



Рекомендательные системы

Uplift-моделирование



Проверить, идет ли запись

Меня хорошо видно && слышно?



Ставим "+", если все хорошо
"-", если есть проблемы



Тема вебинара

Uplift-моделирование



Елена Позднеева

Team Lead Data Scientist, MeraTex

Как подобрать **персональный** тариф?

Как **удержать/привлечь** клиента?

Как принести компании **финансовый** эффект?

Как повышать **лояльность** клиента?

Контакты:

@frideliya

www.linkedin.com/in/elena-pozdneevac

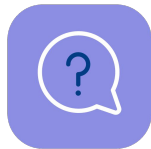
Правила вебинара



Активно
участвуем



Off-topic обсуждаем
в учебной группе **#RecSys-2024-10**
<https://t.me/+tr0h3CYc3xs40Tdi>



Задаем вопрос
в чат или голосом



Вопросы вижу в чате,
могу ответить не сразу

Условные обозначения



Индивидуально



Время, необходимое
на активность



Пишем в чат



Говорим голосом



Документ



Ответьте себе или
задайте вопрос

Маршрут вебинара

— Что такое uplift-моделирование?

— Как правильно собрать данные для разработки

— Подходы с одной моделью

— Подходы с двумя моделями

— Метрики качества

— Практика

— Рефлексия



Цели вебинара

К концу занятия вы сможете

1. формализовать задачу Uplift - моделирования
2. научиться решать задачу Uplift - моделирования при помощи построения одной и нескольких моделей

Смысл

Зачем вам это уметь

1. Понимать, когда нужно задействовать аппарат uplift-моделирования
2. Как правильно собрать данные для uplift
3. Как разработать модель и оценить ее качество



Что вы знаете про uplift-моделирование?

- 1 - ничего не знаю
- 2 - что-то знаю / слышал
- 3 - как рыба в воде!



Что такое uplift-моделирование?

Введение



коммуникации
не было

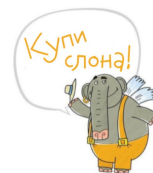
коммуникация
была

не купил	купил
00	01
10	11



пришли за
слоном сами

пришли, потому что
на них повлияла
коммуникация



Введение

	не купил	купил
коммуникации не было	00	01
коммуникация была	10	11

Look-a-like модель: оцениваем вероятность того, что клиент купит слона **сам**. Знаем на ретро тех, кто купил слона - и далее ищем **похожих** на тех, кто купил слона.

Response модель: оцениваем вероятность того, что клиент купит слона **при условии коммуникации**. Знаем на ретро тех, кого убедили купить слона - и далее ищем **похожих** на тех, кто купит слона после коммуникации.



Uplift-модель оценивает чистый эффект от коммуникации, то есть разницу в эффекте при наличии и отсутствии взаимодействия



Как правильно собрать данные для разработки?

Контрольная и экспериментальная группы



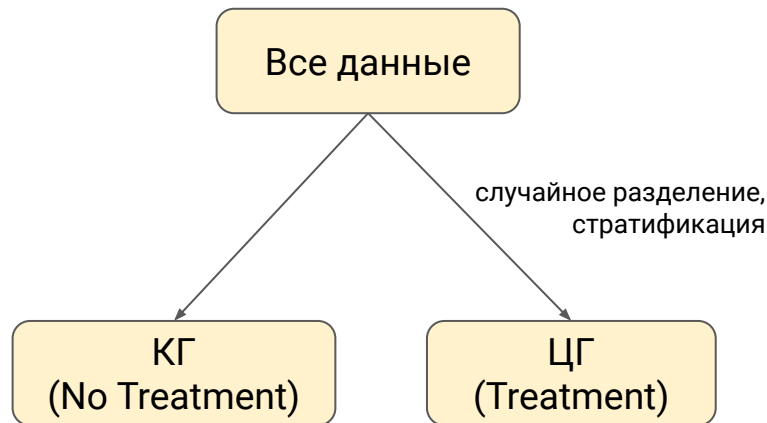
Контрольная группа состоит из субъектов, которые не подвергаются воздействию, оцениваемому в эксперименте.



Экспериментальная (целевая) группа состоит из субъектов, которые подвергаются воздействию, оцениваемому в эксперименте.

