение в автошколе и др.

С введением режима самоизоляции у людей появилось больше свободного времени, которое многие тратят на то, о чем раньше не задумывались, поэтому какие-то вопросы стали задавать в поисковик впервые.

У вас есть 500 KB оперативной памяти, оцените количество уникальных запросов. Правильный ответ не превосходит 100000 и не меньше, чем 50000.

Мы засчитаем решение, если ваш ответ будет отличаться не более, чем на 5%.

Формат ввода

В первой строке указано число n≤500000 — количество запросов, среди которых нужно найти число уникальных. В каждой из n последующих строк содержится по одному запросу. Длина каждого запроса не превосходит 1000 символов.

Формат вывода

Необходимо вывести одно число — оценку количества уникальных запросов. Оценка не обязана быть целой.