

Кейс: рекомендации официантам



С чего все началось

К вам пришёл владелец крупной сети ресторанов:

— Мы хотим предлагать клиентам дополнительные блюда в конце заказа и за счёт этого увеличить прибыль. Можете построить систему рекомендаций блюд?

Первые шаги



A photograph of a young baby, seen from behind, taking its first steps on a dirt path. The baby is wearing a blue and white striped long-sleeved shirt and blue jeans. The background is filled with green foliage and trees.

Первые шаги

- Задавайте произвольные вопросы
- Поймите, что от вас хотят
- Узнайте в чём проблема

A photograph of a young baby, seen from behind, taking its first steps on a dirt path. The baby is wearing a blue and white striped long-sleeved shirt and blue jeans. The background is filled with green foliage and trees.

Первые шаги

- Как будет использоваться система
- Какие доступны данные для обучения
- Какие данные будут доступны на момент предсказания
- Есть ли дополнительные ограничения

A photograph of a young baby with light brown hair, seen from behind, walking away on a dirt path. The baby is wearing a blue and white striped long-sleeved shirt and dark blue jeans. The background is filled with green foliage and trees.

Доступные данные

- Вся информация о чеках
- Вся информация о ресторанах
- В момент заказа вы знаете только содержимое чека и время

Как это можно использовать?



Первая модель



Первая модель

- Какую величину можно было бы предсказывать
- Как измерить качество на исторических данных
- Как обучить модель предсказывать эту величину



Первая модель: что предсказывать

1. Берём чек
2. Выкидываем одну позицию из него
3. Предсказываем, что именно было выкинуто



Первая модель: что предсказывать

1. Берём чек
2. Выкидываем одну позицию из него
3. Предсказываем, что именно было выкинуто

В чём плюсы и минусы такого подхода?



Первая модель: оценка качества

- Оцениваем не конкретное решение
- Метрика должна быть общей для произвольного алгоритма!
- Лучшим методом не обязан оказаться ML



Первая модель: оценка качества

- Метрика качества – $\text{recall}@k$
- Что это такое? Какое k выбрать?



Первая модель: как обучить

- Какие есть варианты?
- Страйтесь называть сразу и плюсы и минусы

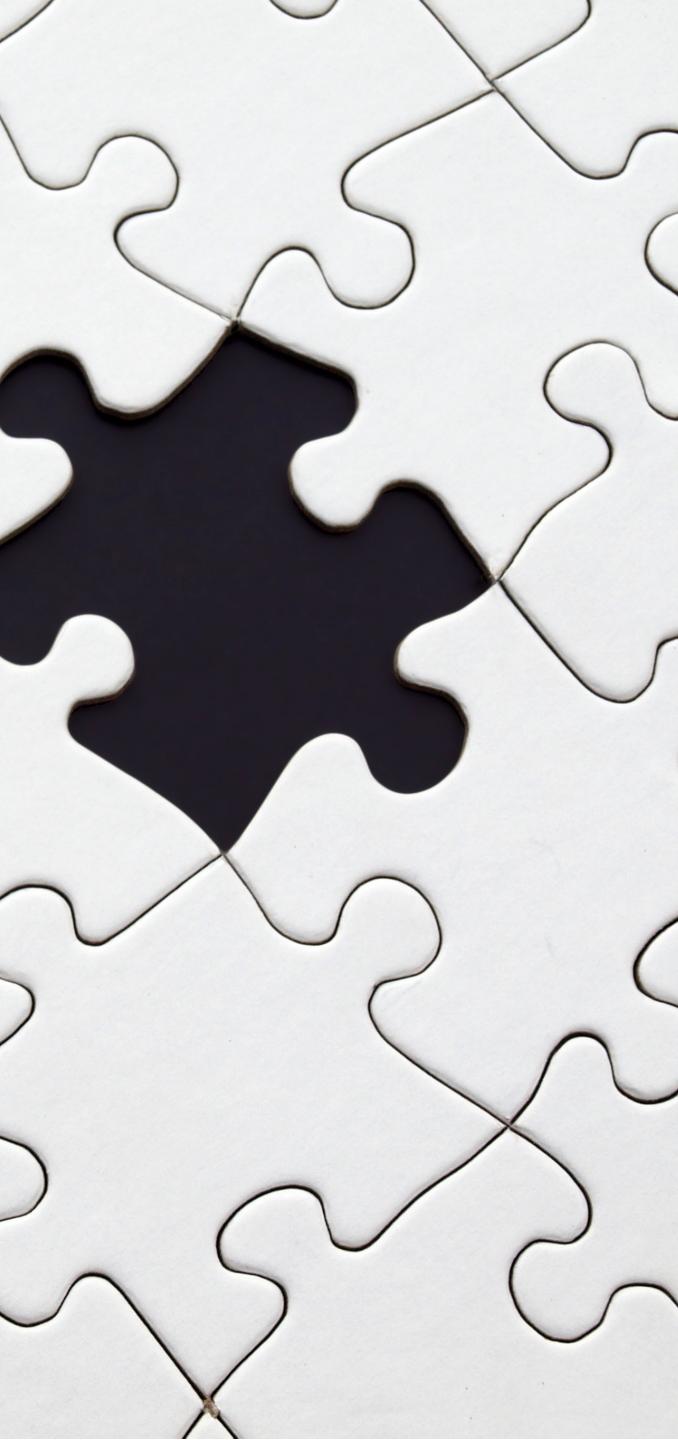
```
In [1]: import sklearn
```

```
-----  
ImportError                                     Traceback (most recent call last)  
<ipython-input-1-8fd979e02004> in <module>()  
----> 1 import sklearn  
  
ImportError: No module named sklearn
```