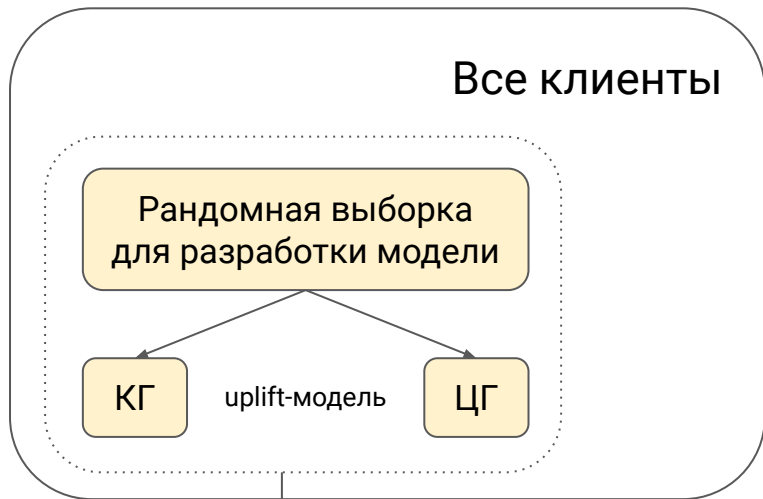


Цель: изолировать эффект воздействия и оценить его чистое влияние на результат



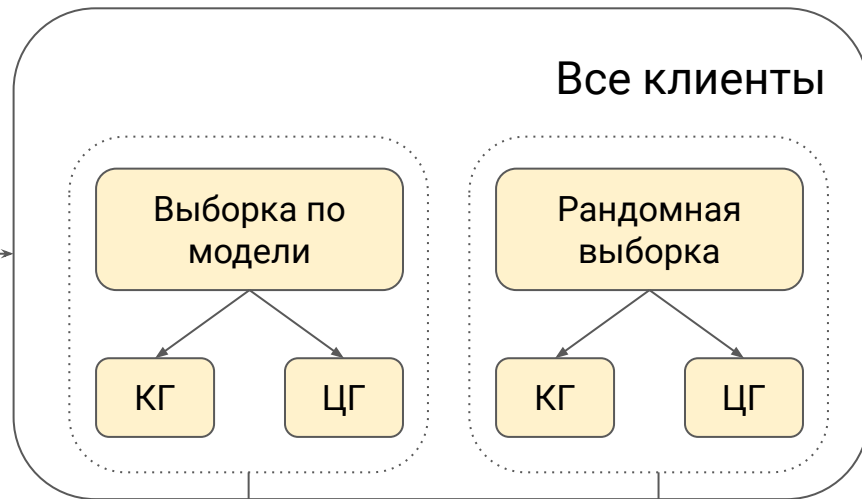
Дизайн эксперимента и эксплуатации

1. Сбор данных и разработка



идем разрабатывать модель

2. Тестирование и эксплуатация



uplift по модели
больше?



Терминология

X_i признаковое описание клиента

Y_i наблюдаемая реакция клиента

W_i наличие воздействия

$$\tau_i = Y_i^1 - Y_i^0$$

causal effect

$$Y_i = \begin{cases} Y_i^1, & \text{если } W_i = 1 \\ Y_i^0, & \text{если } W_i = 0 \end{cases}$$

$$Y_i = W_i Y_i^1 + (1 - W_i) Y_i^0$$

$$CATE = E[Y_i^1 | X_i] - E[Y_i^0 | X_i]$$

Conditional Average Treatment Effect

$$uplift = E[Y_i | X_i = x, W_i = 1] - E[Y_i | X_i = x, W_i = 0]$$

uplift как оценка CATE

Какие методы мы изучим?

Подходы с одной моделью

Мультиклассовая
модель

Трансформация
классов

Одна модель
S-learner

Подходы с двумя моделями

Две независимые
модели T-learner

Две зависимые
модели

Две зависимые
модели X-learner

Вопросы?



