

Look alike & Uplifts

# План

## 1. Ресар и хвостик предыдущего занятия

# О чем говорили

- Задача спроса – ядро ритейл-сервисов
- С алгоритмической точки зрения, почти всегда бустинги
- Валидация временных рядов  $\neq$  K-fold

**Что может помешать?**

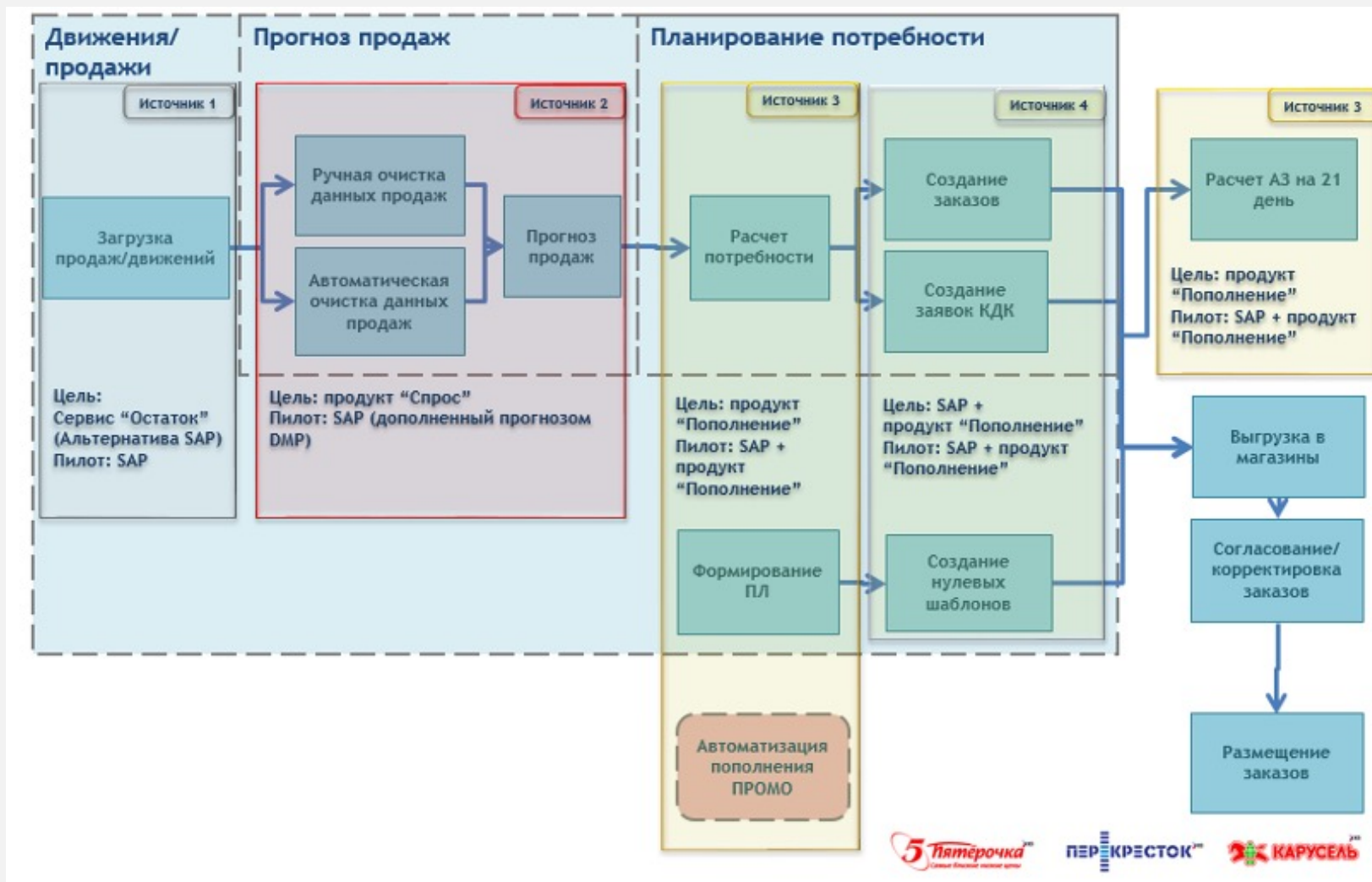
# Что может помешать?

- Проблема «холодного старта»: новые товары/магазины, недостаточно предыстории самого объекта
- Недостаточно наблюдений для редких категорий: мало данных, редко продаются
- Сезонность
- Рост бизнеса: компания масштабируется, продажи растут
- Ловушка заниженной доступности: мало продали, потому что мало привезли, мало продаем
- Перепрогноз / затарка: много заказали, не продали, большие скидки

