

Домашнее задание

Это домашнее задание поможет вам попрактиковаться в применении моделей атрибуции для планирования и оптимизации будущих рекламных кампаний.

Предположим, что у вас есть 1 млн рублей на будущую рекламную кампанию. Помимо этого, у вас есть данные об эффективности различных каналов рекламы в виде таблицы с визитами пользователей. Эта таблица содержит идентификатор пользователя, время визита, затраты на каждый визит и сумму выручки, которую пользователь принес за каждый визит. Визит не обязательно заканчивается покупкой, иногда выручка может быть равна нулю.

Указанная таблица находится в файле data.csv.

Это задание настоятельно рекомендуется выполнять в jupyter notebook с помощью python3.

Чтобы успешно выполнить задание, нужно в attribution_modeling_part2.ipynb

1. загрузить данные и провести первичные расчеты (вы уже разбирали их в предыдущей практике) до пункта “Моделирование оптимальных бюджетов”
2. Начиная с пункта “Моделирование оптимальных бюджетов”, вам нужно запускать ячейки ноутбука, но в некоторых из них предварительно записывать формулы для вычисления различных метрик

Предполагается, что для вычислений вы будете использовать функционал библиотеки pandas. Если нужно освежить знания о том, как работает эта библиотека, воспользуйтесь официальной документацией, например на этой странице:

- https://pandas.pydata.org/pandas-docs/stable/getting_started/10min.html

Решением домашнего задания считаются ответы на следующие вопросы:

- Сколько прибыли получим от использования линейной модели?
- Сколько прибыли получим от использования модели time-decay?
- За счет каких каналов линейная модель будет приносить больше выручки, а за счет каких меньше?
- За счет каких каналов модель time-decay будет приносить больше выручки, а за счет каких меньше?
- Как вы думаете, почему линейная модель и модель time-decay показали такие разные результаты?

Числовые ответы округляйте (по стандартным правилам округления) до целого.

Если нужно перечислить какие-то каналы рекламы, то давайте полный ответ, весь список подходящих каналов.

На открытый вопрос отвечайте свободно, но постарайтесь свести ответ к основным факторам различия между моделями.

Для успешного выполнения домашнего задания нужно ответить правильно на все вопросы.

Немного более подробная инструкция по выполнению задания приведена ниже:

1. Выполним все ячейки с первой и до той, в которой находится заголовок “Моделирование оптимальных бюджетов”
2. Зададим наш рекламный бюджет

Моделирование оптимальных бюджетов

```
In [12]: 1 NEW_BUDGET = 1e6
```

3. Сделаем отдельную табличку, которую будем использовать для расчетов, а также масштабируем наши рекламные бюджеты пропорционально предыдущему их распределению между каналами. В результате суммарно получим бюджет в 1 млн рублей:

```
In [11]: 1 new_totals = totals.copy()  
2 new_totals['new_budget'] = NEW_BUDGET * new_totals['cost'] / new_totals['cost'].sum()
```

4. Посчитаем результаты для линейной модели. Помните, что в некоторых строках нужно домножать на 0.01, чтобы перевести ROI из процентов в доли:
 - a. Для этого сначала выберем нужные для расчета колонки (первая строчка на скриншоте ниже)
 - b. Затем, посчитаем, сколько денег выручки мы бы получили, если бы использовали линейную модель и старое распределение бюджета
 - c. Потом посчитаем дельту между ROI
 - d. Рассчитаем оптимальный бюджет на основе разницы между ROI и старым бюджетом
 - e. Посчитаем, сколько выручки получим при новом бюджете и линейной модели атрибуции
 - f. Посчитаем прибыль от нашей оптимизации – это разность между оптимизированной выручкой и нашей текущей выручкой
 - g. Выведем результаты на экран:

```
linearModel = new_totals[['new_budget', 'lastTouchROI', 'linearROI']]  
  
# Заполните строчки ниже вашими расчетами:  
linearModel['currentValue'] = # ваш код  
linearModel['delta'] = # ваш код  
linearModel['optimalSpend'] = # ваш код  
linearModel['optimalValue'] = # ваш код  
  
linearModel['profit'] = linearModel['optimalValue'] - linearModel['currentValue']  
  
print(f'Суммарный профит от линейной модели = {sum(linearModel["profit"])}')  
  
linearModel
```

5. Аналогичные расчеты сделайте также и в следующей ячейке:

```
tdModel = new_totals[['new_budget', 'lastTouchROI', 'timeDecayROI']]

# Заполните строчки ниже вашими расчетами:
tdModel['currentValue'] = # ваш код
tdModel['delta'] = # ваш код
tdModel['optimalSpend'] = # ваш код
tdModel['optimalValue'] = # ваш код

tdModel['profit'] = tdModel['optimalValue'] - tdModel['currentValue']

print(f'Суммарный профит от time-decay модели = {sum(tdModel["profit"])}')

tdModel
```