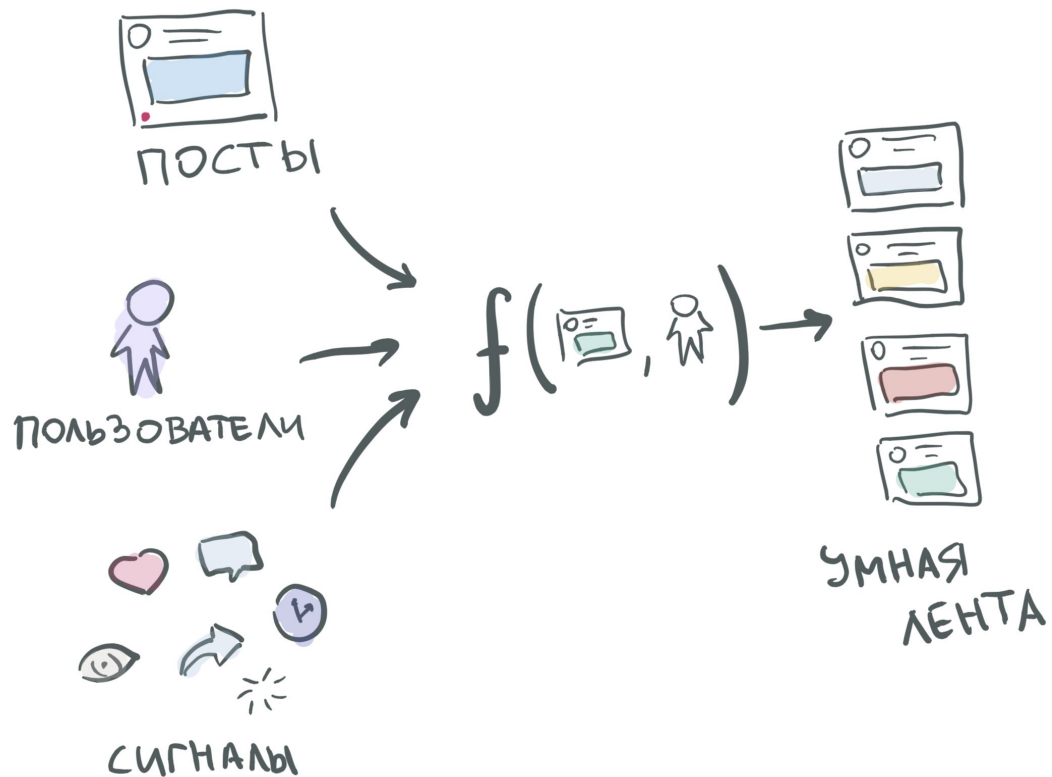


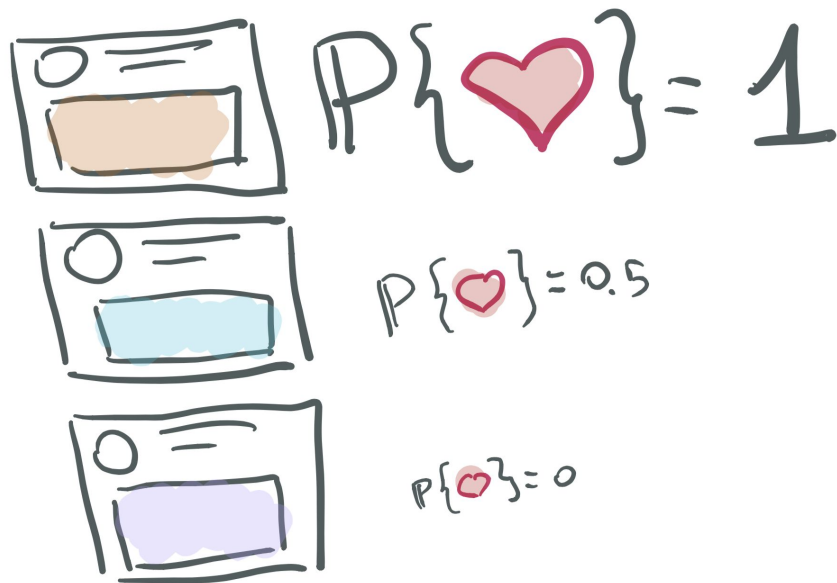
Три перпендикулярные линии или многокритериальная оптимизация в VK