# Ожидание

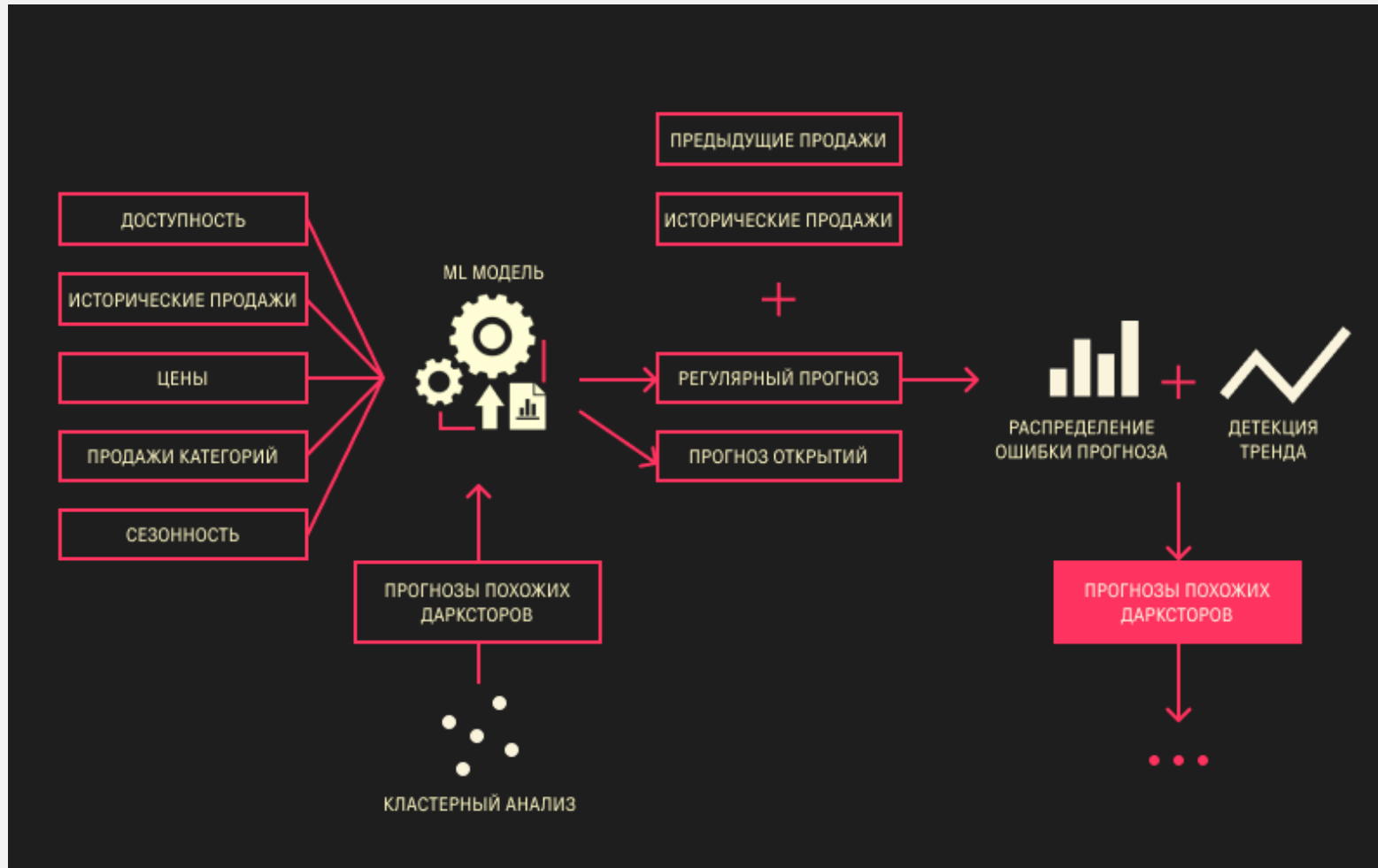


# Схема работы в Х5



[Ссылка](#)

# Схема работы модуля в Самокате



[Ссылка](#)



# Вопрос к обсуждению:

- Сделали модель прогноза спроса
- Как доказать ее бизнес-эффективность?

# План

## 1. Ресар и хвостик предыдущего занятия

# План

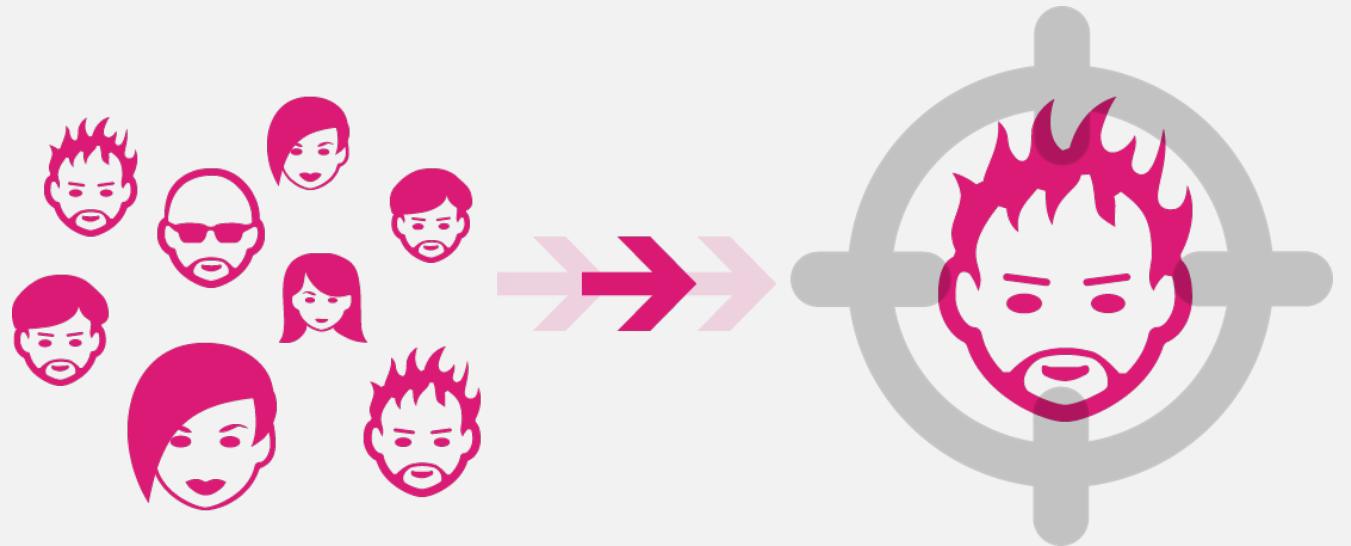
1. Ресар и хвостик предыдущего занятия
2. **O look-alike**

# Найти похожих

- Есть узкий сегмент, хочется его увеличить
- Например, Перекресток выходил на кейс-чемпионат с запросом увеличить аудиторию клуба «Здоровых привычек»
- Допустим, всего участников 10 000. Активная клиентская база – несколько млн. Наверняка есть сегмент(-ы), кто мог бы присоединиться к клубу, но, мб, не знает о нем/не думал/ не видел

# Look-alike

- «Можно же лукэлайки построить, это же легко!»
- Задача: найти похожих пользователей из известного сегмента
- Функционал включен во многие рекламные площадки



# Как сделать?

- Есть узкий сегмент, хочется его увеличить
- Например, Перекресток выходил на кейс-чемпионат с запросом увеличить аудиторию клуба «Здоровых привычек»
- Допустим, всего участников 10 000. Активная клиентская база – несколько млн. Наверняка есть сегмент(-ы), кто мог бы присоединиться к клубу, но, мб, не знает о нем/не думал/ не видел

# Свести к Supervised-задаче

- Есть  $X$  людей с положительными метками
- Можем насэмплировать  $Y$  людей с отрицательными метками
- $\Rightarrow$  классическая задача классификации
- Сложности?

