

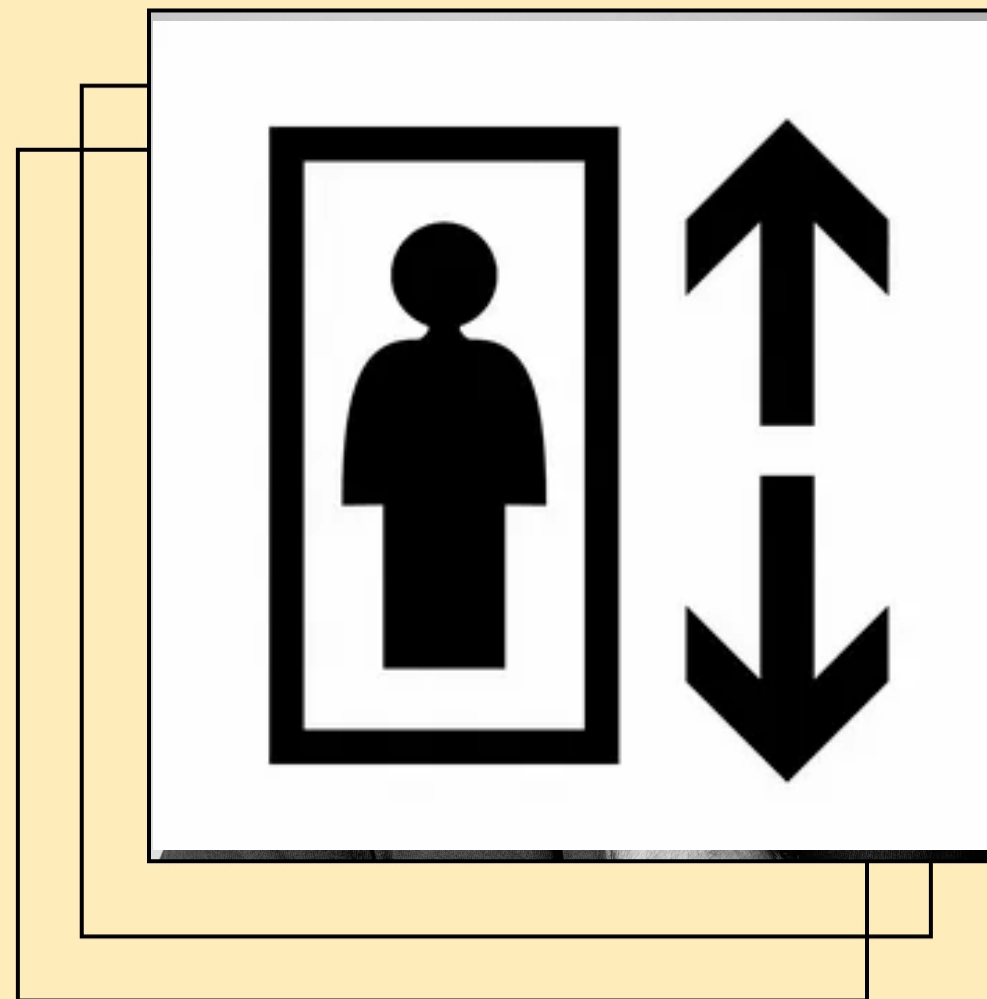
# Индустриальный семинар

## Uplift моделирование

Элен Теванян

Руководитель направления алгоритмического анализа, X5 Tech

[x5retail.tech](https://x5retail.tech)



# Типичный запуск

# Типичный запуск

## Примеры офферов

- Получите 200 баллов при покупке от 400 руб.
- Вам (пред-) начислено 300 баллов. Успеите списать в течение следующей недели
- В 3 раза больше баллов в категории “Сосиски, сардельки”

# Типичный запуск

## Примеры офферов

- Получите 200 баллов при покупке от 400 руб.
- Вам (пред-) начислено 300 баллов. Успейте списать в течение следующей недели
- В 3 раза больше баллов в категории “Сосиски, сардельки”

## Примеры метрик

- Выручка
- Количество активаций
- Маржа
- Etc

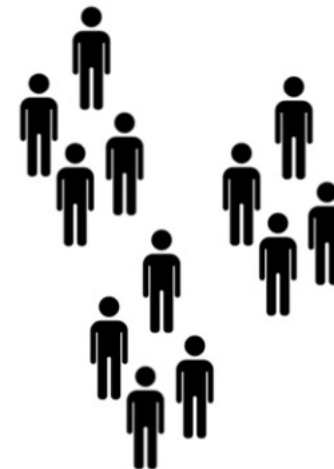
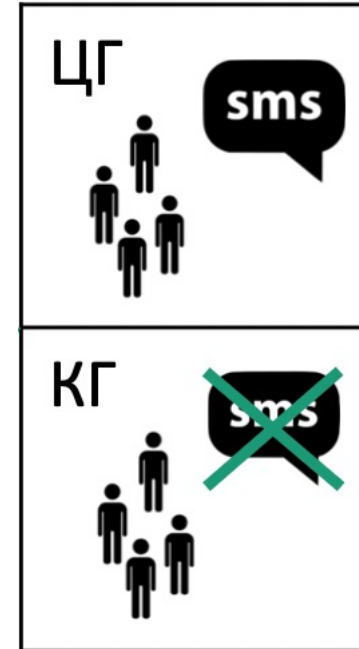
# Типичный запуск