Никита Маршалкин, CoreML

Действующие лица



Давайте отранжируем по лайкам





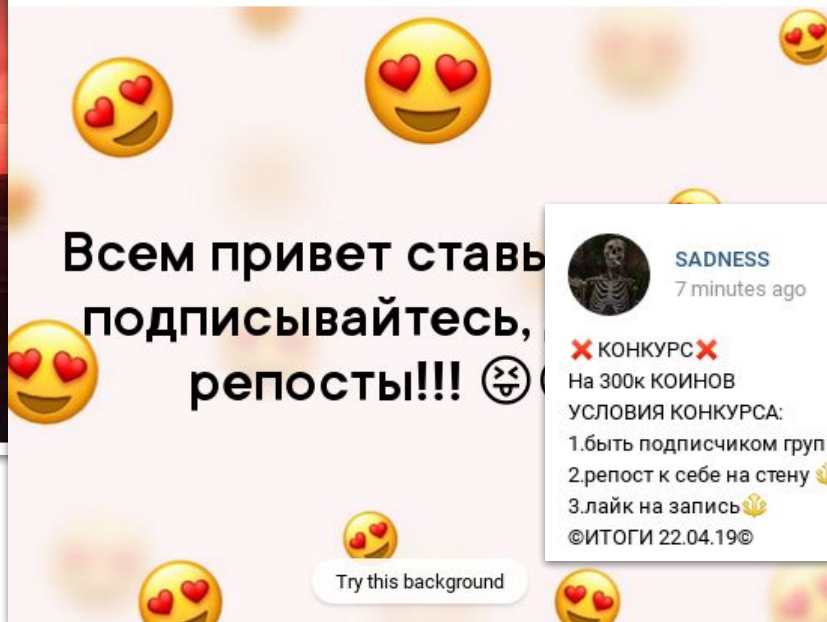
Dmitry Bykadorov
today at 11:18 pm 🌐

+ Follow

Ставь лайк если понравился закат👍



Хвостатые байки
three minutes ago



Try this background



SADNESS
7 minutes ago

❌ КОНКУРС ❌

На 300к КОИНОВ

УСЛОВИЯ КОНКУРСА:

1. быть подписчиком группы 🏆

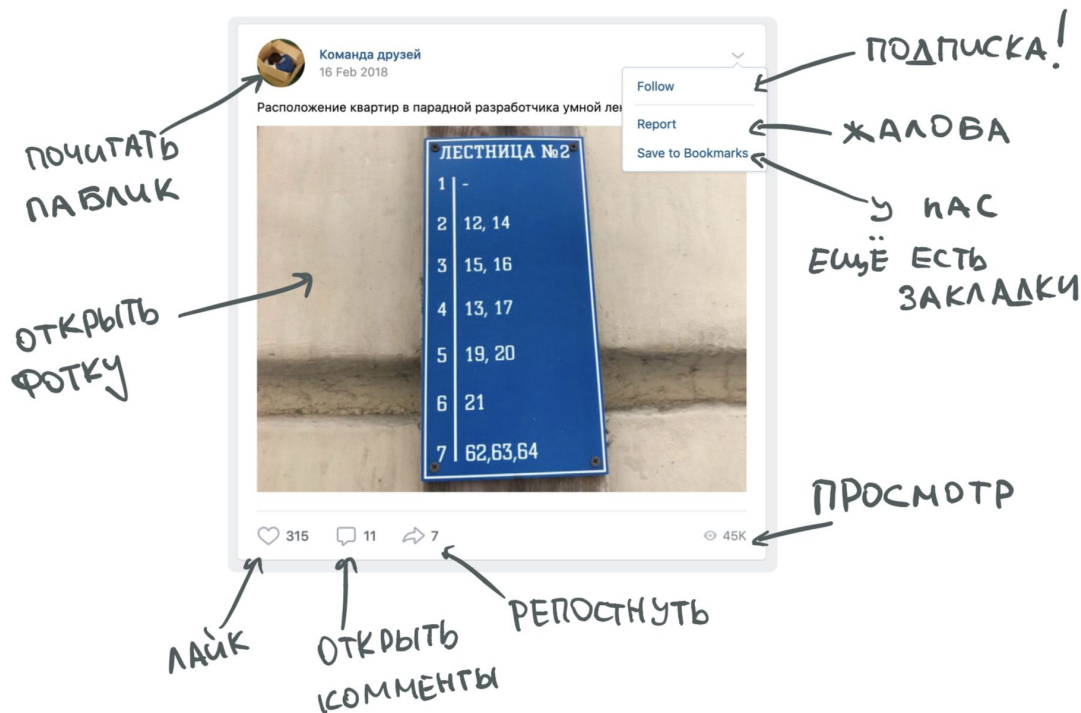
2. репост к себе на стену 🏆

3. лайк на запись 🏆

©ИТОГИ 22.04.19©

Неявные действия

⌚ ВРЕМЯ ПРОСМОТРА



Баланс в ML

$$P\{\heartsuit\} = \sigma(f(\heartsuit))$$

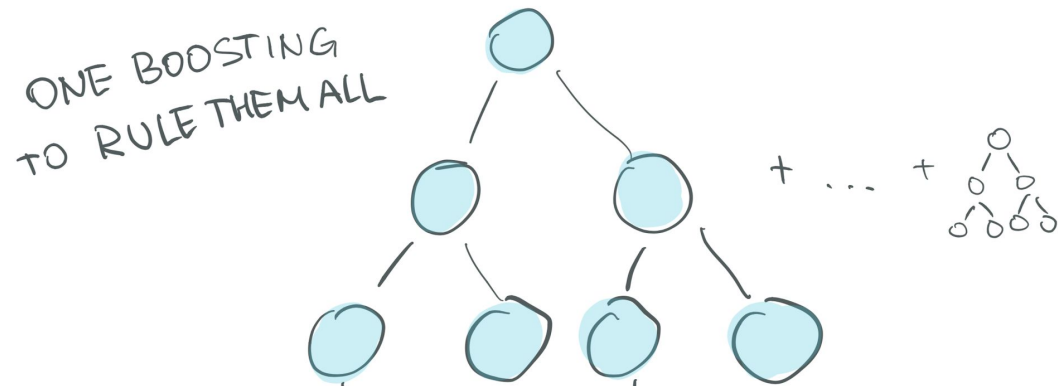
$$P\{\text{💬}\} = \sigma(f(\text{💬}))$$

$$\text{score} = \sigma(w_1 \cdot f(\heartsuit) + w_2 \cdot f(\text{💬}))$$

Кураленок, Поваров. Некоторые способы построения целевой функции

Tang et al., An Empirical Study on Recommendation with Multiple Types of Feedback

Быстрый баланс в ML



$$P\{\heartsuit\} = G(s_L) = G(w_1 l_1 + w_2 l_2 + w_3 l_3 + w_4 l_4)$$

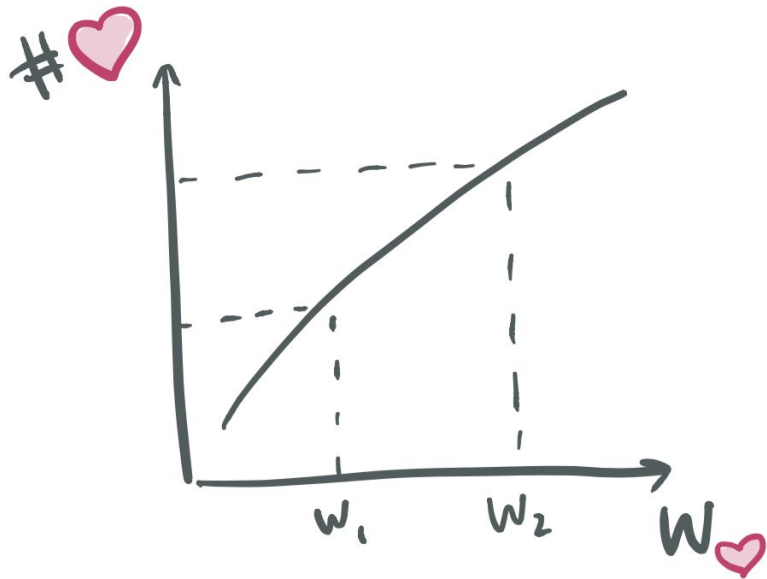
$$P\{\spadesuit\} = G(s_C) = G(v_1 l_1 + v_2 l_2 + v_3 l_3 + v_4 l_4)$$

$$G(as_L + bc_C) = G((aw_1 + bv_1)l_1 + (aw_2 + bv_2)l_2 + (aw_3 + bv_3)l_3 + (aw_4 + bv_4)l_4)$$