Первая модель: как обучить

- По паре (чек, товар) предсказываем, был ли этот товар скрыт
- Много 0, используем негативное сэмплирование
- Как именно сэмплировать – параметр задачи



Подводные камни

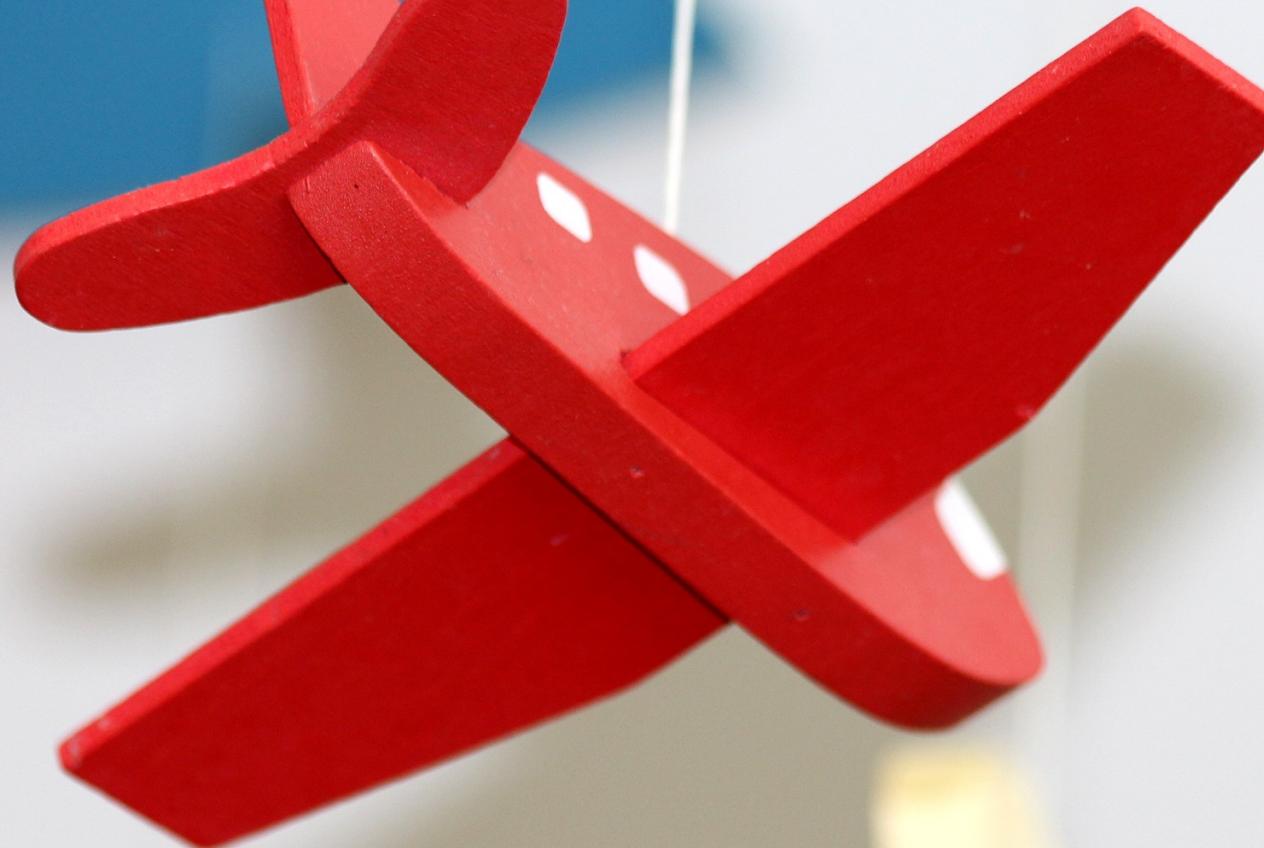
- Конечное решение принимает официант
- Оптимизируете не деньги
- Какие ещё проблемы вы видите?