# Свести к Supervised-задаче

- Есть  $X$  людей с положительными метками
- Можем насэмплировать  $Y$  людей с отрицательными метками
- $\Rightarrow$  классическая задача классификации
  
- Сложности?
  - Легко скатиться к несбалансированной задаче
  - Отрицательные метки – не факт, что отрицательные



# Свести к Supervised-задаче

- Есть  $X$  людей с положительными метками
- Можем насэмплировать  $Y$  людей с отрицательными метками
- $\Rightarrow$  классическая задача классификации
  
- Сложности?
  - Легко скатиться к несбалансированной задаче
  - Отрицательные метки – не факт, что отрицательные

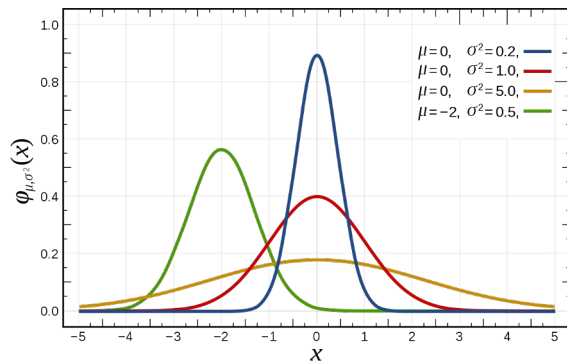
# Свести к One-Class Learning

- В 1996 году представили понятие One-class classification – обучении модели только на одном классе данных
- Примеры такого подхода не только в LAL, но и:
  - Детекция аномалий
  - Поиск выбросов
  - Обнаружение новизны
- Частные примеры: аварии в двигателях, критические ситуации на АЭС, поломки в нефтяных скважинах и т.п.

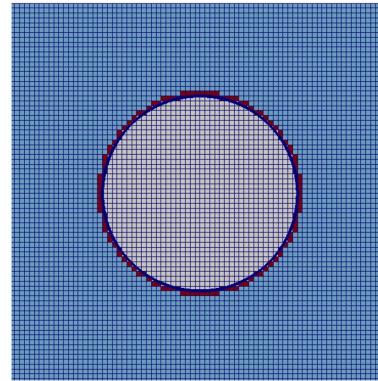
# План

1. Ресар и хвостик предыдущего занятия
2. O look-alike
- 3. One-Class Learning**

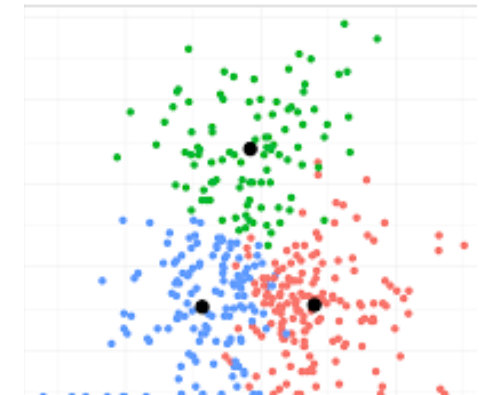
# Свести к One-Class Learning



- Density Methods
- Gaussian Mixtures, Parzen Density



- Boundary Methods
- K-centers, Nearest Neighbor, SVDD



- Reconstruction Methods
- K-means Clustering, Learning Vector Quantization, Self-Organizing Maps

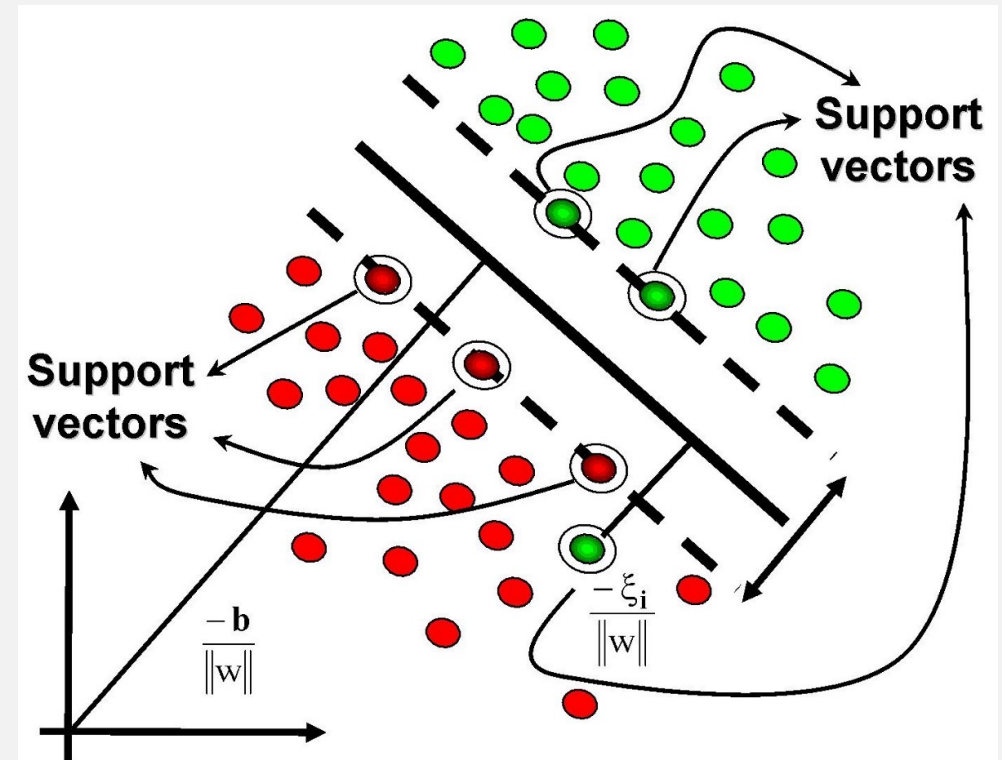
# SVM-классический

- Классический SVM проводит гиперплоскости, чтобы разделить данные одного класса от другого

$$\min_{w, b, \xi_i} \frac{\|w\|^2}{2} + C \sum_{i=1}^n \xi_i$$

subject to:

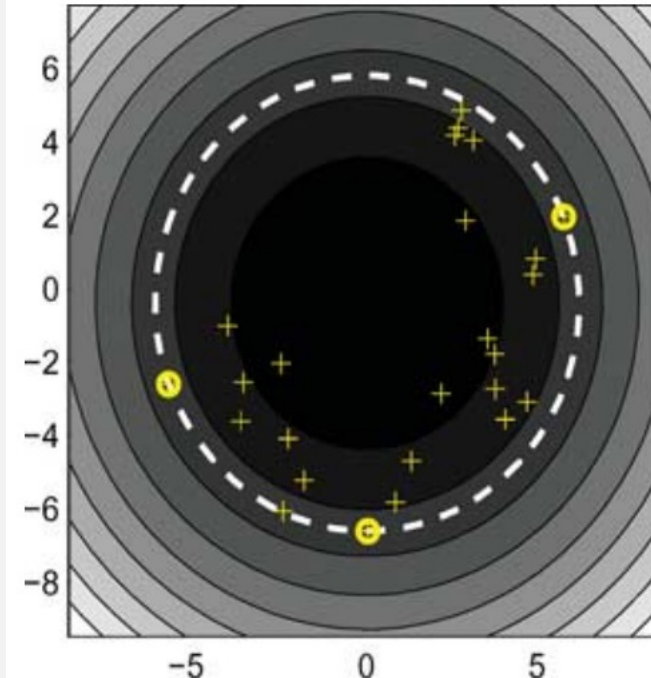
$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i \quad \text{for all } i = 1, \dots, n$$
$$\xi_i \geq 0 \quad \text{for all } i = 1, \dots, n$$