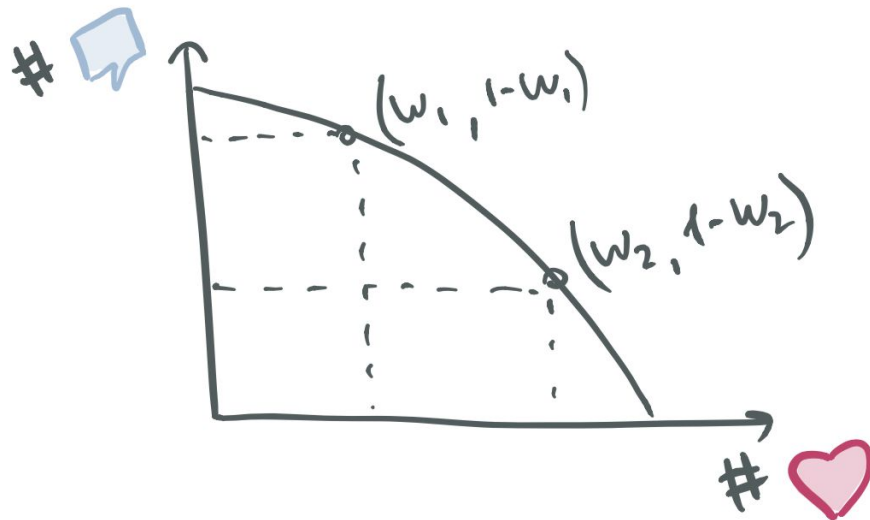
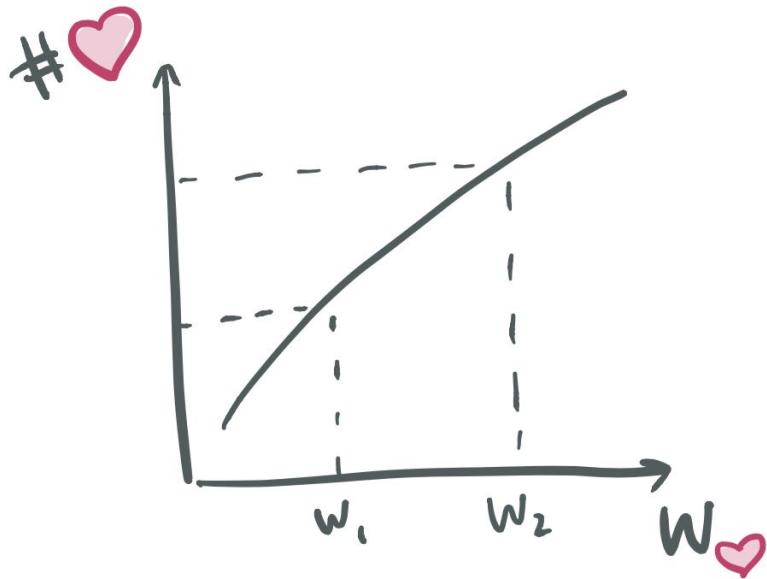
Ling et al., Model Ensemble for Click Prediction in Bing Search Ads



Какие веса брать? $\sigma(w_1 \cdot \{\text{❤️}\} + w_2 \cdot \{\text{💬}\})$



Какие веса брать? $\sigma(w_1 \cdot \{\text{❤️}\} + w_2 \cdot \{\text{💬}\})$



Какие веса брать?

- Растить все клики одновременно не получится - парето фронт
- Не знаем ценность кликов - не скаляризировать

Нужен один критерий...

TIME SPENT!

- + Дольше сидит - счастливее!
- + Время - деньги!
- + Сложно заабыюзить!
- + Проводит время - значит любит!

Быстро!

Качественно!

Недорого !

