

Продуктовые и ML метрики

Зачем нам о них знать?

Позволяют понять:

- Как связана операционная (бизнесовая) и ml-метрики?
- Какую ml-метрику оптимизировать?
- Какую ml- метрику не уронить?
- Какие доп метрики важны?
- Какое значение метрики выбрать в качестве достаточной?

1. Определения

Продуктовая метрика – метрика, отражающая качество и финансовую эффективность продукта в зависимости от событий/действий пользователей. Наиболее часто встречаются следующие метрики:

- WAU – среднее за неделю количество активных пользователей
- MAU - среднее за месяц количество активных пользователей
- DAU - среднее за день количество активных пользователей
- Revenue – оборот
- Profit – прибыль = выручка - издержки
- Retention – удержание клиентов = # пользующихся клиентов в сервис в день / # клиентов, пользующихся сервисом N дней назад
- MRR - ежемесячная регулярная сумма дохода от пользователя (когда бизнес построен на подписке)
- New users – # новых клиентов
- Intensity – сколько действий в сервисе делает клиент в месяц
- Conversion – конверсия (в воронке). Например, из добавления в корзину в покупку.
- Retention rate – возвращаемость клиента = (# клиентов на конец периода - # новых) / # клиентов на начало
- Churn Rate – отток = # ушедших клиентов в последний месяц/ # всех оплативших за последний месяц
- Leave – количество ушедших из сервиса клиентов = #неактивных клиентов / #зарегистрированных
[https://vk.com/away.php?to=https%3A%2F%2Fdata36.com%2Fwp-content%2Fuploads%2F2019%2F10%2Fdata_dictionary_tomi_mester.pdf&cc_key=]
- Win-back - количество клиентов, вернувшихся в сервис после длительного отсутствия
- Frequency – частота использования – среднее количество активных действий в продукте
- ARPU – средняя выручка с пользователя (вся выручка / всех пользователей)
- ARPPU – средняя выручка с платящего пользователя (вся выручка / всех платящих пользователей)
- OPEX – постоянные расходы на обеспечение работы компании
- CAPEX — затраты на покупку и модернизацию основных средств
- CAC (CPA, CPL) – стоимость привлечения 1 клиента
- cLTV – доход с 1 клиента за все время
- ROI – эффективность стратегии привлечения (доля затрат на привлечения 1 клиента возвращается компанией) = (LTV-CAC)/CAC
- CSI – удовлетворенность клиента
- Liquidity Ratio - способность платформы обеспечивать ликвидность и выполнение транзакций = sum транзакций/sum средств на платформе
- Gross Merchandise Volume, GMV - Общий объем продаж
- Network Effects Strength - Сила сетевых эффектов (коэффициенты Меткала) - как изменение числа участников влияет на ценность платформы для каждого из них
- Repeat Purchase Rate - процент клиентов, совершающих повторные покупки на платформе