Ставим “+”,
если вопросы есть



Ставим “-”,
если вопросов нет



Подходы с одной моделью

Одна модель S-Learner

The training process:

$$\text{fit} \left(\begin{matrix} x_{11} & \cdots & x_{1k} & w_1 & y_1 \\ \vdots & \ddots & \vdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nk} & w_n & y_n \end{matrix} \right)$$

X_{train} W_{train} Y_{train}

Признаки = фичи + флаг коммуникации
Таргет = отклик

The process of applying the model:

$$\text{predict} \left(\begin{matrix} x_{11} & \cdots & x_{1k} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ x_{m1} & \cdots & x_{mk} & 1 \end{matrix} \right)$$

X_{test}

W_1

Признаки = фичи + "1" как
наличие коммуникации

$$\text{predict} \left(\begin{matrix} x_{11} & \cdots & x_{1k} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ x_{m1} & \cdots & x_{mk} & 0 \end{matrix} \right)$$

X_{test}

W_0

$uplift$

Признаки = фичи + "0" как
отсутствие коммуникации

$$= \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}$$



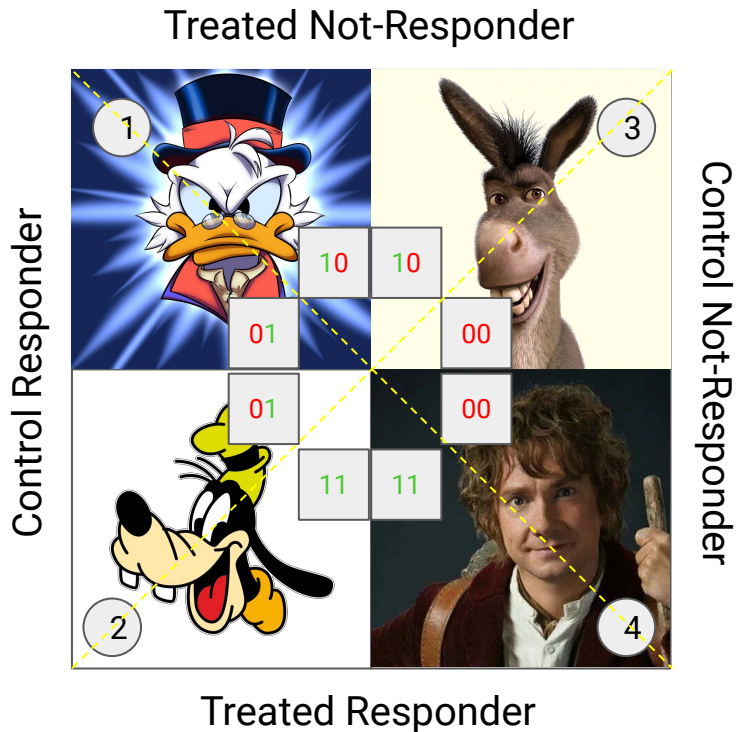
КГ ~ ЦГ

Типы клиентов



1. **Не беспокоить:** негативная реакция на коммуникации
2. **Лояльный:** всегда реагирует положительно на коммуникации
3. **Потерянный:** не совершит целевое действие, независимо от коммуникаций
4. **Убеждаемый:** совершит целевое действие только при коммуникации

Многоклассовый подход



коммуникации
не было

коммуникация
была

не купил

купил

00	01
10	11

W	Y	Y'
0	0	CN
0	1	CR
1	0	TN
1	1	TR

$$uplift = (P(TR) + P(CN)) - (P(TN) + P(CR))$$

вероятность
принадлежности
классам TR и CN

вероятность
принадлежности
классам TN и CR

Трансформация классов

W	Y	Y'	Z
0	0	CN	1
0	1	CR	0
1	0	TN	0
1	1	TR	1



Идея: собрать новый класс

$$Z_i = \begin{cases} 1, & \text{если } W_i = 1 \text{ и } Y_i = 1 \\ 1, & \text{если } W_i = 0 \text{ и } Y_i = 0 \\ 0, & \text{в остальных случаях} \end{cases}$$

$$Z_i = W_i \cdot Y_i + (1 - W_i) \cdot (1 - Y_i)$$

$$uplift = 2 \cdot P(Z = 1) - 1$$



КГ ~ ЦГ

Подробнее про вывод формулы: https://www.uplift-modeling.com/en/latest/user_guide/models/revert_label.html

Вопросы?