Онлайн эксперимент

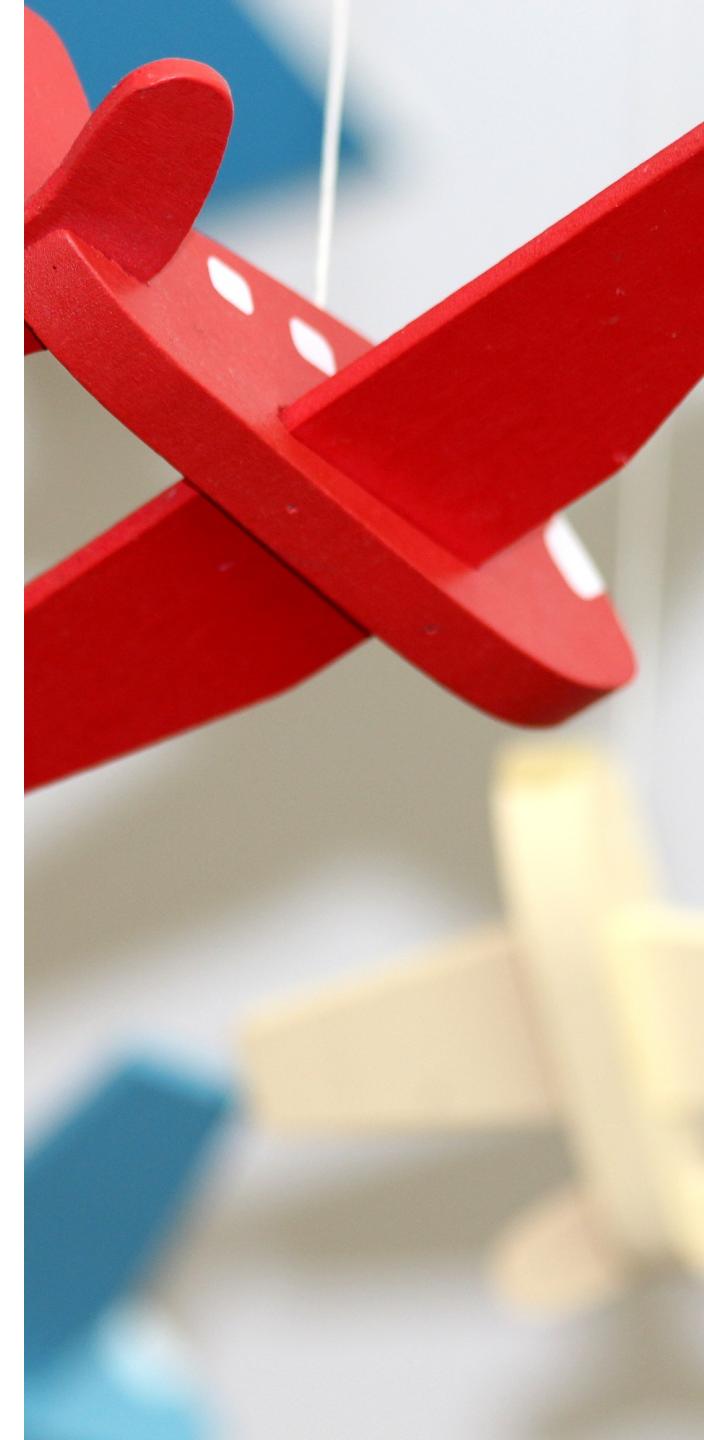
- Не будем останавливаться подробно
- Опишите примерный дизайн эксперимента