## Примеры офферов

- Получите 200 баллов при покупке от 400 руб.
- Вам (пред-) начислено 300 баллов. Успейте списать в течение следующей недели
- В 3 раза больше баллов в категории “Сосиски, сардельки”

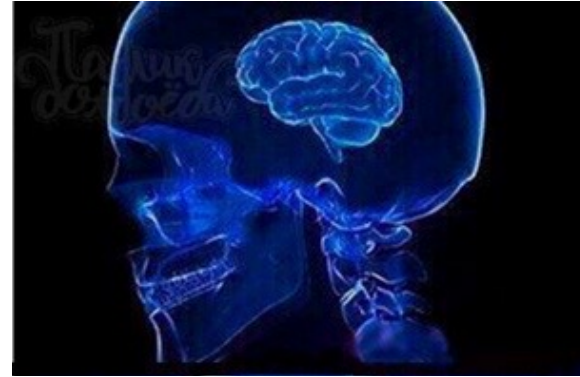
## Примеры метрик

- Выручка
- Количество активаций
- Маржа
- Etc



# Результаты глазами маркетологов

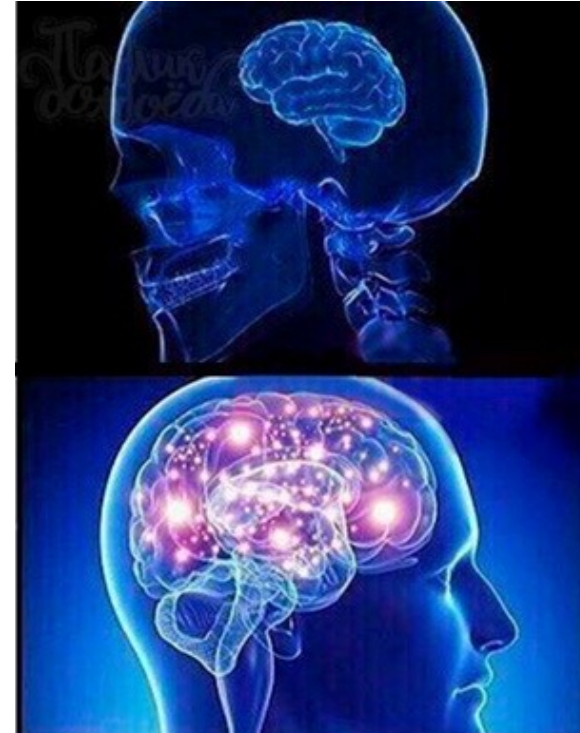
Конверсия в активацию оффера – 30!  
Была бы еще больше, если бы не  
выделяли КГ



# Результаты глазами маркетологов

Конверсия в активацию оффера – 30!  
Была бы еще больше, если бы не  
выделяли КГ

Конверсия в активацию оффера 30%  
в ЦГ, 20% в КГ. Значит, мы заработали  
на 10% дополнительной конверсии



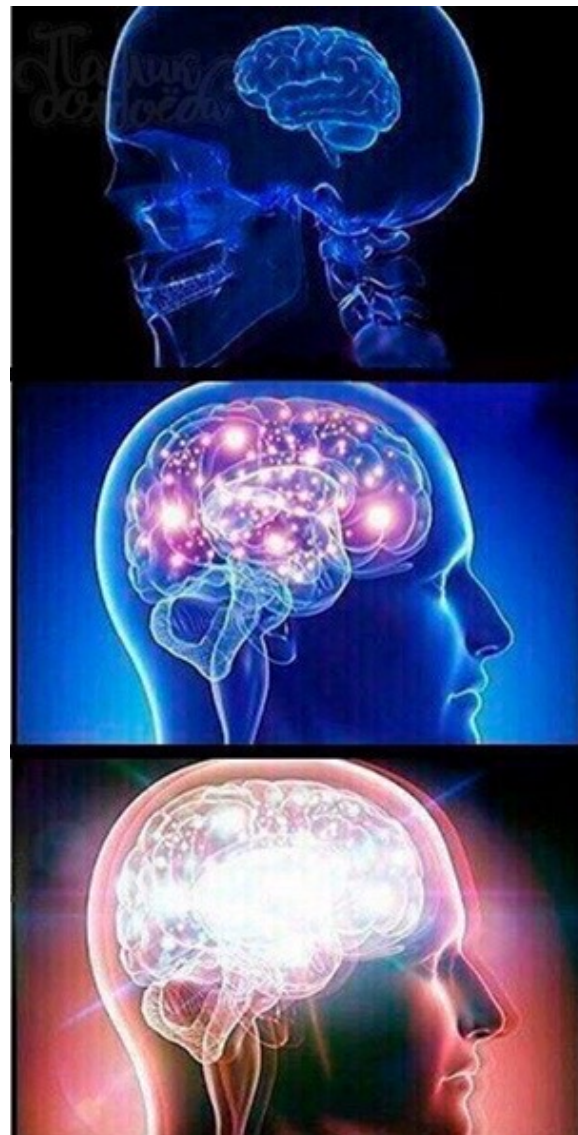
# Результаты глазами маркетологов

Конверсия в активацию оффера – 30!  
Была бы еще больше, если бы не выделяли КГ

Конверсия в активацию оффера 30%  
в ЦГ, 20% в КГ. Значит, мы заработали  
на 10% дополнительной конверсии

Конверсия в активацию оффера 30%  
в ЦГ, 20% в КГ. Значит, мы заработали  
на 10% дополнительной конверсии.