- Нельзя напрямую оптимизировать в ML
- Не самая чувствительная штука

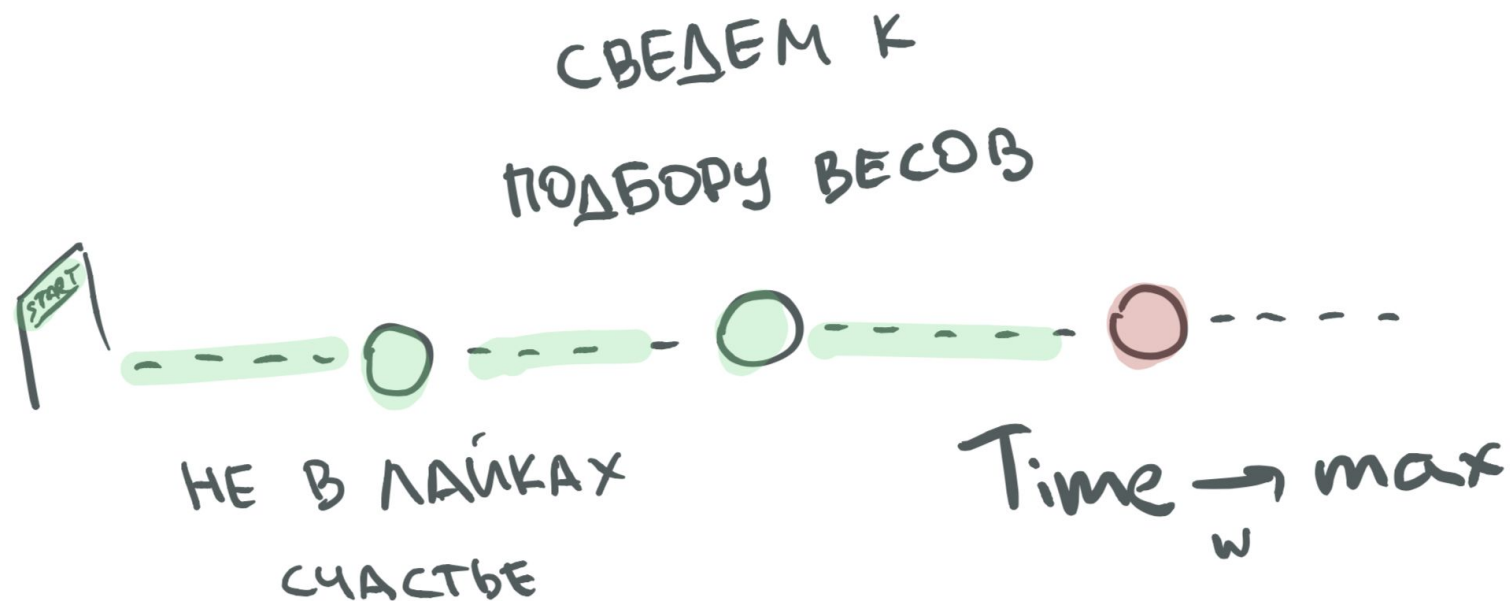
Метрики

- **Time Spent:** сессии на пользователя, время между сессиями, время сессии в постах, пикселях, минутах.
- **Клики:** лайки, комменты, скрытия, репосты, гейзы. CTR, позиции первого клика, время до клика.

Техника экспериментов

- Метрики на уровне пользователя, t-test в большинстве, иногда Mann-Whitney
- Delta method для процентов
- CUPED для чувствительности
- Фиксированные слоты + иногда помешиваем
- Следим за FP на AA
- Когда стремно, катим через обратные эксперименты

Где мы



Параметры всюду

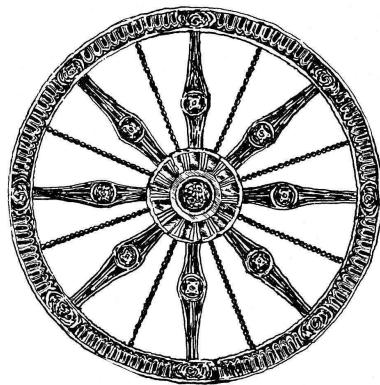
- Конфигурация кандидатов, доли
- Окна/гравитация в разнообразии
- Трешхолды в фильтрах
- Коэффициенты бустов/занижений
- Баланс контента
- Эвристики переранжирования
- Вставка блока рекомендации друзей
- Версии эмбедингов
- Количество постов на разных этапах
- Блоки “Вам может быть интересно”
- Критерии показа инлайн комментариев



Ручной подбор параметров

ВЫБИРАЕМ ВЕСА

(интуитивно, aka экспертное мнение)



ДУМАЕМ

СТАВИМ
АБ-ТЕСТ

ЖДЕМ 1-7 дней

Ручной подбор параметров



Одна гипотеза за раз :(

Полуавтоматический подбор параметров



Много гипотез за раз :)

Полуавтоматический подбор параметров



①

ДЕТЕРМИНИРОВАНО ОТОБРАЖАЕМ
В ПАРАМЕТР

$$w_i = f(\text{md5}(\text{uid}))$$

Полуавтоматический подбор параметров

$$\text{ЛЕНТА}(\text{человек}, w) =$$


2

СТРОИМ
ЛЕНТУ С ПАРАМЕТРОМ

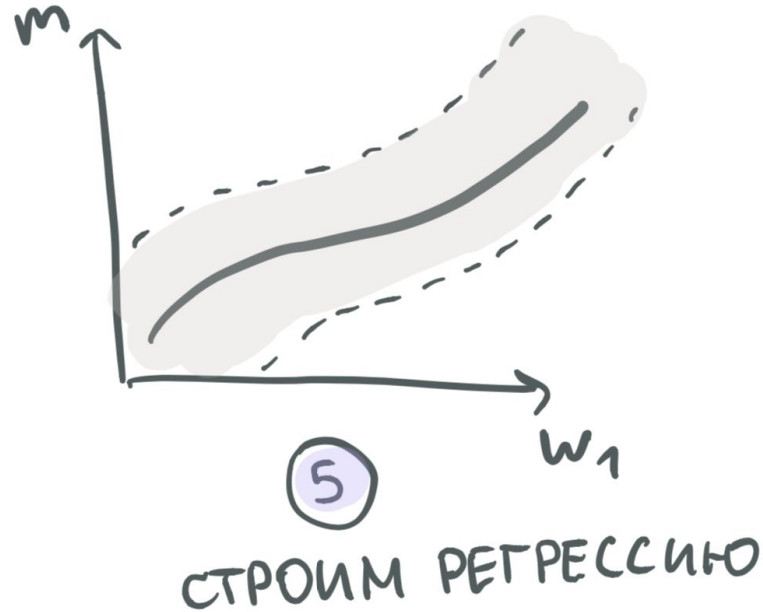
Полуавтоматический подбор параметров



③

СЧИТАЕМ
метрики по пользователям

Полуавтоматический подбор параметров



Полуавтоматический подбор

Модель должна:

1. Давать степень уверенности
2. Уметь отклонять зависимость
3. Быть гибкой (непараметрической?)