# Support Vector Data Description

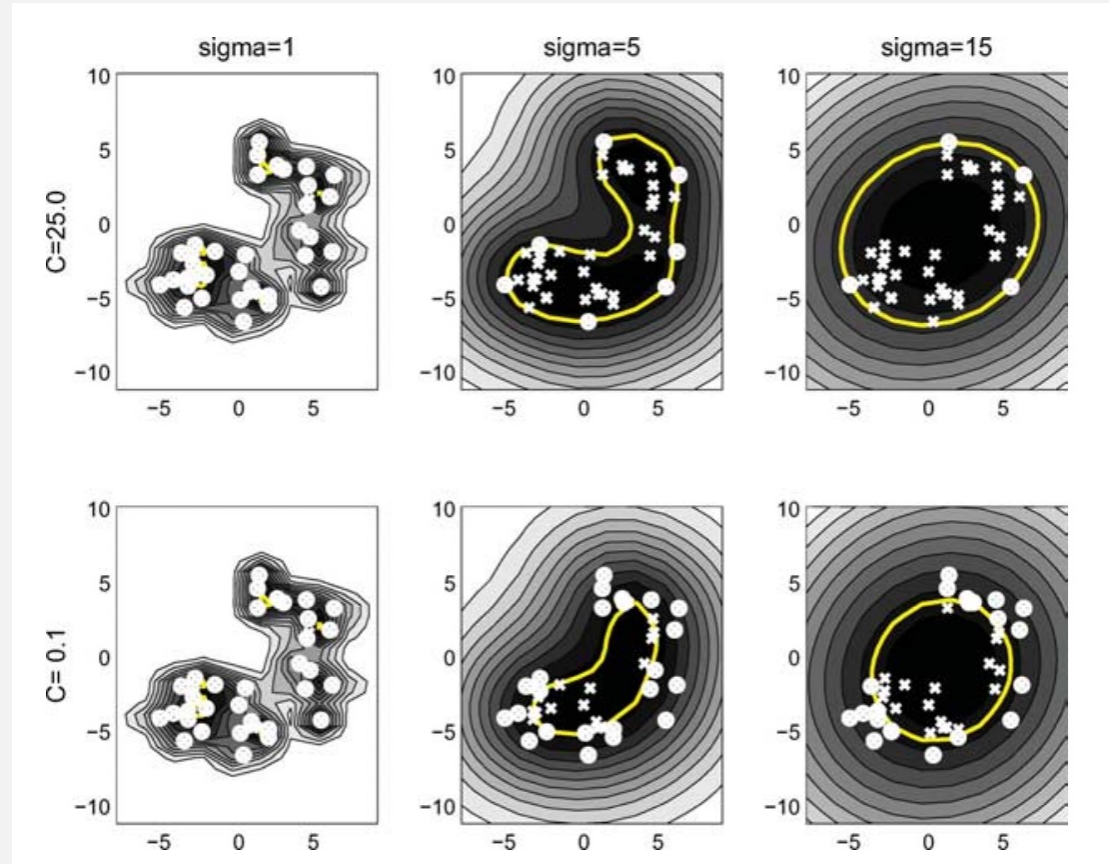
- Алгоритм восстанавливает сферу вокруг данных.  
Минимизируется объем сферы

$$\begin{aligned} \min_{R, \mathbf{a}} R^2 + C \sum_{i=1}^n \xi_i \\ \text{subject to:} \\ \|x_i - \mathbf{a}\|^2 \leq R^2 + \xi_i \quad \text{for all } i = 1, \dots, n \\ \xi_i \geq 0 \quad \text{for all } i = 1, \dots, n \end{aligned}$$



# Support Vector Data Description

## Kernel Trick



# План

1. Ресар и хвостик предыдущего занятия
2. O look-alike
3. One-Class Learning
- 4. Uplift**