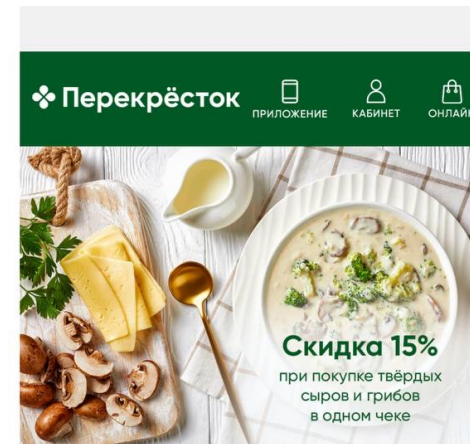
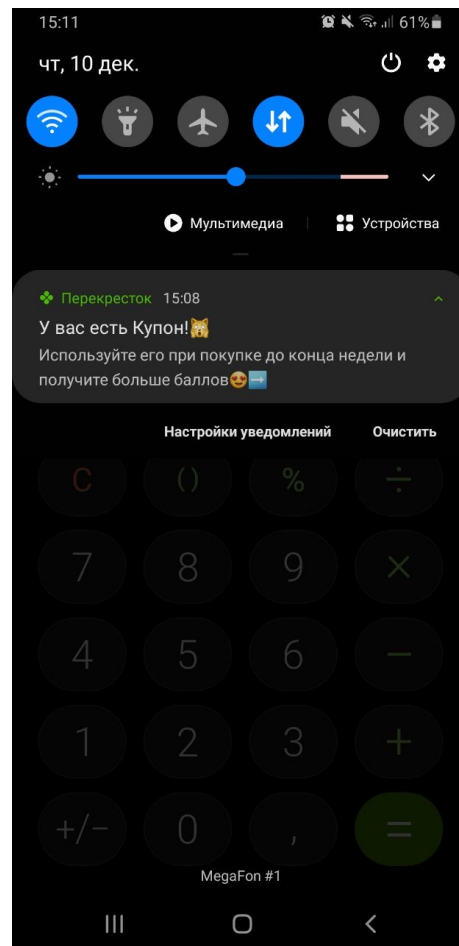
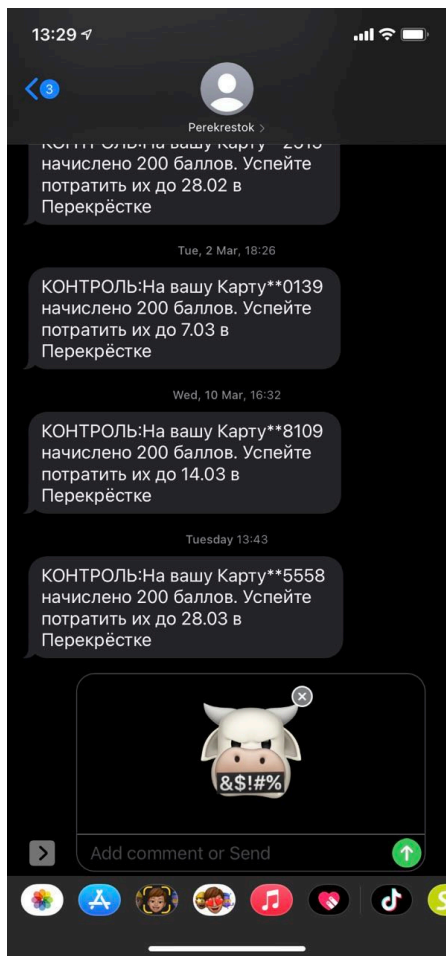
Но мы потратили 20% на ЦГ, которые  
и так откликнулись бы.





# Коммуникации

Под коммуникацией понимаем любой контакт в доступной канале с каким-то оффером



Для истинных гурманов — **скидка 15%** при покупке твёрдых сыров и грибов в одном чеке. Они идеально подойдут для вкусного ужина: приготовьте жаркое, жульен или запеканку для всей семьи.