Полуавтоматический подбор параметров

- $SparseGP^1$, если параметров немного и есть непрерывность
- $RF + jackknife^2 + PDP^3 + SHAP^4$, если параметров много и они разные

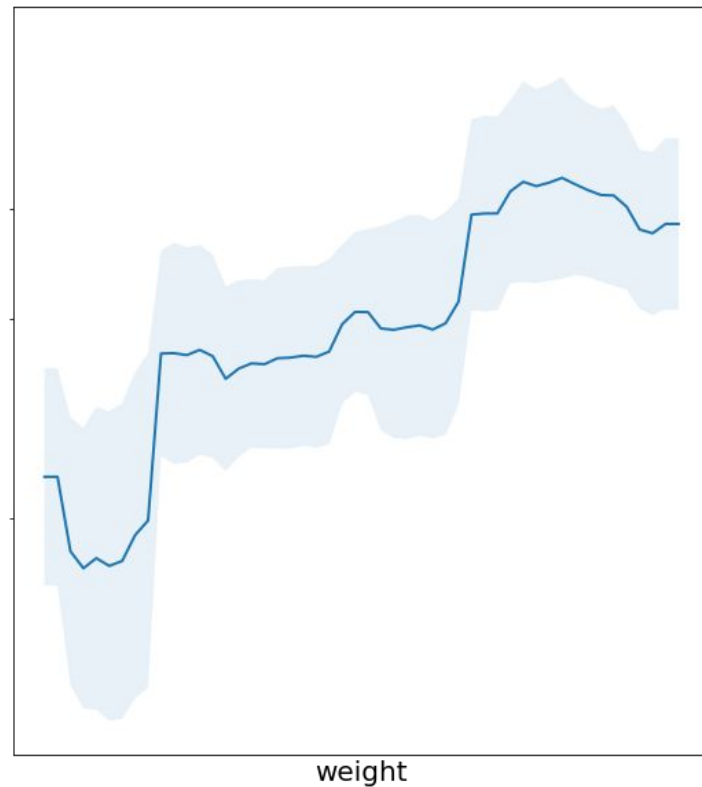
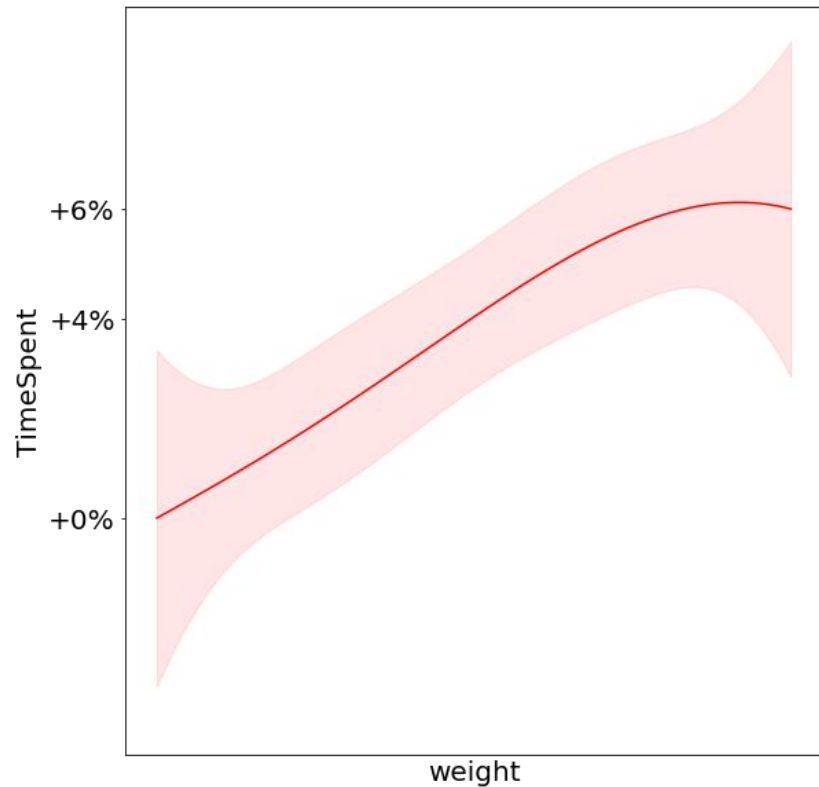
¹Titsias, Variational Model Selection for Sparse Gaussian Process Regression

²Wager et al., Confidence Intervals for Random Forests: The Jackknife and the Infinitesimal Jackknife

³Friedman, Greedy function approximation: A gradient boosting machine

⁴Lundberg, Consistent Individualized Feature Attribution for Tree Ensembles

Полуавтоматический подбор параметров



А еще

Можно домешать соцдем, историю, кластеры пользователя, метрики до эксперимента, префиксы, суффиксы, время захода в эксперимент и многое другое...

$$m = f(w) + \epsilon_m$$

$$\hat{m} = g(\underbrace{\text{человек}, \text{History}, \dots}_{\text{не зависит от } w}) + \epsilon_{\hat{m}}$$

не зависит от w

$$m := m - \hat{m} + \mathbb{E} \hat{m}$$

А что еще?