1. Определения

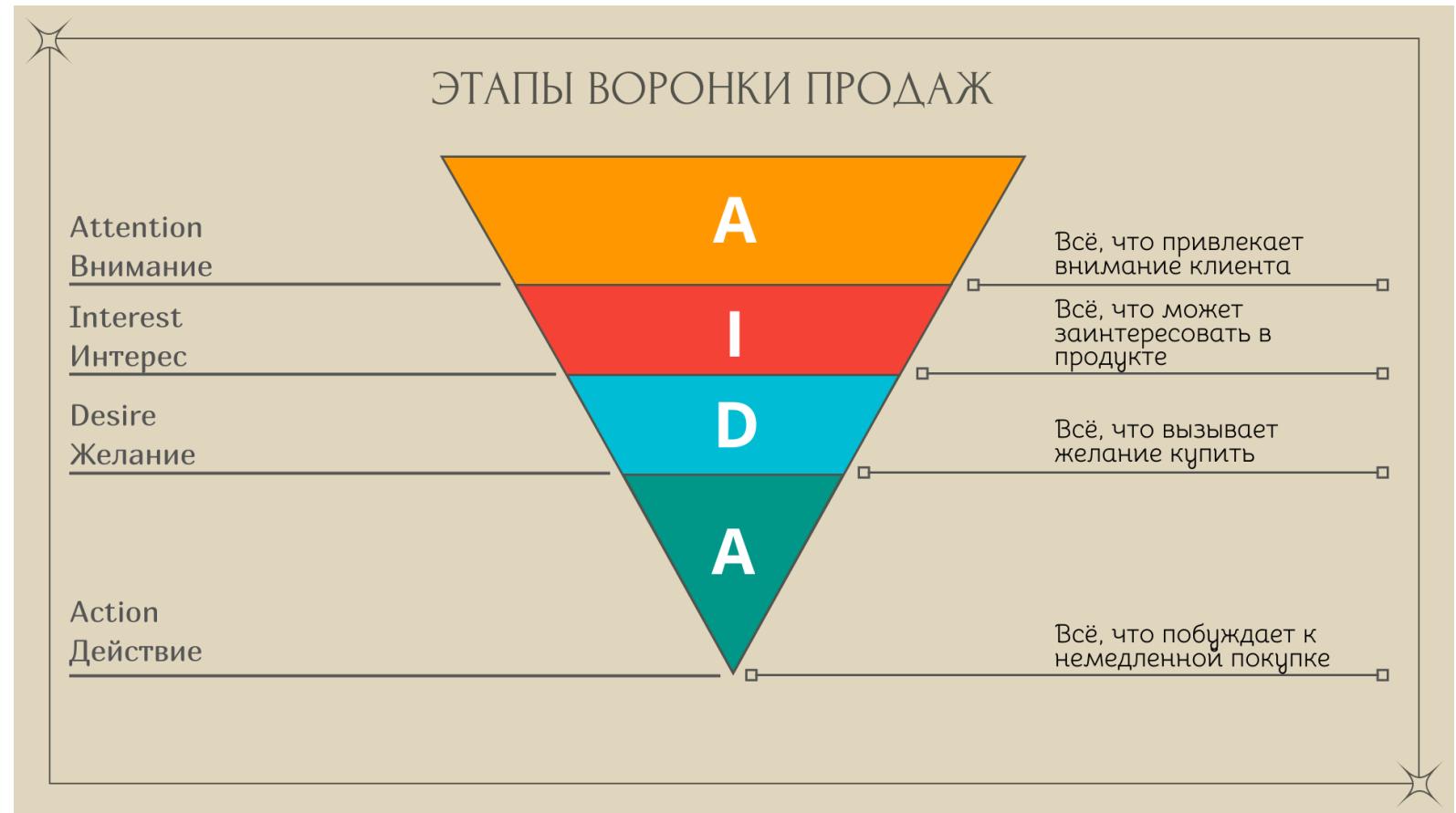
Для соц. проектов:

- 1. Социальный эффект (Social Impact). Оценка воздействия продукта на целевую аудиторию и общество в целом. Это может включать в себя изменения в образе жизни, поведении, здоровье, образовании и других аспектах.
- 2. Покрытие (Reach). Количество людей или сообществ, которые используют продукт или получают доступ к его услугам. Это помогает оценить, насколько широко продукт распространен и какое количество людей он достигает.
- 3. Вовлеченность (Engagement). Уровень вовлеченности пользователей в использование продукта. Это может включать в себя частоту использования, время, проведенное в приложении или на сайте, а также уровень активности и взаимодействия с функциями продукта.
- 4. Изменения в поведении (Behavior Change). Оценка того, насколько успешно продукт влияет на изменение поведения или привычек своих пользователей в желаемом направлении. Например, уменьшение потребления вредных продуктов, повышение активности или участие в образовательных программах.
- 5. Качество жизни (Quality of Life). Оценка влияния продукта на улучшение качества жизни своих пользователей. Это может включать в себя аспекты, такие как улучшение доступа к услугам здравоохранения, образованию, жилью, питанию и другим базовым потребностям.
- 6. Эффективность решения проблемы (Problem-solving Effectiveness). Оценка того, насколько успешно продукт решает конкретные социальные проблемы или потребности своих пользователей. Это может включать в себя сокращение бедности, борьбу с неравенством, улучшение доступности образования и многое другое.
- 7. Устойчивость и долгосрочность (Sustainability). Оценка устойчивости и долгосрочного воздействия продукта на общество и окружающую среду. Это включает в себя его способность продолжать действовать и принести пользу на протяжении длительного времени.

2. Выбор бизнес - метрики

Метрики нужно выбирать исходя из целей бизнеса.

Например, Retention для службы спасения не актуален. Кроме того, бизнес – модель требует, чтобы cLTV был сильно больше, чем САС (иначе продукт будет убыточен). Поэтому САС, уровень персонального подхода и конверсию считают на каждом шаге воронки.