[ВЫБРАТЬ БЛИЖАЙШИЙ ПЕРЕКРЁСТОК](#)

**ЧТО ДЕЛАТЬ?**



# Предыстория задачи



# Предыстория задачи



# Что хотим смоделировать

- **X** – признаки, описание клиента
- **T** – флаг оффера; 1 – был оффер, 0 – отсутствие
- **Y(0)** – целевая переменная (реакция) клиента в той версии Мультивселенной, где ему не сделали оффер
- **Y(1)** – целевая переменная (реакция) клиента в той версии Мультивселенной, где ему сделали оффер

# ЧТО ХОТИМ СМОДЕЛИРОВАТЬ

- **X** – признаки, описание клиента
- **T** – флаг оффера; 1 – был оффер, 0 – отсутствие
- **Y(0)** – целевая переменная (реакция) клиента в той версии Мультивселенной, где ему не сделали оффер
- **Y(1)** – целевая переменная (реакция) клиента в той версии Мультивселенной, где ему сделали оффер
- $\tau = Y(1) - Y(0)$  – causal effect

# ЧТО ХОТИМ СМОДЕЛИРОВАТЬ

- **X** – признаки, описание клиента
- **T** – флаг оффера; 1 – был оффер, 0 – отсутствие
- **$Y(0)$**  – целевая переменная (реакция) клиента в той версии Мультивселенной, где ему не сделали оффер
- **$Y(1)$**  – целевая переменная (реакция) клиента в той версии Мультивселенной, где ему сделали оффер
- $\tau = Y(1) - Y(0)$  – causal effect

# Проблема Causal Inference

- $\tau = Y(1) - Y(0)$  – causal effect
- Большой эффект – супер, таким клиентам срочно отправляем оффер
- Им же срочно не отправляем оффер
- Приходим к противоречию: мы не можем посчитать  $\tau$
- Простую регрессию не построим

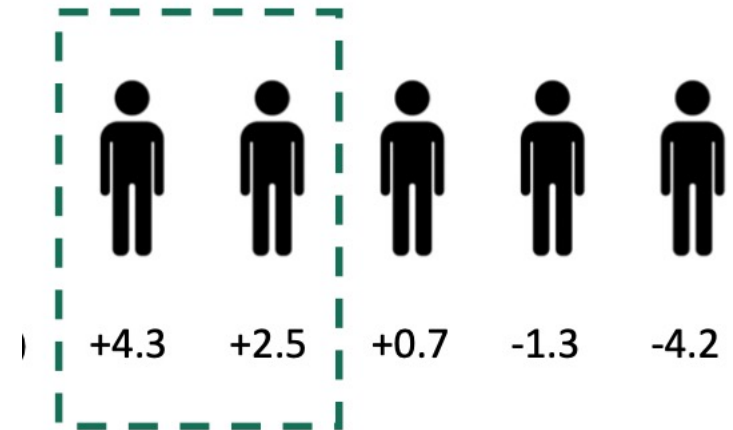


# Таргет

- $Y = TY(1) - (1 - T)Y(0)$  – целевая переменная
- $uplift(x_i) = \widehat{CATE} = \{ \text{Conditional Average Treatment Effect} \}$

$$= E[Y|X = x_i, T = 1] - E[Y|X = x_i, T = 0]$$

На сколько больше в среднем клиент с описанием  $x_i$  реагирует при коммуникации, чем без коммуникации





# Откуда взять данные для обучения?

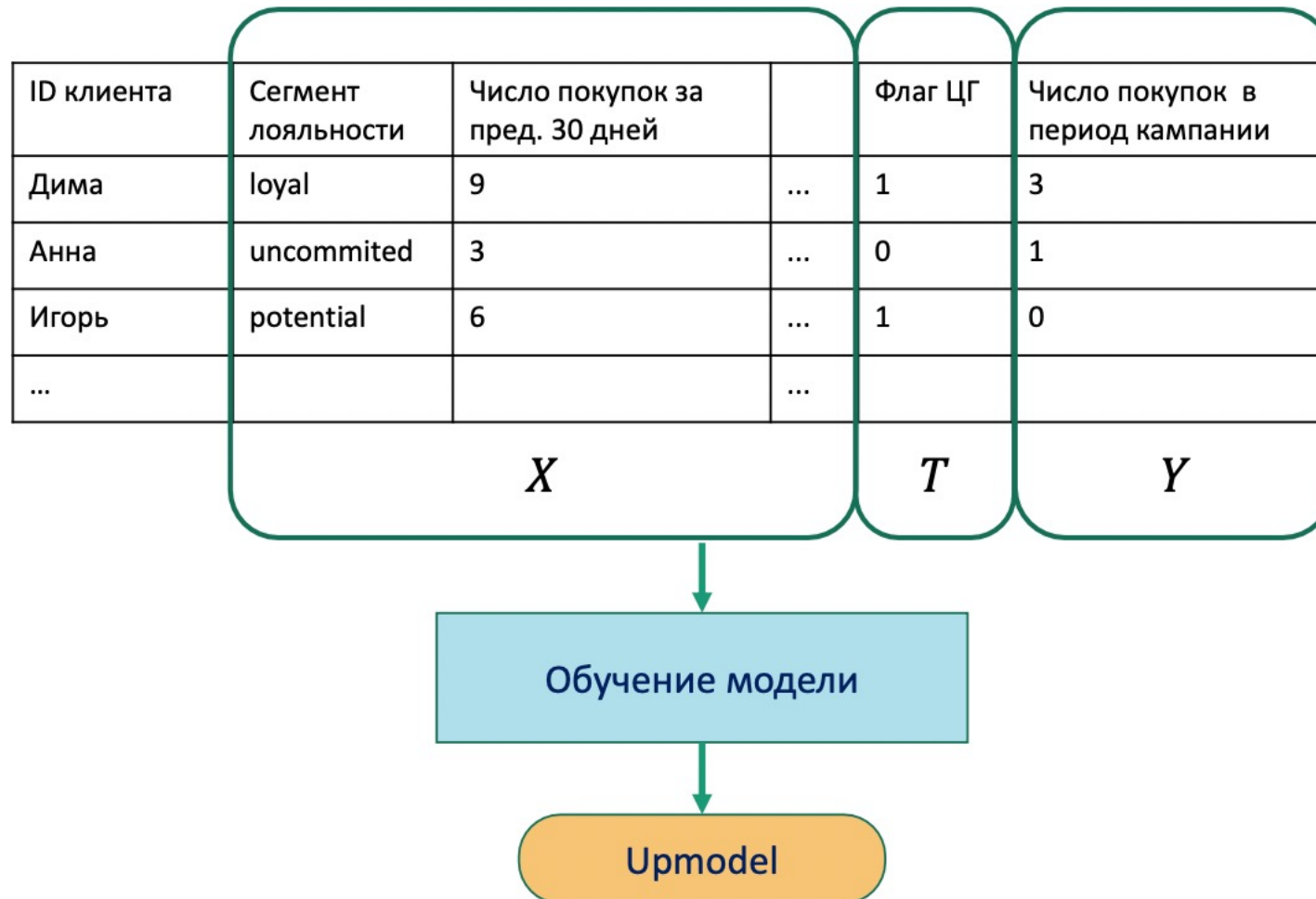
- $Y = TY(1) - (1 - T)Y(0)$  – целевая переменная
- $uplift(x_i) = \widehat{CATE} = \{ \text{Conditional Average Treatment Effect} \}$