# Типичная кампания

# Типичная кампания

## Примеры офферов

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

# Типичная кампания

## Примеры офферов

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

## Примеры метрик

- Выручка
- Количество активаций
- Маржа
- Etc

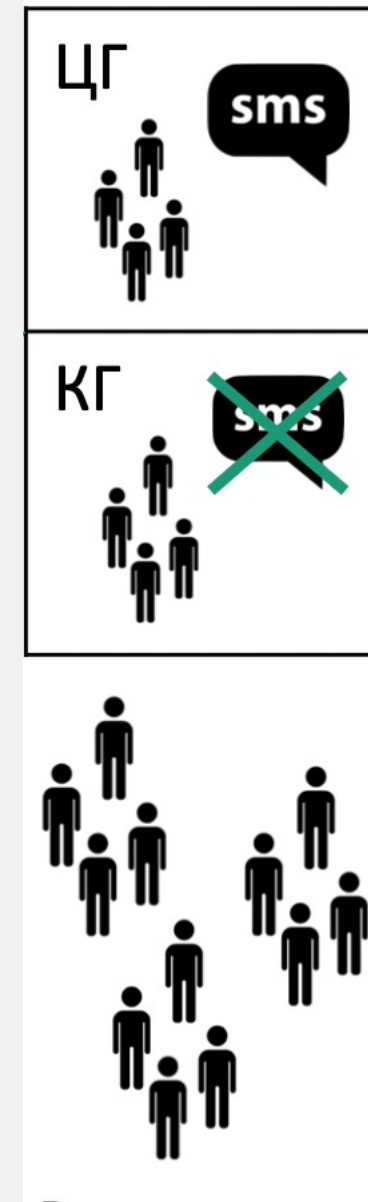
# Типичная кампания

## Примеры офферов

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

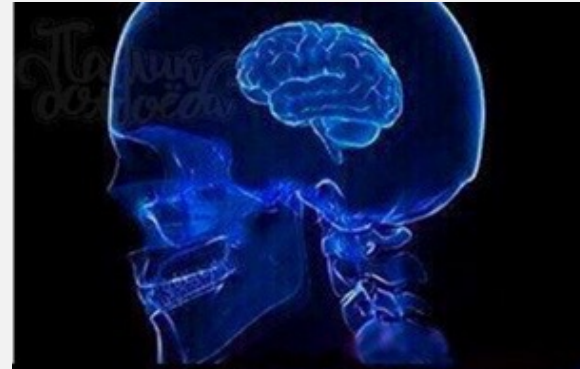
## Примеры метрик

- Выручка
- Количество активаций
- Маржа
- Etc



# Результаты глазами бизнеса

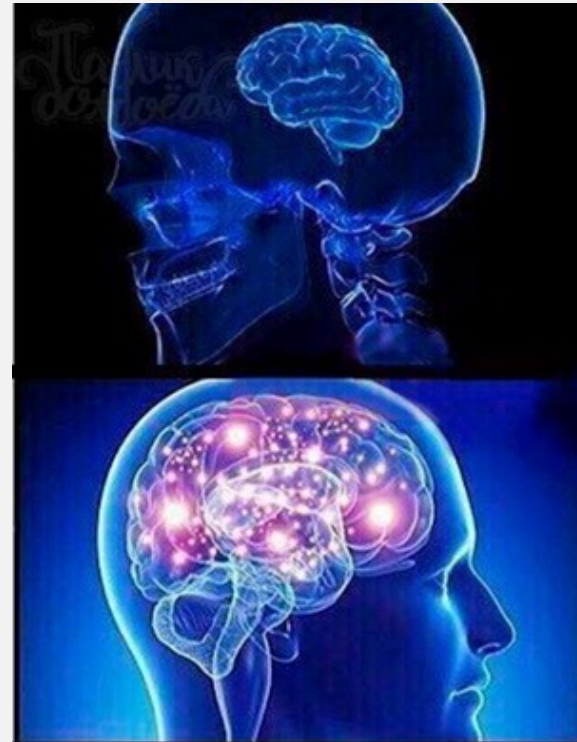
Конверсия в активацию оффера – 30!  
Была бы еще больше, если бы не  
выделяли КГ



# Результаты глазами бизнеса

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



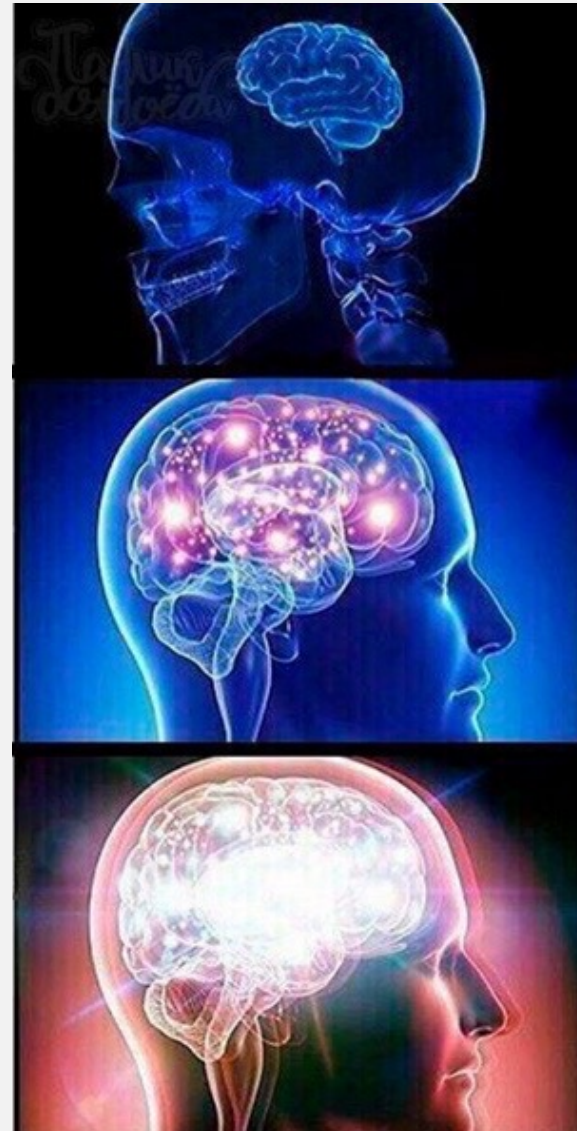
# Результаты глазами бизнеса

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

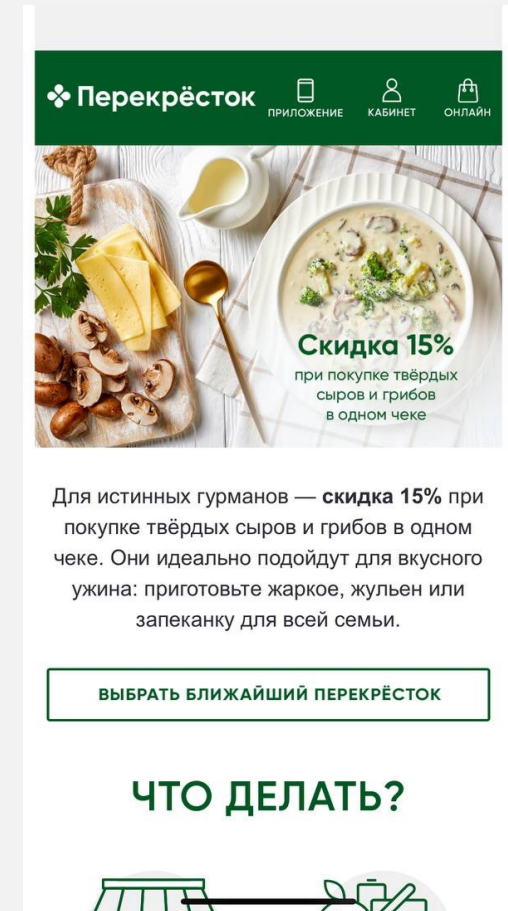
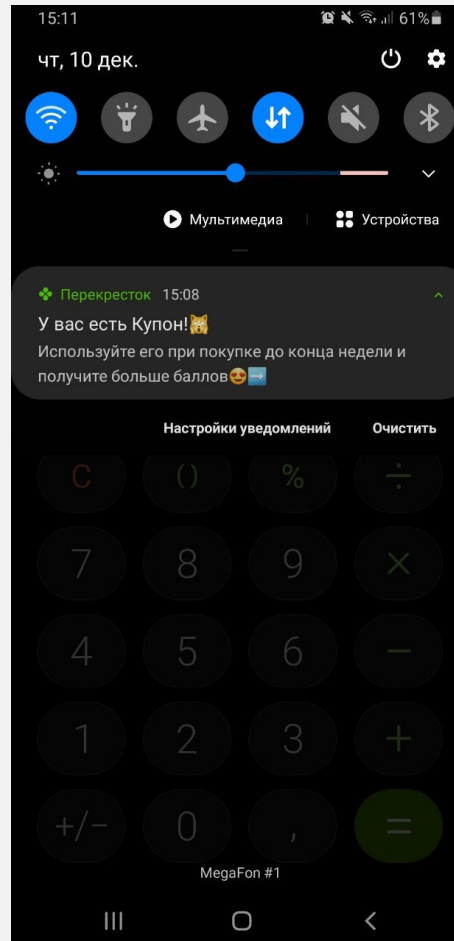
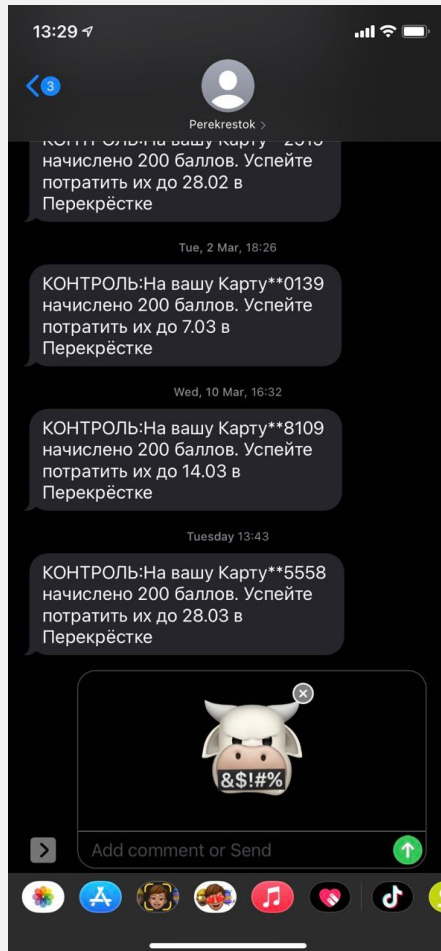
Конверсия в активацию оффера  
30% в ЦГ, 20% в КГ. Значит, мы  
заработали на 10% дополнительной  
конверсии