Выбор бизнес - метрики

Выбор метрики зависит от:

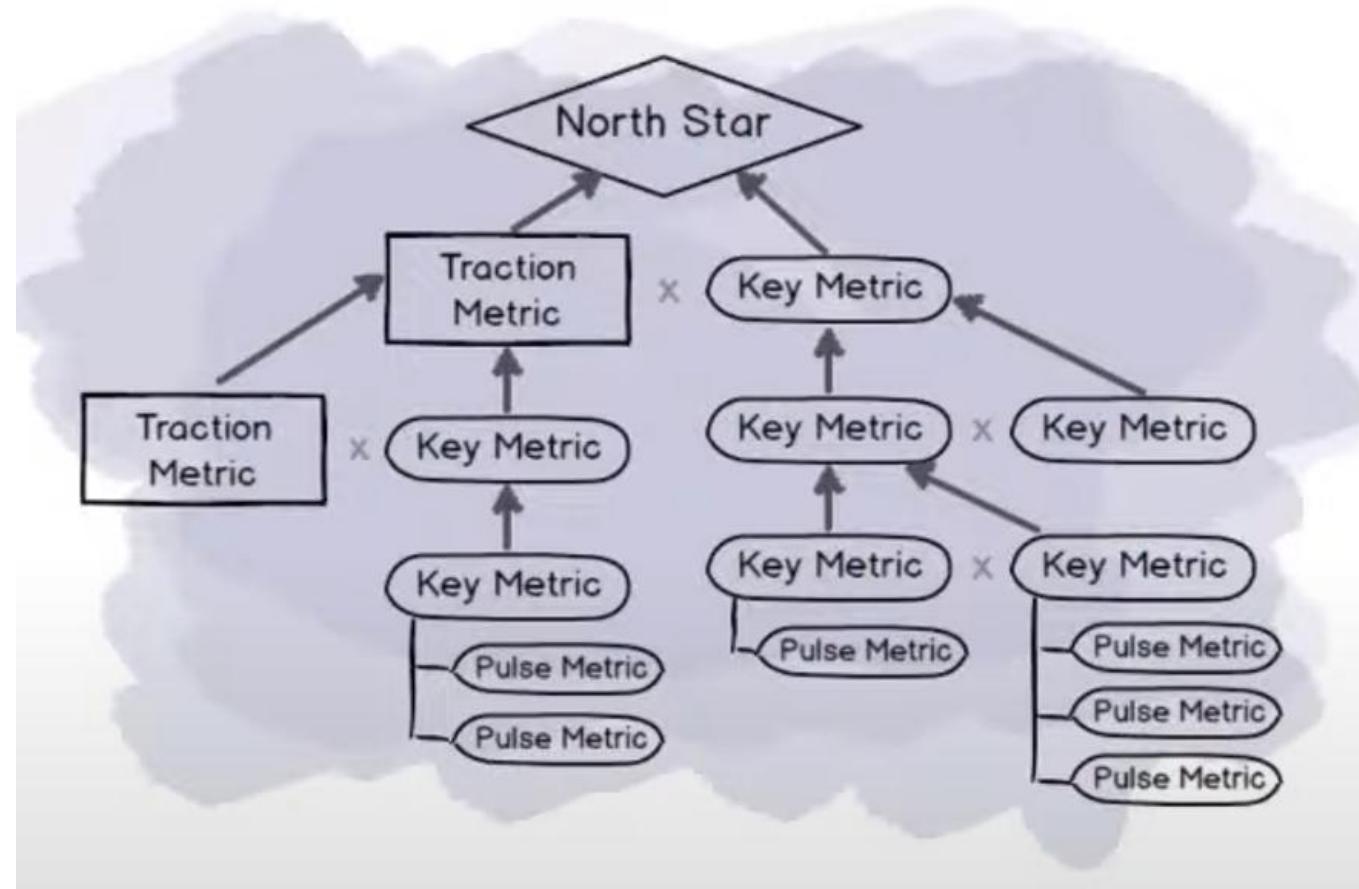
- Тип продукта
- Этап развития продукта
- Бизнес-цели (в каждой ситуации – свои)

Тип	Метрики	Заметки	Примеры компаний
Software as a service (SaaS)	MRR/ARR, Growth Rate, Net revenue retention, CAC	Все преимущества регулярного дохода, B2B обычно	Adobe, Zoom, Slack
Transactional	GTV, Net revenue, user retention, CAC	Обычно это финтех – компании и платежные системы. Успешные компании за счет большого трафика получают высокий стабильный доход	Stripe, Coinbase, Wise
Marketplaces	GMV, Net revenue, user retention, Growth Rate	Трудно сдвинуться с мертвой точки, но если бизнес начал работать, то он занимает доминантное положение на рынке.	Airbnb, Doordash, Etsy
Subscription	MRR/ARR, Growth Rate, User retention, CAC	Возвратный доход – самый ценный, часто B2C, часто ежемесячно	Amazon Prime, classpass, Netflix
Enterprise	Booking, Revenue, ACV, Pipeline	Мало клиентов, крупные сделки, покупатель не всегда потребитель.	Workday, sap
Usage Based	Monthly revenue, Growth Rate, Revenue Retention, Gross margin	Развитие по мере роста клиентов	Aws, twilio, algolia
E-commerce	Monthly revenue, Growth Rate, Gross margin / Unit Economics , CAC	Необходимо привлекать новых клиентов и управлять unit/operations-экономикой	Amazon, warby parker, D2C, Spotify
Advertising	DAU, MAU, User retention, CPM/CPC	Клиент – рекламодатель, а не пользователь; нужно большое количество показов рекламы каждый месяц	Facebook, twitter, reddit
Hardtech/ Bio/ Moonshots	Milestones, Signed contracts, LOIs	Необходимо много времени, чтобы получить готовый продукт из-за регуляторных или технических рисков	Boom, Pfizer, cruise
For internal usage	CSI	В любом случае ресурсы тратятся на создание собственной системы.	Любой продукт для внутреннего использования
Data	?	Законодательство должно разрешать подобное использование агрегированных данных.	Консультации партнерам, дашборды для рекламодателей (Meta), для магазинов – рекламодателей (Zalando)
Income share agreement (ISA)	?	Y combinator	Образовательные платформы