Вторая модель



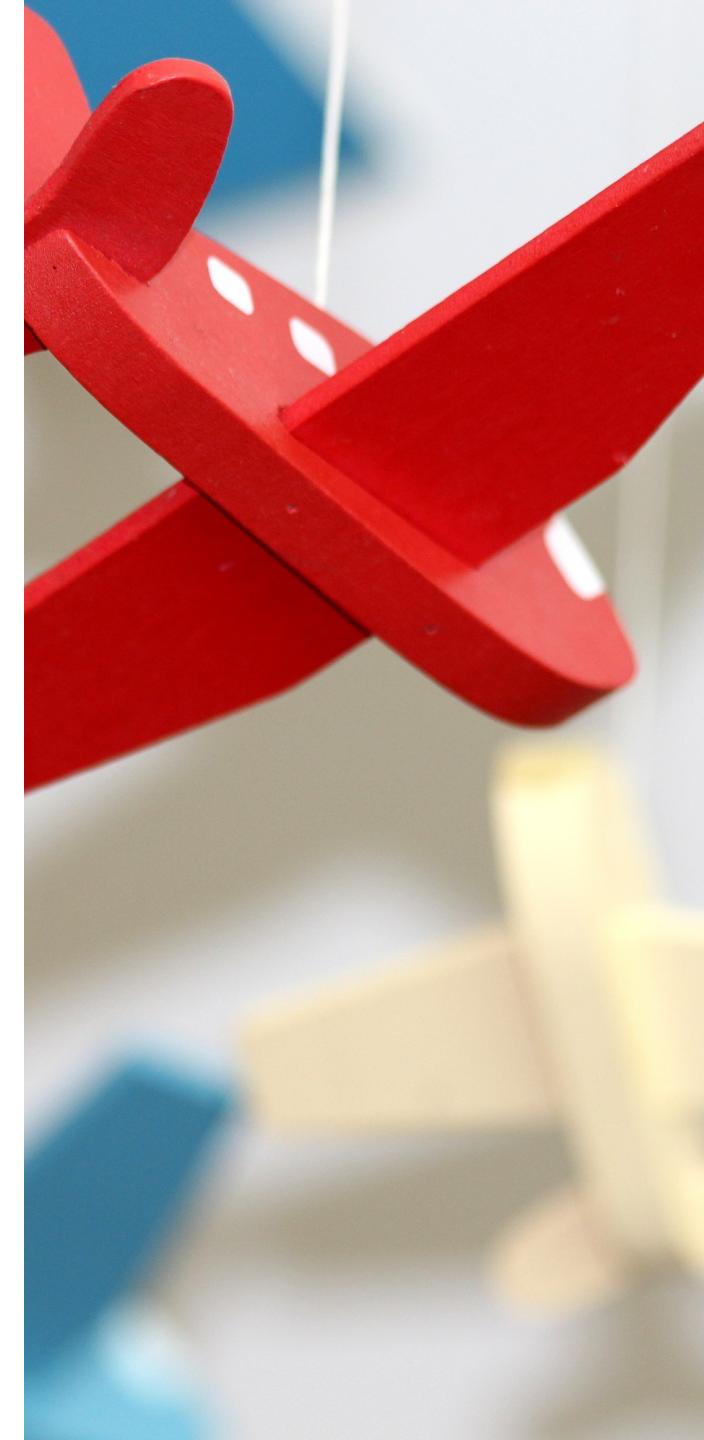
Вторая модель

- Допустим, вы сможете провести эксперимент, чтобы собрать данные
- Как вы измените дизайн эксперимента?
- Теперь цель – собрать данные, а не только улучшить продажи



Два этапа: сбор данных

- Что именно вы хотите собрать?
- Как добиться максимального качества собираемых данных?
- Как скорректировать первую модель, чтобы этого добиться?



Два этапа: новая модель

- Какую модель обучить на новых данных?
- Как оценить её качество?
- Как спроектировать новый АБ тест?
- Предлагаем вам обсудить эти вопросы
совместно в чате