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.

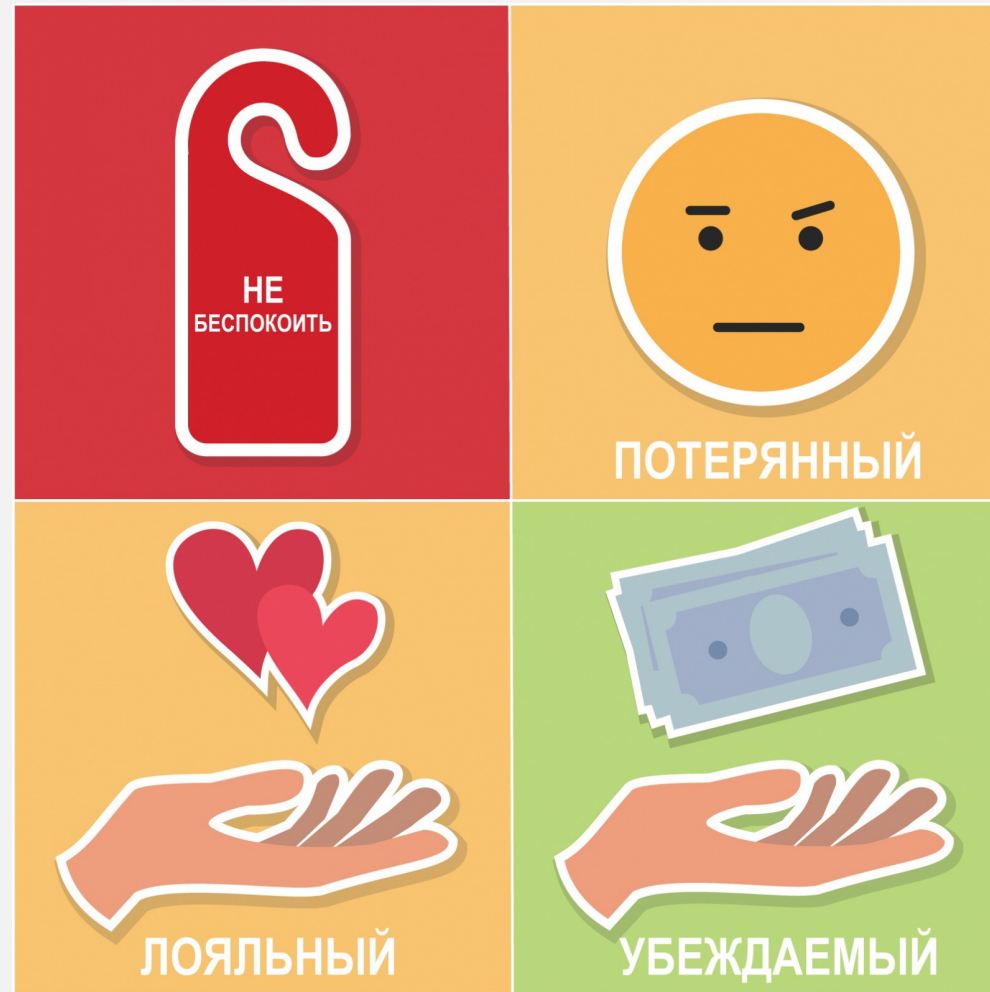


# Коммуникации

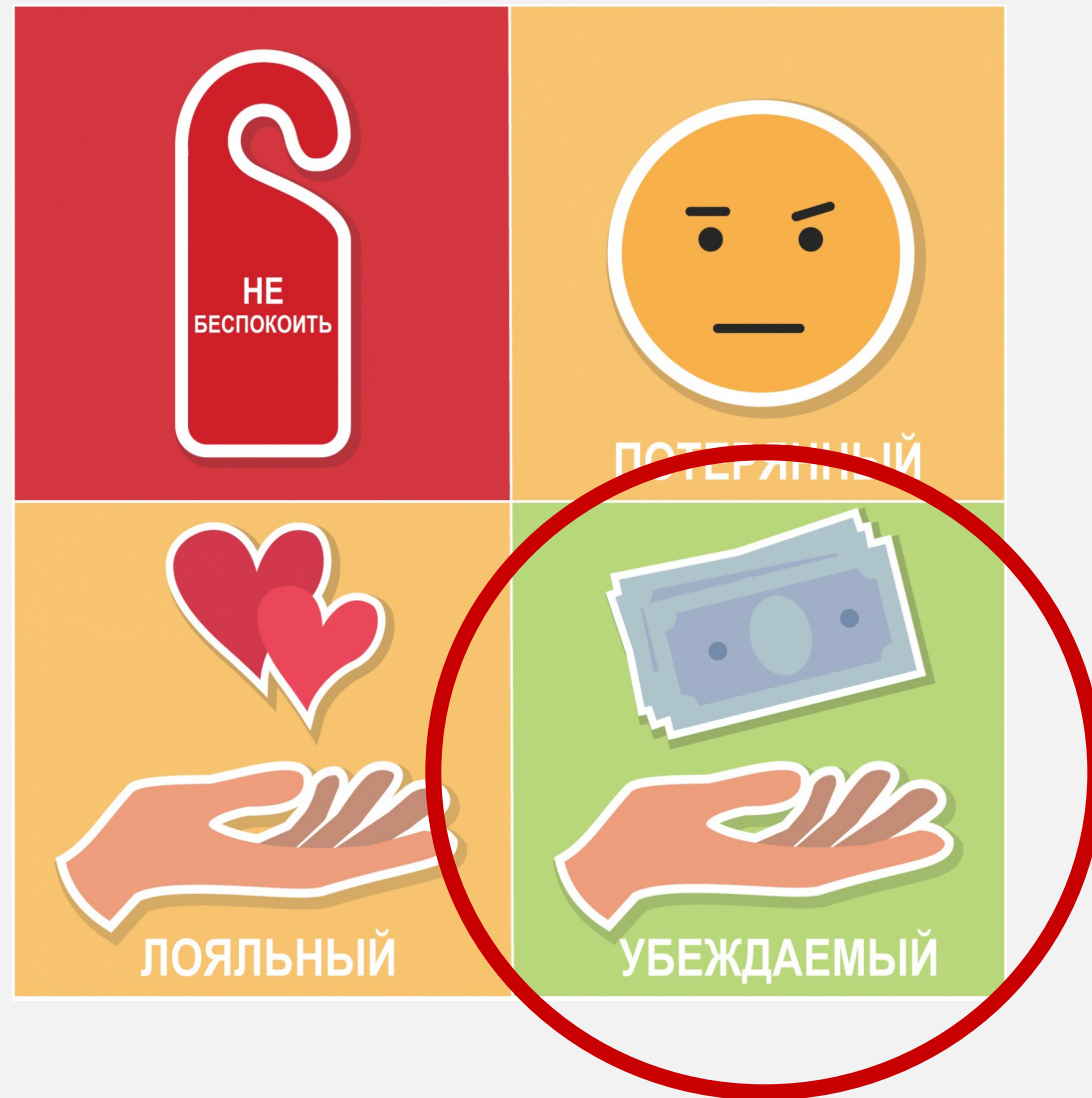




# Основы задачи



# Основы задачи



# Что моделируем

- $X$  – 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 моделируем