- Заменить оптимизацию времени на комбинацию кликов без пула экспериментов
- Насемплировать $\text{argmax}(m(w))$ из апостериорного распределения

Слайд с выводами

- Корректно поставленная задача - роскошь
- Лучше решать правильную задачу плохо, чем неправильную хорошо
- Онлайн > офлайн

Тут слайды



В CoreML@VK
летняя стажировка
Пишите, будет
кайфово.

vk.com/cmdf

Сюда писать



vk.com/support

press@vk.com

А еще (2)



А еще (2)

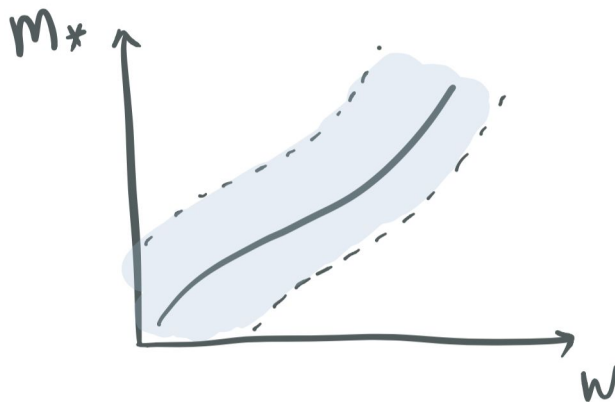
Можно предсказывать изменение времени изменением кликов

$$\mathbb{E}_{\text{exp}}(m_{ts} \mid m_0 \dots m_n) = g(m_0 \dots m_n)$$

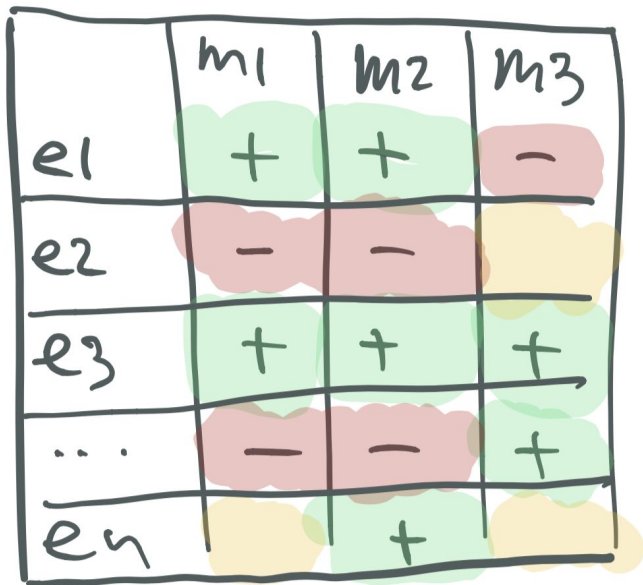
А еще (2)

Можно предсказывать изменение времени изменением кликов

$$m_* = g(m_0 \dots m_n)$$



Пул экспериментов



A hand-drawn table with 5 rows and 4 columns. The columns are labeled m1, m2, and m3. The rows are labeled e1, e2, e3, ..., and e4. The cells contain '+' or '-' signs. The background of the cells is colored: green for '+' and red for '-'. There are also yellow highlights in some cells.

	m1	m2	m3
e1	+	+	-
e2	-	-	
e3	+	+	+
...	-	-	+
e4		+	

Соберем все эксперименты за последние N лет

Найдем, с чем меняется Time Spent

Kharitonov et al., Learning Sensitive Combinations of A/B Test Metrics

Deng et al., Data-Driven Metric Development for Online Controlled Experiments: Seven Lessons Learned