Ставим “+”,
если вопросы есть



Ставим “-”,
если вопросов нет



Подходы с двумя моделями

Две независимые модели T-Learner

The training process:

$$\begin{matrix} model^T = fit \left(\begin{matrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pk} \end{matrix}, \begin{matrix} y_1 \\ \cdots \\ y_p \end{matrix} \right), & model^C = fit \left(\begin{matrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{q1} & \cdots & x_{qk} \end{matrix}, \begin{matrix} y_1 \\ \cdots \\ y_q \end{matrix} \right) \\ X_{train_treat} & Y_{train_treat} & X_{train_control} & Y_{train_control} \end{matrix}$$

The process of applying the model:

$$\begin{matrix} model^T \\ predict \\ proba \end{matrix} \left(\begin{matrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mk} \end{matrix} \right) - \begin{matrix} model^C \\ predict \\ proba \end{matrix} \left(\begin{matrix} x_{11} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mk} \end{matrix} \right) = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}$$

X_{test} X_{test} $uplift$

Проскорьте клиентов
моделью целевой группы

Проскорьте клиентов
моделью контрольной группы

1. Обучить модель на данных целевой группы
2. Обучить вторую модель на данных контрольной группы
3. Чтобы получить значение uplift для тестовых клиентов:
 - a. проскорить их первой моделью
 - b. проскорить их второй моделью
 - c. найти разницу двух векторов



КГ и ЦГ могут быть
с дисбалансом

Две зависимые модели

The training process:

$$model^C = fit \left(\begin{matrix} x_{11} & \dots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{q1} & \dots & x_{qk} \end{matrix}, \begin{matrix} y_1 \\ \dots \\ y_q \end{matrix} \right),$$

$X_{train_control}$ $Y_{train_control}$

$$P^C = \begin{matrix} model^C \\ predict \\ proba \end{matrix} \left(\begin{matrix} x_{11} & \dots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{p1} & \dots & x_{pk} \end{matrix} \right), \quad model^T = fit \left(\begin{matrix} x_{11} & \dots & x_{1k} & P_1^C & y_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{p1} & \dots & x_{pk} & P_p^C & y_p \end{matrix} \right),$$

X_{train_treat} X_{train_treat} P^C Y_{train_treat}

The process of applying the model:

$$\begin{matrix} model^T \\ predict \\ proba \end{matrix} \left(\begin{matrix} x_{11} & \dots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mk} \end{matrix} \right) - \left(\begin{matrix} model^C \\ predict \\ proba \end{matrix} \left(\begin{matrix} x_{11} & \dots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mk} \end{matrix} \right) \right) = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}$$

X_{test} $P^C(X_{test})$ X_{test} $uplift$

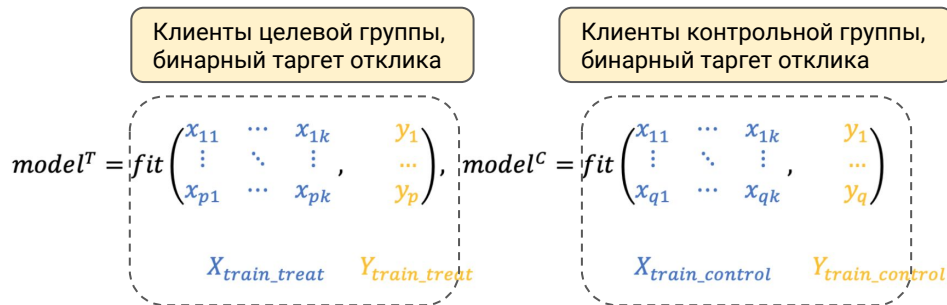
Проскорьте клиентов моделью контрольной группы, сохраните переменную, - этот предикт используется дважды

1. Обучить модель на данных **контрольной** группы
2. С помощью этой модели получить предикт на клиентах **целевой** группы
3. Обучить вторую модель на признаках **целевой** группы + **предикт** с прошлого шага как доп.фича
4. Чтобы получить uplift для тестовых клиентов:
 - a. проскорьте клиентов первой моделью
 - b. добавьте полученный предикт как фичу и проскорьте второй моделью
 - c. найдите разность пунктов **b** и **a**