$$= E[Y|X = x_i, T = 1] - E[Y|X = x_i, T = 0]$$

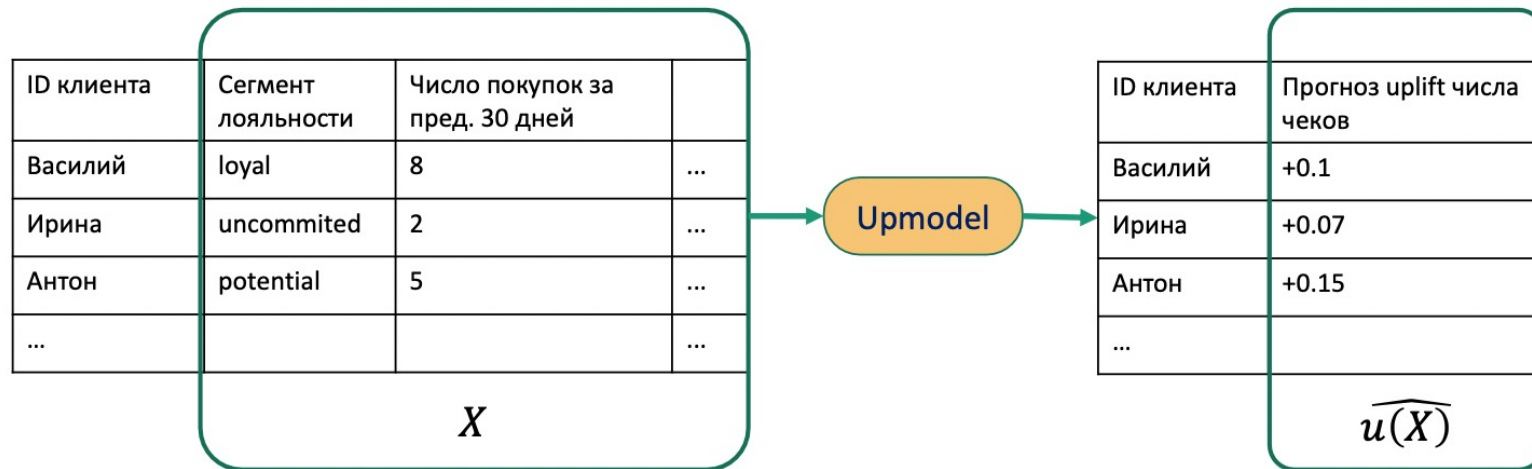
На сколько больше в среднем клиент с описанием  $x_i$  реагирует при коммуникации, чем без коммуникации



# Общая схема обучения



# Общая схема применения

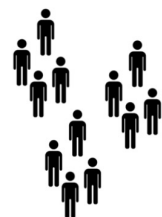
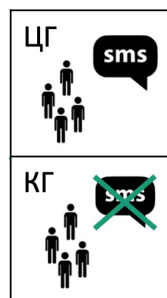


# Детали обучения

- Выборка нужна большая и лишенная смещений
- Как правило, запускаются холостые кампании:
  - На случайной, репрезентативной подвыборке из базыв



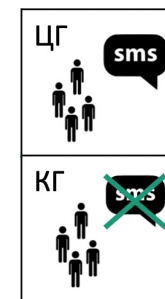
# Детали обучения



Случайная аудитория

$T$   
1  
0  
1  
0  
0  
1  
1  
0  
1

Uplift model



Отобранная моделью аудитория

# Методы обучения

## Solo Learner (S-Learner, Метод с 1 моделью)

$$\text{fit}\left(\begin{array}{cc} x_{11} & x_{1m} t_1 y_1 \\ x_{n1} & x_{nm} t_n y_n \end{array}\right)$$