Этап	Критерий успешности	Что считаем
Выход на рынок	x пользователей в день / у покупок / z отзывов положительных	# новых пользователей, # покупок, # установок в день; общее # установок, все конверсии; отзывы на открытых ресурсах; NPS (удовлетворенность пользователей сервисом)
Стабилизация аудитории	retention 30 дней x %; z клиентов совершили покупку 1 раз	DAU, WAU, MAU; N-day retention (n=7, 14, 30); churn rate; # активных (лояльных) клиентов
Наращивание аудитории	Рост # новых клиентов за неделю x %, z% органических пользователей тратят менее 5000 руб. на одну покупку.	# новых пользователей в день/неделю; средний чек; продажи за периоды VS прошлые периоды; полная воронка продаж с конверсиями на каждом этапе
Насыщение рынка	Увеличение среднего чека на % процентов ежемесячно, z % клиентов совершили покупку 1 раз	Общая выручка продукта; LTV, Churn, retention; ARPU, ARPPU; Стоимость льда (CPL)
Смена парадигмы	Цикл начинается заново	

3. Дерево метрик

- Чаще всего выбранные продуктовые метрики связаны между собой и представимы через другие метрики, операционные.
- Для структуризации и визуализации этих связей формируется дерево метрик. Оно требует упорядочивания всех метрик для продукта в целом и конкретного business-unita.
- Вершиной дерева является ключевая метрика качества продукта North Star Metric.



Дерево метрик

Принципы построения дерева:

- Иерархичность
- К метрикам уровня l применяются математические операции для получения метрик $l+1$
- Иерархия метрик = ценность продукта, описанная через подметрики
- Позволяет видеть прозрачную картинку
- Имеет смысл, где есть дашборды старше года (для отслеживания исторических данных)

При грамотном формировании дерева на одном из уровней можно будет встретить метрику успеха для подпроекта, где реализовано ML-решение, используемое в продакшене. Так метрика ML модели будет напрямую влиять на какую-то операционную метрику, которая влияет на качество продукта в целом. Однако низкое качество модели несет в себе репутационные риски, которые конвертируются в отношение клиента к бренду, под которым работает продукт. Сильный бренд в свою очередь позволяет сокращать расходы на привлечение и удержание клиентов, поэтому плохое качество ML-модели не только в текущем моменте уменьшает прибыль с клиента, но и увеличивает затраты на его удержание/привлечение.

Дерево метрик

Например, для сервиса ответа на вопросы и последующей продажи в рамках большой компании упрощенное дерево метрик может выглядеть так:

- BREND Value & accuracy < - ML-model
- REVENU <- LTV – CAC
- LTV = conversion * accuracy * ...
- CAC = manager cost + external marketing cost + ...
- external marketing cost = TV marketing cost + Radio marketing cost + ... - BREND Value

Ошибки модели могут по-разному влиять на бизнес-метрики: если неверный ответ модели является ограничением покупательской способности клиента, то это оказывается на недополученной прибыли компании.

Например, модель консервативна и на вопросы об условиях выдачи кредита выдает информацию с завышенными требованиями (например, упустив исключения), то человек не будет даже пытаться подать заявку (и передавать свои личные данные). Поэтому в данном случае необходимо детализировать ошибку и изменять ML-метрику.

Выбор прокси-метрик для ML-решения

Требования:

- Прокси-метрики должны быть связаны с долгосрочным целевым показателем, например, с LTV. Как минимум, должно быть понятно, в какую сторону изменится показатель при изменении метрик;
- Хорошие прокси-метрики сложно улучшить, делая «неправильные» вещи. Например, в YouTube вместо конверсии в клики (CTR) используют метрику long CTR, где клик засчитывается, только если пользователь просмотрел существенную часть видео. Это связано с тем, что CTR можно увеличить, используя «кликбейт» в названии и вводя пользователя в заблуждение о содержании ролика, что в итоге снижает удовлетворенность сервисом;
- Прокси-метрики должны быть чувствительны к изменениям, которые влияют на долгосрочный целевой показатель. Например, пользователи отдельного типа устройств могут существенно влиять на долгосрочный целевой показатель, но увеличить эту долю довольно сложно;
- Вычисление прокси-метрик не требует много ресурсов. К примеру, в качестве таких метрик не подойдут опросы общественного мнения;
- Критерий, сформированный на основе прокси-метрик, может учитывать новые сценарии. Пример негибкого критерия — в магазине продавали телевизоры и взяли метрику, основанную на диагонали экрана. Затем в ассортименте появились пылесосы, к которым эта метрика не применима, поэтому ее нельзя использовать для учёта нового вида товара.