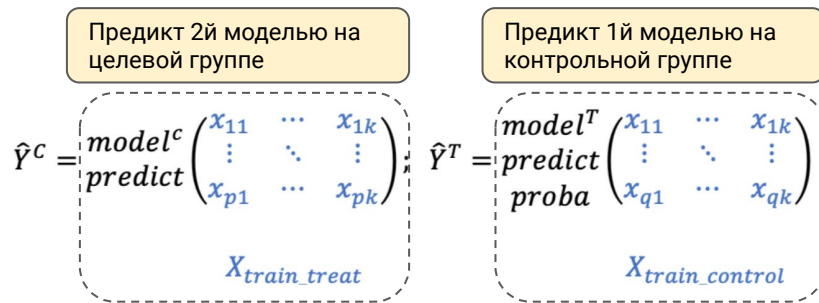


КГ и ЦГ могут быть с дисбалансом

Две зависимые модели X-learner (1)



1. Обучите независимо модели на ЦГ и КГ



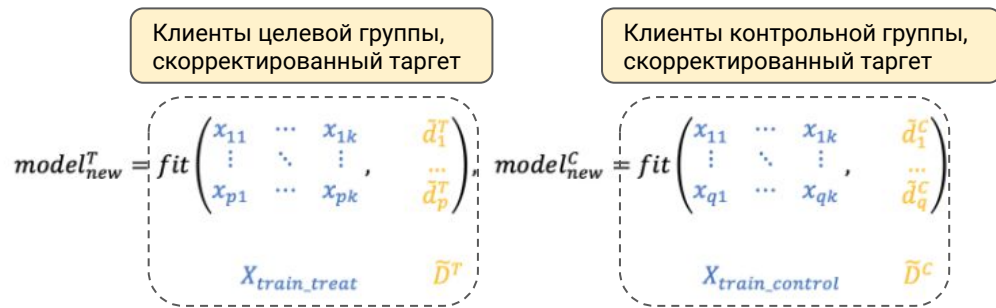
2. Преобразуйте таргет:

- проскорьте второй моделью целевую группу, сохраните вектор
- проскорьте первой моделью контрольную группу, сохраните вектор
- соберите два новых "таргета" как разность истинного и полученного в пунктах а и и

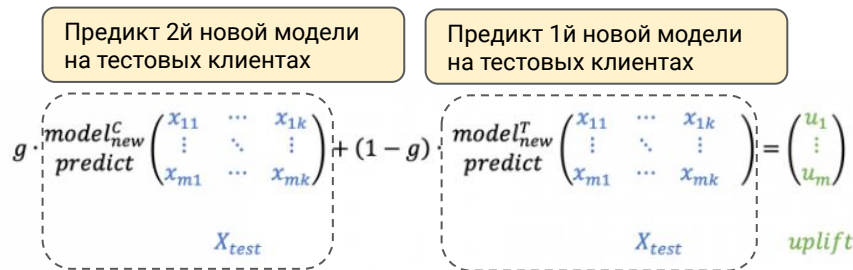
$$\tilde{D}^T = Y_{train_treat} - \hat{Y}^C; \tilde{D}^C = \hat{Y}^T - Y_{train_control}$$

Разница исходных таргетов и полученных предиктов

Две зависимые модели X-learner (2)



3. Полученные на прошлых шагах “таргеты” используйте как целевые переменные и обучите на них две независимых модели



4. Чтобы получить uplift при фиксированном g:
- проскорьте тестовых клиентов моделью контрольной группы с прошлого шага, с множителем g
 - проскорьте тестовых клиентов моделью целевой группы с множителем (1-g)
 - найдите сумму этих векторов



КГ и ЦГ могут быть с дисбалансом

Сравнение подходов

Подход	Плюсы	Минусы
Мультикласс	Позволяет учитывать сложные взаимодействия	Сложность интерпретации, небаланс классов
Трансформация классов	Обладает высокой точностью	Требует дополнительных действий по трансформации таргета
Одна модель	Простота реализации	Чувствительна к небалансу классов
Две независимые модели	Простота реализации, можно использовать в случае дисбаланса классов	Не учитывает взаимодействия между моделями
Две зависимые модели	Выше точность за счет учета взаимодействий	Более сложная реализация
Две зависимые модели X-learner	Высокая точность за счет учета взаимодействий	Самая сложная реализация



Вопросы?