- $X$  – 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

- $\tau = Y(1) - Y(0)$  – 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



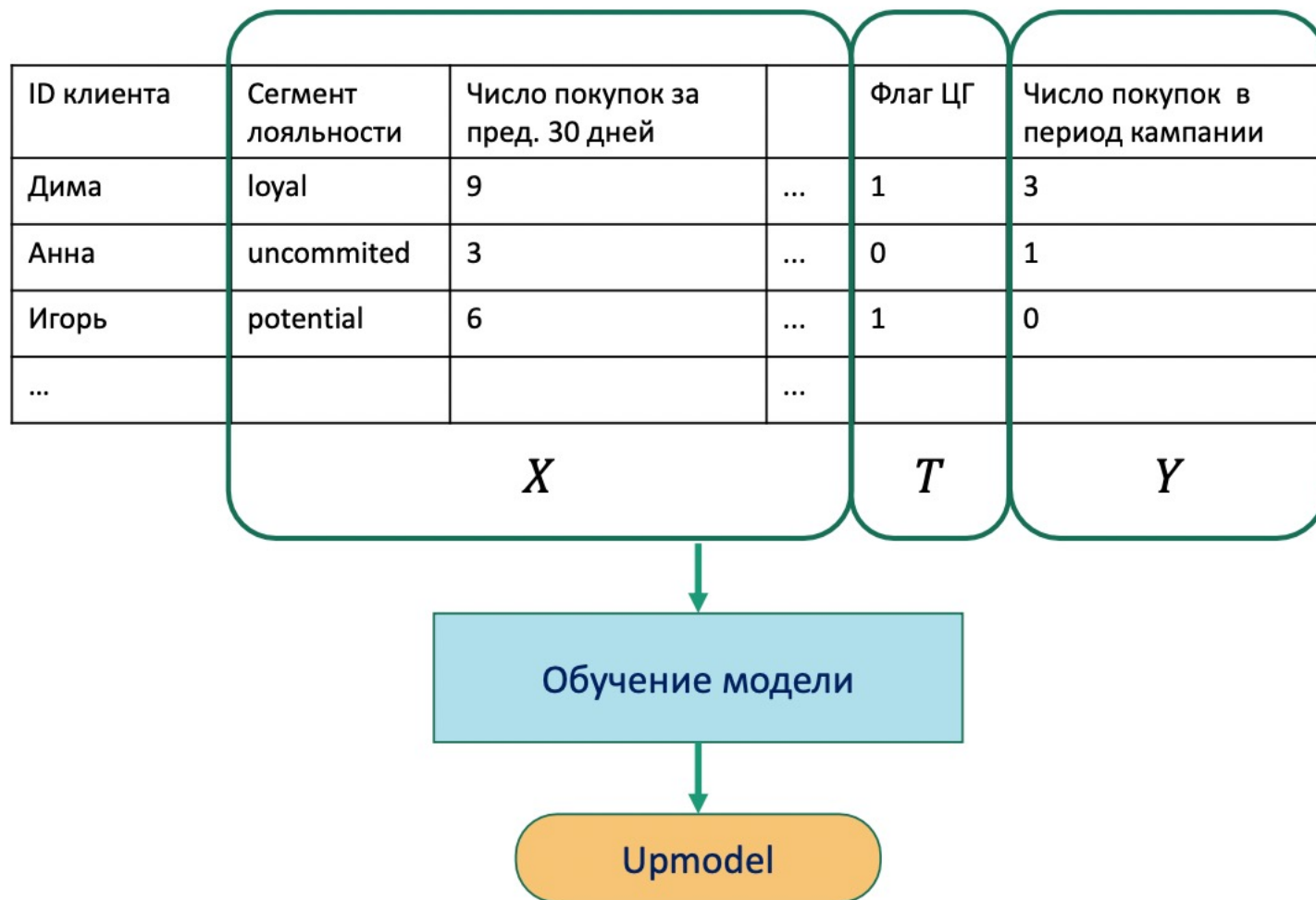
# Таргет

- $Y = TY(1) - (1 - T)Y(0)$  — 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