# Методы обучения

## Solo Learner (S-Learner, Метод с 1 моделью)

$$fit\left(\begin{array}{cc} x_{11} & x_{1m} t_1 y_1 \\ x_{n1} & x_{nm} t_n y_n \end{array}\right)$$

$$predict\left(\begin{array}{cc} x_{11} & x_{1m} 1 \\ & 1 \\ x_{n1} & x_{nm} 1 \end{array}\right) - predict\left(\begin{array}{cc} x_{11} & x_{1m} 0 \\ & 0 \\ x_{n1} & x_{nm} 0 \end{array}\right) = uplift$$

# Методы обучения

## Two models approach (T-Learner, Метод с 2 моделями)

$$model_{treatment} = fit \left( \begin{array}{cc} x_{11} & x_{1m} y_1 \\ x_{p1} & x_{pm} y_p \end{array} \right)$$

$$model_{control} = fit \left( \begin{array}{cc} x_{p1} & x_{pm} y_p \\ x_{n1} & x_{nm} y_n \end{array} \right)$$

$$predict_{treatment} \begin{pmatrix} x_{11} & x_{nm} 1 \\ & 1 \\ x_{n1} & x_{nm} 1 \end{pmatrix} - predict_{control} \begin{pmatrix} x_{11} & x_{nm} 0 \\ & 0 \\ x_{n1} & x_{nm} 0 \end{pmatrix}$$



# Методы обучения

## Transformed outcome

- Целевая переменная подвергается подобной трансформации

$$Y^{new} = Y_i \frac{(T_i - p)}{p(1 - p)}$$

- P – propensity score, вероятность отнесения к контрольной или целевой группе
- Обучаемся как на обычную регрессию
- Тут показывают, что матожидание трансформированного таргета и есть искомый causal effect, т.е.

$$\tau = uplift = E[Y^{new} | X]$$

# Метрики качества

## Uplift@k

- Отсекаем из всей тестовой выборки по порогу  $k\%$
- Сортируем выборку по величине предсказанного uplift
- Смотрим на разницу в среднем по таргету ( $response\_rate\_$  между целевой и контрольной группой)