Увязка создания ценности и метрик



1 – стоимость подписки на сервисы GigaPlatform, амортизация стоимости разработки, стоимость поддержки агента

Альтернативы деревьям

Проблемы:

- Слишком **сильно упрощает взаимосвязи между метриками** -
Локальные изменения менее заметны:

Выручка = ARPPU * Число клиентов (за период)

Решение: Смотреть отдельно по когортам (частота использования в рамках конкретного Use Case)

- **Фокус на локальной оптимизации:** дерево создает иллюзию, что есть волшебная таблетка, чтобы синизить SAC, увеличить LTV или поднять C1. Но на деле это не так;
- — **Сложность адаптации к изменениям:** составить дерево — полбеды, а вот его поддерживать и анализировать — реальная проблема;
- — **Неприменимость** в реально **больших продуктах** как следствие предыдущего пункта.
- Решение: модель роста (Growth Model)

Growth Model

В модели роста заложено все (с нужной именно сейчас детализированностью):

- Удержание
- Монетизация
- Привлечение
- Рост

- ✓ Дает ответ на вопрос: «что будет, если я улучшу удержание на 7-й день на 20%»
- ✓ Можно найти точки роста, аргументировать фичи
- ✓ Можно найти узкие места и креативно их устарнить (Growth hacking)

Реализация:

- Качественная: схема
- Количественная: таблица с формулами

Ai-решение может

Core LLM

Напрямую влиять на прибыль – если влияет на монетизацию

- Генерировать более качественные лиды
- Предлагать трейдеру более выгодные вложения
- Быстрее принимать решения

Можно считать unit-экономику => прямая связь с качеством ml

Косвенно влиять на прибыль – остальные элементы

- Улучшать только качество услуги, если прибыль идет с рекламы, то качество решения увеличивает поток, а уже поток уже увеличивает прибыль

Связь не с unit экономикой, а только с macro-метриками => if-else связь

Не влиять на текущую прибыль – т.к. бизнес работает хорошо и без модели

- Не нужная фича
- Фича развита гораздо лучше у конкурентов
- Нет средств для поддержания
- Не укладывается в стратегию компании

Репутация – связь только с брендом => метрики из бенчмарков

Монетизация

ратить продукт в денежную машину

service (SaaS), Enterprise, Data, E-commerce - Прямые продажи : юзер
раз, получает продукт навсегда.

- **Подписка:** платит регулярно, пока пользуется.

аза бесплатно, а за полезные фичи придется платить (один раз /
Фичи могут быть бесплатными в демо-периоде

I, Usage Based, Marketplaces - Транзакционная модель: берешь % с
ации / кейсов использования. Чем больше операций, тем больше денег.

платит рекламодатель. Обычно стоимость размещения рекламы
льна квантам потоку просмотров / подписчиков / ...

o/ Moonshots – Большой и дорогой Research, потом одна из моделей выше
(прямые продажи)

agreement (ISA) – Как прямые продажи, но оплата после достижения

sage – ЗП, не больше, не меньше

Не любая монетизация = конверсия из качества ml-модели

Важна красивая презентация, cherry-pick
Прямое влияние на репутацию
Чем более редкие продажи, тем меньше
связь качества ml с метрикой

Прямое влияние

Во время демо – прямое.
Далее – в зависимости от модели
монетизации

Прямое влияние

Косвенное влияние

В зависимости от модели монетизации

Во время agreement можно отказаться от
пользования продукта => прямое

Важна красивая презентация, cherry-pick
Если есть Валидация, то прямое (-премия
или KPI)

Заработка на Advertising - уточнение

- Просто баннеры / виджеты от «рекламных мсорубок» / стать частью чужого омниканального продвижения - косвенное core-сервиса
- Персонализация рекламы по цифровому следу отдельной моделью внутри сервиса / uplift-моделирование – прямое влияние рекламной модели, косвенное core-сервиса
- Выступать как «лидер мнений» - репутация core-сервиса

Как выбрать метод монетизации?