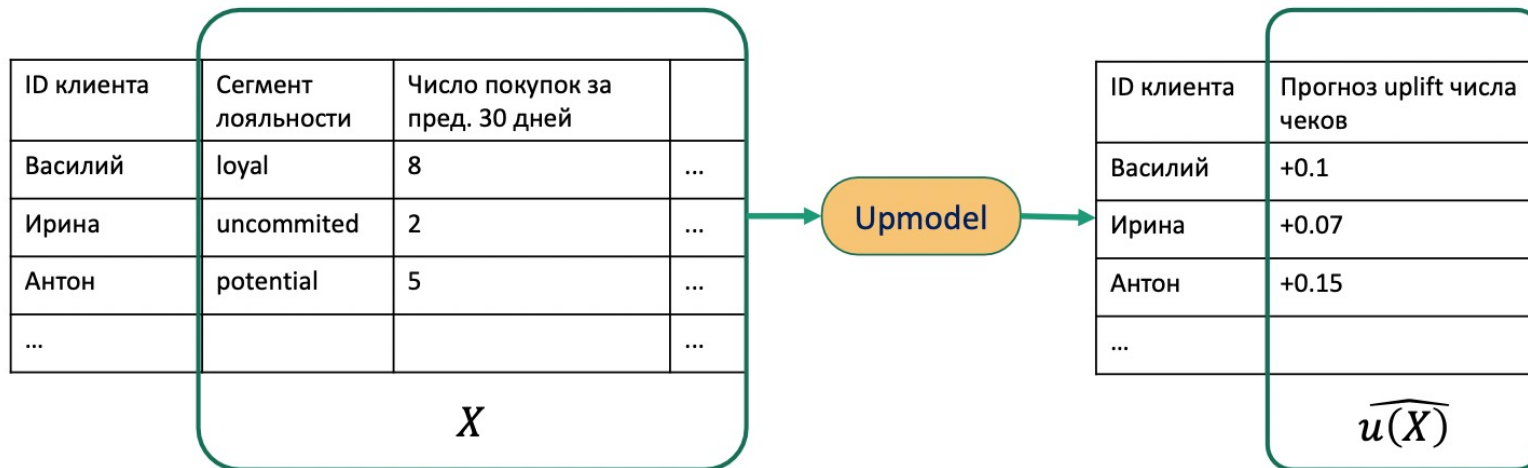
# Данные для обучения?

# Данные для обучения?



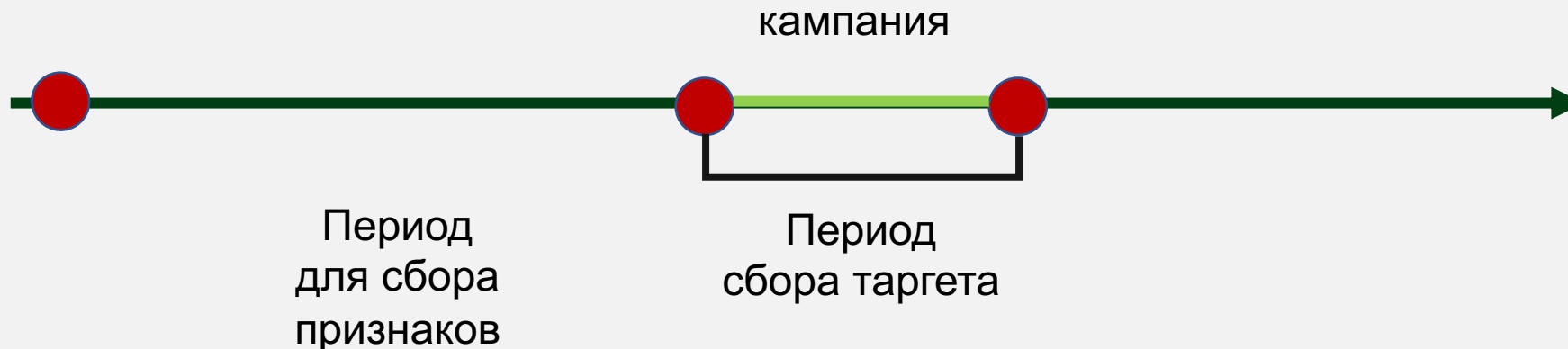


# Общая схема применения

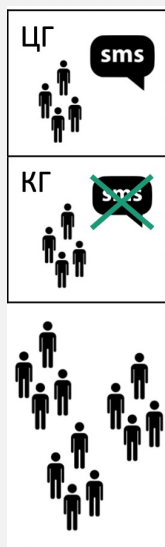
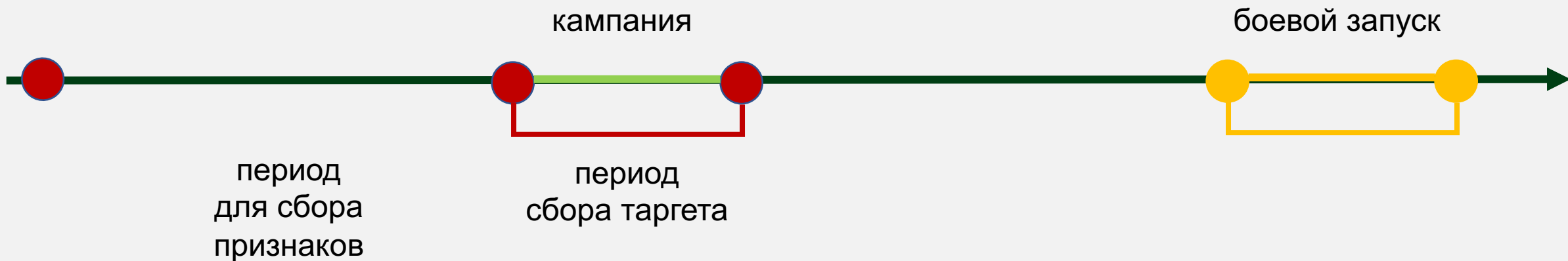


# Детали обучения

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



# Детали обучения



$T$

1

0

1

0

0

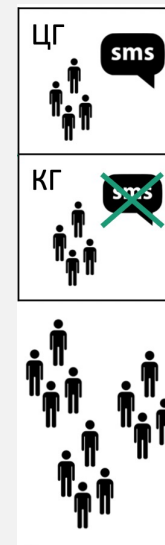
1

1

0

1

Uplift model



Случайная аудитория

Отобранная моделью аудитория