$$uplift@k = response\ rate@k_{(treatment)} - response\ rate@k_{(control)}$$

$$response\ rate@k = mean(Y@k)$$

$Y@k$  — таргет на  $k\%$

# Метрики качества

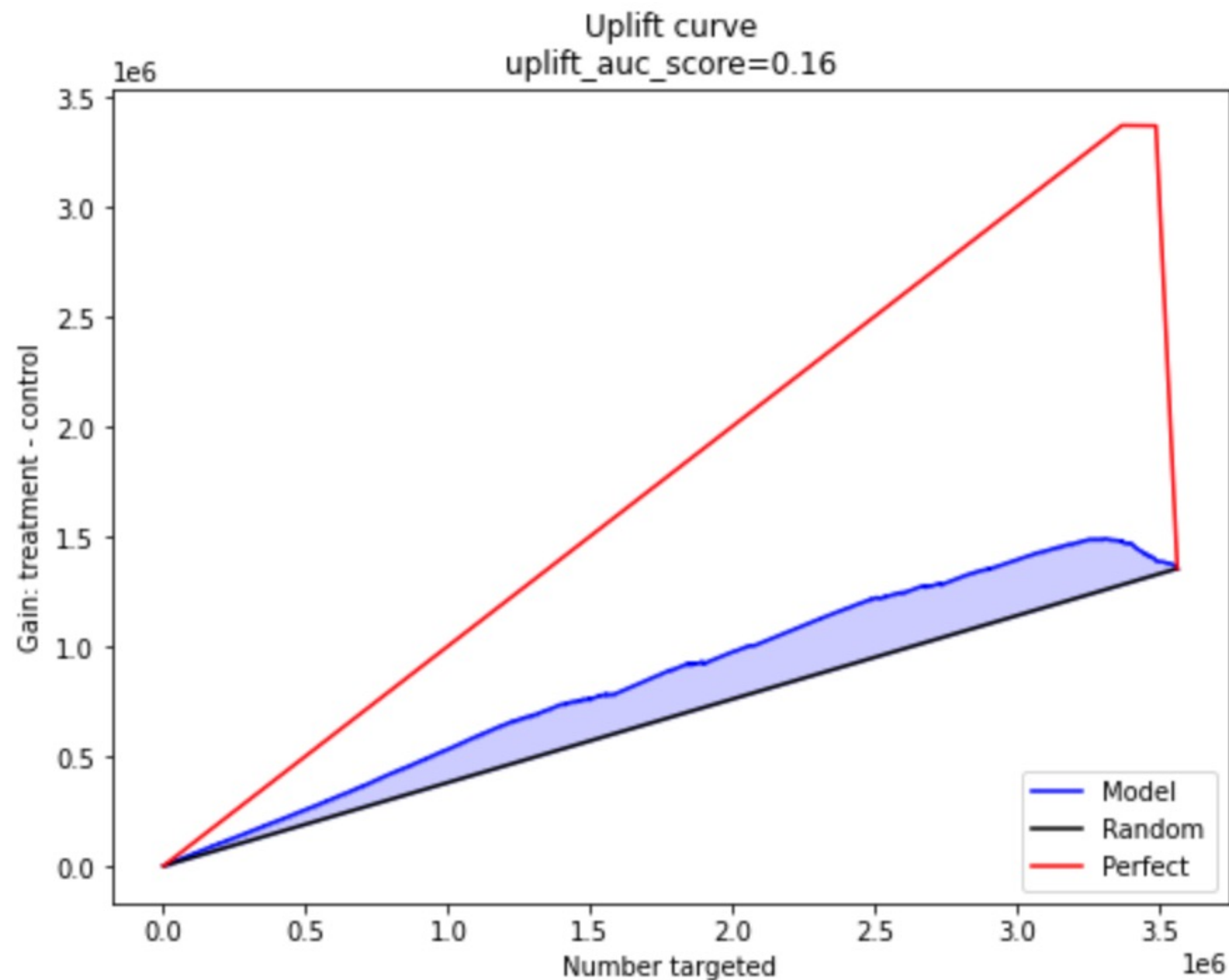
## Uplift curve

$$uplift\ curve(t) = \left( \frac{Y_t^T}{N_t^T} - \frac{Y_t^C}{N_t^C} \right) (N_t^T + N_t^C)$$