- **Аудитория.** Если продукт для корпораций, они легко заплатят 100к в месяц. Если для студентов — даже 100 рублей могут быть много.
- **Анализ рынка.** Если все вокруг используют подписки, то просить платить сразу 10к нет смысла
- **Рефлексия.** Какая проблема решается? Если без продукта людям больно, они готовы платить больше. Если просто «немного удобнее» — ждать миллиона смысла нет.

Как выбрать нужную операционную метрику?

Надо выбрать несколько!



	Обычный человек	DS
ML-метрика		Показатель качества генерализации данных моделью
Операционная метрика (низ дерева)	Какой-то показатель качества	Таргет
Остальные бизнес-метрики		Какой-то показатель качества

Выбор прокси-таргета для ML-решения

Формализуем понятия.

ML-метрика $G: x \rightarrow z$ – это прокси к операционной $F: x \rightarrow y$, $T: z \rightarrow y$ – аппроксимация (конверсия) событий z в бизнес-пользу y (например, конверсия добавления в корзину в факт заказа покупки). Искомая F – композиция G и T .

Выявление целевой бизнес-метрики y происходит сильно дольше, чем прокси метрики z . Поиск прокси-метрик сводится к следующему алгоритму:

1. Собираем:

- Набор больших датасетов $\{D\}$: $\{x\} \sim \{z\}$,
- Датасет $d: \{z\} \sim y$
- $\{x\}$ – набор фичей пользователя и продукта
- $\{z\}$ – набор возможных прокси показателей (быстрые операционные метрики, фичи сессии) – ml target
- m – метрика качества предсказания значения z
- y – оптимизируемая бизнес-метрика

2. Обучаем модели $\{G\}$ для предсказания $\{z\}$, оптимизируя метрики $\{M\}$.

3. Отбираем среди моделей $\{G\}$ те $\{G^*\}$, которые имеют наибольшее качество (и прокси-показатели, соответственно, $\{z^*\}$).

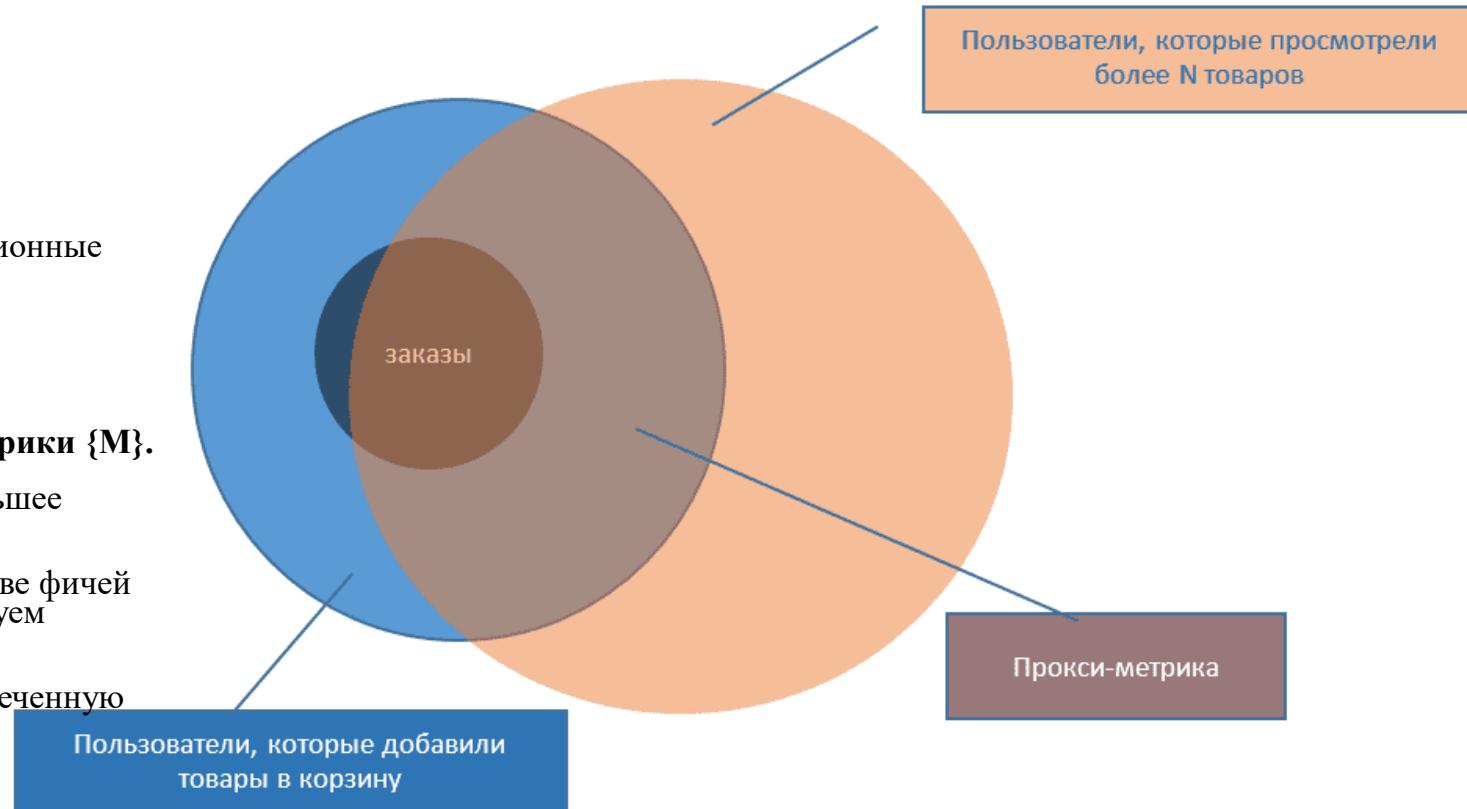
4. Обучаем интерпретируемую модель T на d , используя в качестве фичей $\{z^*\}$ (лин рег, деревья, бустинг на пеньках, сплайны?). Используем регуляризацию и др. методы от переобучения.