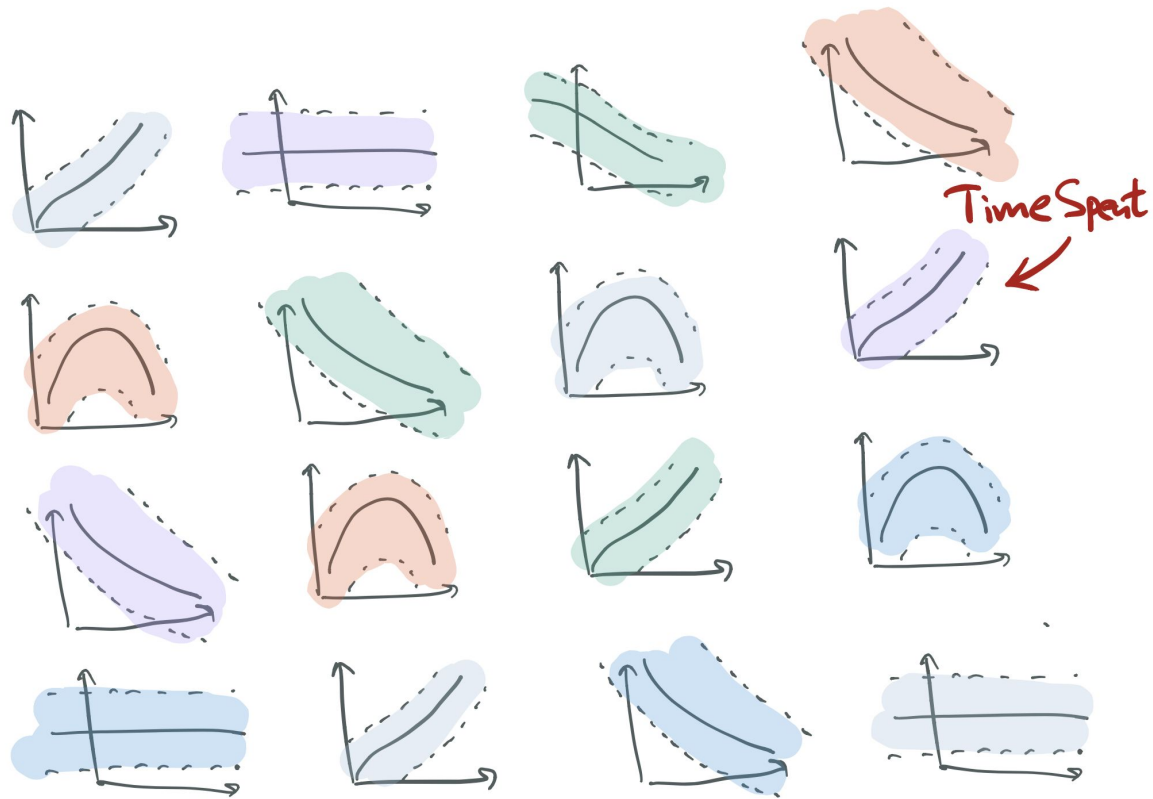
Параметры всюду

- Конфигурация кандидатов, доли
- Окна/гравитация в разнообразии
- Трешхолды в фильтрах
- Коэффициенты бустов/занижений
- Баланс контента
- Эвристики переранжирования
- Вставка рекомендации друзей
- Версии эмбедингов
- Количество постов на разных этапах
- Блоки “Вам может быть интересно”
- Критерии показа инлайн комментариев

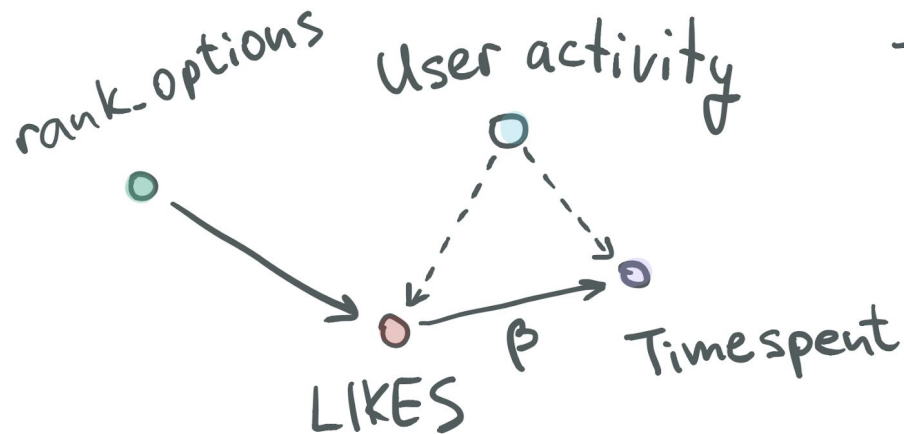
Метрики

- **Time Spent:** сессии на пользователя, время между сессиями, время сессии в постах, пикселях, минутах.
- **Клики:** лайки, комментарии, скрытия, репосты. CTR, has, позиции первого клика, время до клика. Гейзы, долгие клики

Куча метрик



Instrumental Variables & TSLS



$$LIKES = f(r-o) + \epsilon_L$$

$$\overline{LIKES} = \widehat{f}(r-o)$$

$$TS = g(\overline{LIKES}) + \epsilon_{TS}$$

Stock et al., A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments

Peysakhovich et al., Learning causal effects from many randomized experiments using regularized instrumental variables