Ставим “+”,
если вопросы есть



Ставим “-”,
если вопросов нет



Метрики качества

Uplift@k

	n_treatment	n_control	response_rate_treatment	response_rate_control	uplift
percentile					
0-10	3346	2656	0.733413	0.640437	0.092976
10-20	3098	2904	0.665913	0.634642	0.031272

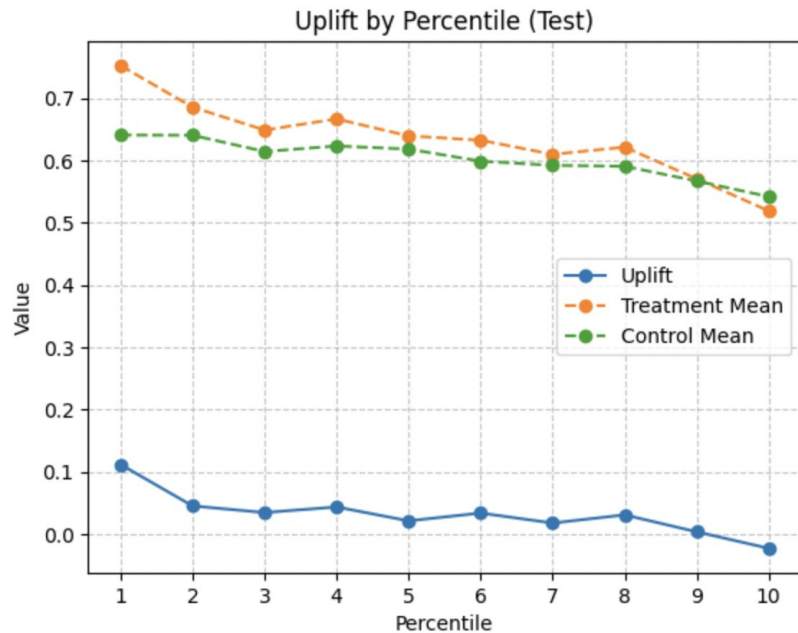
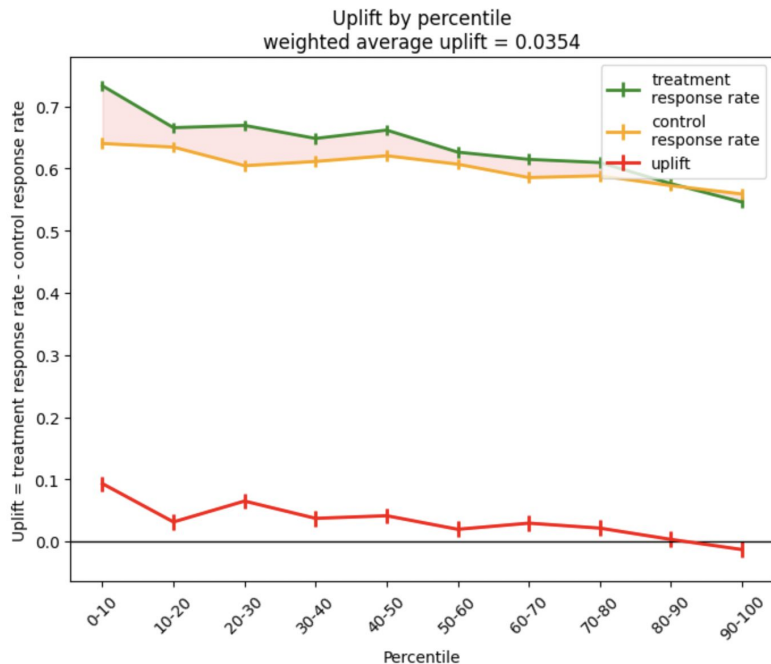
Доля откликов в
целевой группе

Доля откликов в
контрольной группе

uplift как разница
двух response_rate

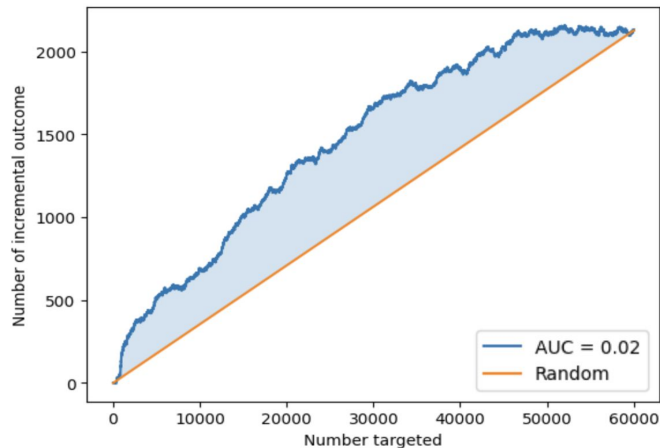
График plot_uplift_by_percentile

```
from sklift.viz import plot_uplift_by_percentile - ПЛОХО!
```