5. Объединяем $\{D\}$ в один датасет, скорим при помощи T неразмеченную часть, обучаем $F: x \rightarrow y$.

6. Пилотируем модель F , собираем новый датасет $\{d\}$.

Бонусом получаем Модель роста T , если выбрали показательные Z

Вдохновлено: <https://www.chernobrovov.ru/articles/kak-uskorit-i-udeshevit-reklamnye-eksperimenty-predskazav-povedenie-polzovatelej-internet-magazina-s-pomoshhyu-proksi-metrik.html>



Примеры:

- Позволяет избавиться от кликбейтового контента:
 - <https://www.youtube.com/watch?v=X1a5jNgqzGk>
 - <https://www.youtube.com/watch?v=4JhAiJ2W1i4>
 - <https://habr.com/ru/companies/yandex/articles/651751/> - нелинейная комб.
- Избавиться от биаса: клик = показ(релевантность*(1-каннибализация)+(1-релевантность)*(доверие к рекомендашке +(1-доверие)*популярность*склонность к клику на популярность)).
 - <https://www.youtube.com/watch?v=dQu86dAsLHc>
- Ridge Regression для взвешивания таргетов для RL
 - <https://dl.acm.org/doi/10.1145/3460231.3474247>

А если нет данных d для {z}->y?

1. Собираем use cases и применяем метод Fogg Behavior Model (FBM). Модель, объясняющая поведение через три фактора: мотивация, способность и триггер.
2. Собираем use cases и применяем Фреймворк AAARRR (Acquisition, Activation, Retention, Revenue, Referral) или HEART (Happiness, Engagement, Adoption, Retention, Task success) для анализа воронки или пользовательского опыта как части дерева метрик
3. Поиск доп. связей в дереве через SWOT-анализ (разделение факторов, влияющих на User на четыре категории: Strengths, Weaknesses, Opportunities, Threats).
4. Стесс-тест дерева метрик через S.C.A.M.P.E.R: задаем 7 вопросов к i-той итерации дерева для перехода к i+1: Substitute, Combine, Adapt, Modify, Put to another use, Eliminate, Reverse.
5. Упрочнение (утяжеление) связей в дереве метрик: Triangulation Testing - использование нескольких разных методов (опросы, A/B тесты, интервью) для проверки одной и той же гипотезы.
6. Growth Hacking Experiments – небольшие a/b тесты для сбора конкретных семплов для d.
7. Fake Door Testing – добавление UI без backend'а для проверки заинтересованности user в фиче. Можно добавить нецелевой вариант backend для более точного анализа заинтересованности.

Связь ml и операционной стратегии

Мало выбрать метрику, нужно еще выбрать «критерии приемки качества»

Стратегия	Цель	Бизнес-методы	ML(?)-задачи	Метрики
Стратегия низких издержек	Минимизировать затраты и предложить конкурентоспособные цены на рынке.	Оптимизация цепочек поставок Массовое производство Экономия на масштабе Автоматизация процессов Walmart, IKEA	<ul style="list-style-type: none">Решение задач оптимизации направленных взвешенных графовМодели спроса, ценообразованияУмный парсинг конкурентов (KVI-товары)	Могут быть низкие
Стратегия гибкости	Быстро адаптироваться к изменениям на рынке и в потребностях клиентов.	Гибкость в производстве, быстрые изменения в продуктовой линейке, адаптация к новым трендам. Zara (быстрая мода), Amazon.	<ul style="list-style-type: none">Модели ценообразованияУмный парсинг конкурентов (KVI-товары)Анализ трендов через новости (LLM)	Не ниже среднего