$t$  — накопившееся количество объектов

$Y_t^T, Y_t^C$  — таргет в *treatment* группе, таргет в *control* группе

$N_t^T, N_t^C$  — размер *treatment* группы, размер *control* группы



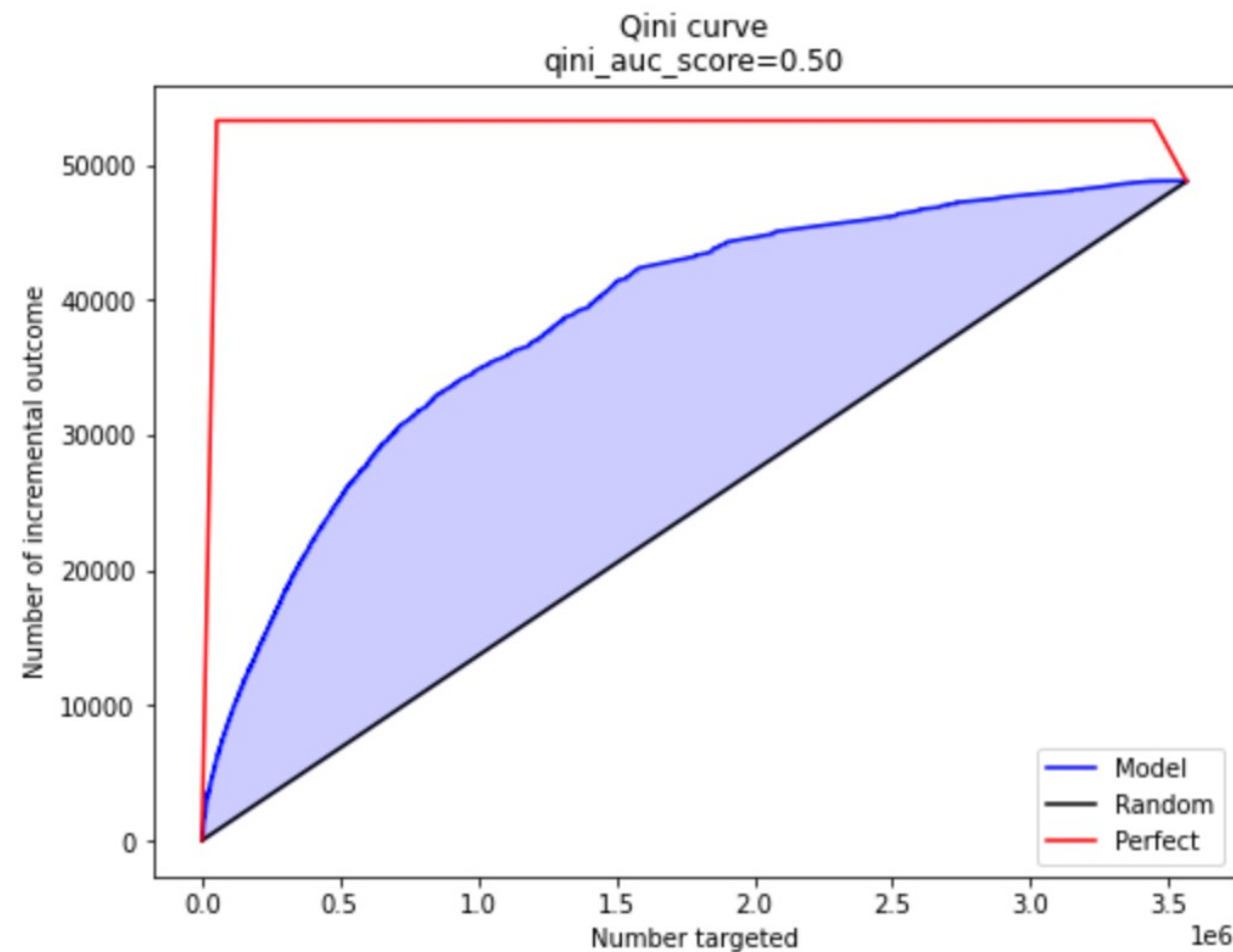
# Метрики качества

## Qini curve

$$qini\ curve(t) = Y_t^T - \frac{Y_t^C N_t^T}{N_t^C}$$

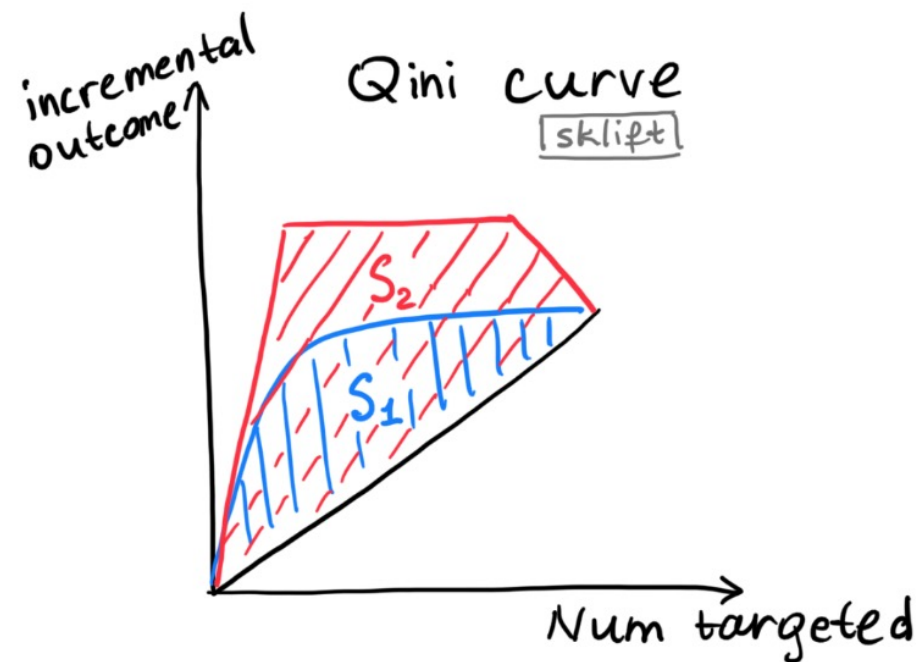
$Y_t^T, Y_t^C$  — таргет в *treatment* группе, таргет в *control* группе

$N_t^T, N_t^C$  — размер *treatment* группы, размер *control* группы



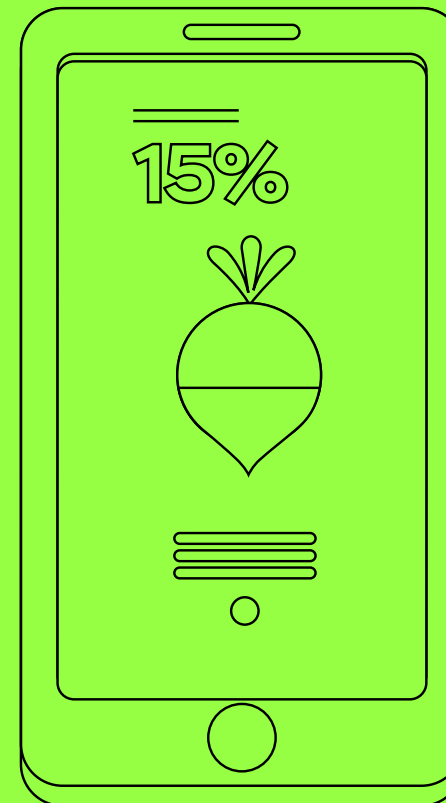
# Метрики качества

## Откуда берутся скоры



$$\text{Qini coefficient} = \frac{S_1}{S_2}$$

# мудрое наставление



# ССЫЛКИ

## **Подробная серия статей-объяснений:**

[https://habr.com/ru/company/ru\\_mts/blog/485980/](https://habr.com/ru/company/ru_mts/blog/485980/)

## **Курс по uplift-моделированию**

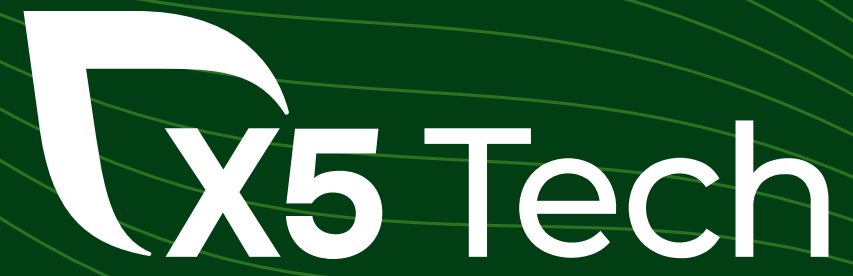
<https://ods.ai/tracks/uplift-modelling-course/blocks/69a5fb73-458b-4fa0-9edf-201eb10b1538>

## **Библиотеки**

<https://github.com/uber/causalml>

<https://github.com/wayfair/pylift>

<https://www.uplift-modeling.com/en/v0.3.0/index.html>



20 августа 2021