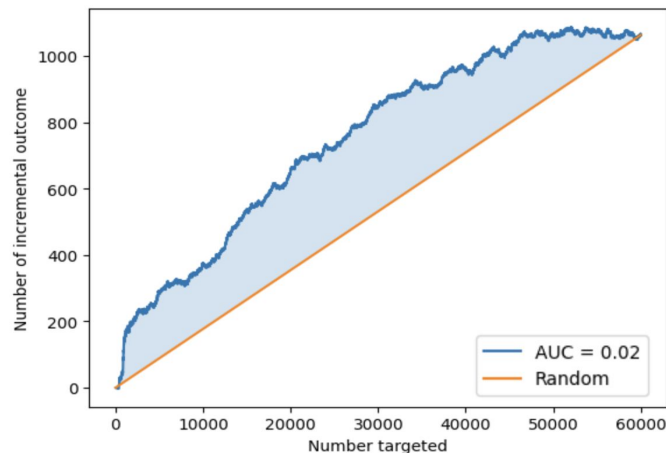


Uplift and qini curve

```
from sklift.viz import plot_uplift_curve, plot_qini_curve
```



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



$$qini_curve(t) = Y_t^T - \frac{Y_t^C \cdot N_t^T}{N_t^C}$$

t — количество объектов

Y_t^T, Y_t^C — таргет в целевой и контрольной группах

N_t^T, N_t^C — размер целевой и контрольной групп

LIVE

Домашнее задание: X-learner



На данных, рассмотренных на вебинаре, реализовать uplift-модель по алгоритму X-learner. Сравнить с моделями, реализованными на вебинаре по метрике uplift@10%

1. Повторить метод X-Learner перекрестной зависимости.
2. Загрузить данные, разбить на выборки
3. Обучите модель X-Learner без использования uplift-библиотек. Используйте для обучения моделей, например, градиентный бустинг
4. Оцените качество полученной модели X-learner на тестовой выборке при разных значениях параметра g . Сравните по метрике uplift@10% с моделями, полученными на вебинаре.
- 5*. Замерьте качество полученной модели по другим метрикам, которые вы знаете или изучили самостоятельно.
6. Оформите решение в виде notebook с наличием подробных комментариев и выводов.



Цели вебинара

Проверка достижения целей

1. формализовали задачу uplift-моделирования как набор нескольких моделей
2. изучили построение в виде одной модели
3. изучили построение в виде нескольких моделей

Вопросы для проверки

По пройденному материалу всего вебинара

1. Чем отличается задача uplift-моделирования от look-a-like и response?
2. Приведите пример подхода с одной моделью
3. Приведите пример подхода с двумя моделями
4. Какие метрики для uplift-моделирования вы знаете?



Ключевые тезисы занятия

Подведем итоги

1. Задача uplift-моделирования помогает выявить чистый эффект от коммуникации, или выявить убеждаемых клиентов
2. Для разработки могут использоваться подходы с одной или несколькими моделями
3. При обучении необходимо отслеживать качество моделей по релевантным метрикам

Полезные ссылки

1. Библиотека Double ML
<https://docs.doubleml.org/stable/intro/install.html#install-python>
2. Библиотека CausalML <https://causalml.readthedocs.io/en/latest/installation.html>
3. Обзор литературы [Causal Inference and Uplift Modeling](#)
4. Чуть более формальная статья про рассмотренные подходы [Uplift Modeling for Multiple Treatments with Cost Optimization](#)
5. Статья с продвинутыми методами uplift-моделирования [Estimating Individual Treatment Effect in Observational Data Using Random Forest Methods](#)

Рефлексия

Рефлексия



Будете ли применять на практике то,
что узнали на вебинаре?



Следующий вебинар



21 ноября 2024

Введение в рекомендательные системы



Ссылка на вебинар
будет в ЛК за 15 минут



**Заполните, пожалуйста,
опрос о занятии
по ссылке в чате**