Связь ML и операционной стратегии

Стратегия	Цель	Бизнес-методы	ML(?)-задачи	Метрики
Стратегия фокусирования	Сфокусироваться на узком сегменте рынка и удовлетворить его потребности лучше, чем конкуренты.	Глубокое понимание рынка, нишевые продукты, высокое качество обслуживания. Ferrari, Rolex	<ul style="list-style-type: none">• А нужен ли вообще?• Автопилот• Навигатор по борделям	Определяются законодательством (если в 50% сбываем людей – долго бизнес не протянет)
Стратегия устойчивости	Ведение бизнеса с учетом экологической и социальной ответственности.	Экологически чистое производство, социально ответственные практики, устойчивое управление ресурсами Patagonia, Unilever.	<ul style="list-style-type: none">• ML в производстве для минимизации брака / выбросов	Не важны? Минимум углеродного следа от GPU Fairness

Связь ml и операционной стратегии

Стратегия	Цель	Бизнес-методы	ML(?)-задачи	Метрики
Стратегия дифференциации	Предложить уникальный продукт или услугу, которая выделяется на рынке.	Инновации в продуктах, высокий уровень сервиса, сильный бренд, уникальные характеристики. Apple, Nike	<ul style="list-style-type: none">• Ai-based core бизнес• Ai-based фичи• CRM с доп источниками	Должны быть высокие как продуктовые, так и на бенчмарках Косвенное влияние на бизнес
Стратегия инноваций	Быть лидером в разработке новых продуктов и технологий.	Инвестирование в исследования и разработки, стимулирование креативности, быстрый вывод на рынок новых продуктов. Google, Tesla	<ul style="list-style-type: none">• Ai-based core бизнес• Ai-based фичи	Должны быть высокие как продуктовые, так и на бенчмарках Прямое влияние на бизнес

Связь ML и операционной стратегии

Стратегия	Цель	Бизнес-методы	ML(?)-задачи	Метрики
Стратегия обслуживания	Предоставление исключительного уровня сервиса клиентам	Обучение персонала, ориентированность на клиента, индивидуальный подход. Ritz-Carlton, Zappos.	<ul style="list-style-type: none">CRM с доп источникамиLLM-based тренажеры общения с клиентом	Могут быть низкие
Стратегия качества	Обеспечить высокое качество продукции и услуг	Постоянное улучшение процессов (Continuous Improvement), контроль качества, вовлечение сотрудников в процессы улучшения. Toyota (Total Quality Management), Siemens	<ul style="list-style-type: none">CV-детекция дефектовCV-детекция охрана трудаДетекция миссейлингаАнтифрод	Должны быть высокие

Полезные ссылки

- ProductStar [https://www.youtube.com/channel/UCN97-wQgHScYUhnz31_NdeA]
- ProductSense [<https://www.youtube.com/c/ProductSense>]
- Gopractice – онлайн симулятор - упор на продуктовую аналитику, если в бюджете вашей компании есть бюджет – очень рекомендую [https://simulator.gopractice.ru/?utm_source=blog&utm_campaign=mainmenu]
- Читаем книгу «Спроси маму» [https://leader-id.storage.yandexcloud.net/event_doc/297226/627d2f800defe600717902.pdf]
- Статья на Go Practice о проверке гипотез ценности без разработки: [https://gopractice.ru/product/testing_hypotheses_without_development/]
- Основы Customer Development. Иван Замесин [<https://www.youtube.com/watch?v=ymsswuloXz8>]
- Статья о Дизайн Мышлении [<https://tilda.education/courses/web-design/designthinking/>]
- Тест о навыке управления продуктом [<https://gopractice.ru/sat/>]

Семинарская часть

- <https://www.evidentlyai.com/ml-system-design>
- A New Era of Creativity: Expert-in-the-loop Generative AI at Stitch Fix (генерация headline товаров в режиме online)
- <https://multithreaded.stitchfix.com/blog/2023/03/06/expert-in-the-loop-generative-ai-at-stitch-fix/>

Семинарская часть

Гиперпараметр	Значение
Целевая метрика	(безопасность + продающая способность + полнота + нет галлюцинаций)/4 Разметка асессорами
Loss-функция	API GigaChat => нет Loss. Или Lora дообучение со стандартным loss, но метки асессоров берем по компонентам целевой метрики Для RAG используем стандартный loss
Min значение допустимое	70% (требование моего бизнес-owner)
Max значение достижение	100%
Принцип разбиения данных	id item, ~равномерно по тематикам
Дополнительная (-ые) метрика качества	Компоненты целевой метрики Скорость ml-решения Стоимость инференса ml-решения при единичном вызове Качество работы RAG (для поиска контекста для описания): MP@3≥80% (требование моего бизнес-owner)

вопросы



КОНЕЦ ЛЕКЦИИ 09.02