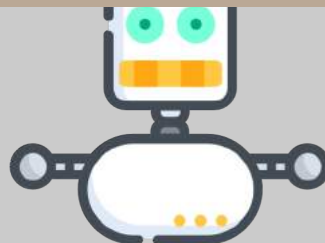


Постановка задач



Что входит в постановку задачи



Какое
целевое
значение



Как
используются
прогнозы



Как
оценивается
качество



Доп.
требования

Пример постановки (для churn prediction)

- Отток - разрыв договора подключения к сервису
- Модель - бинарная классификации
- Горизонт прогнозирования – 2 недели
- Методика оценки – метрика AUC
- Применение – удержание топ-k% по вероятности
- Дизайн эксперимента – А/Б тестирование на 10% сегменте случайных пользователей
- Требования к модели: вероятностная модель

Пример постановки (для churn prediction)

- Отток - разрыв договора подключения к сервису
- Модель - бинарная классификации
- Горизонт прогнозирования – 2 недели
- Методика оценки – метрика AUC
- Применение – удержание топ-k% по вероятности
- Дизайн эксперимента – А/Б тестирование на 10% сегменте случайных пользователей
- Требования к модели: вероятностная модель

Примечание: в этом примере не все хорошо, далее мы это обсудим

План

Таргет и прогнозы

Качество на исторических данных

Экономический эффект

Оценка после внедрения

Таргет и прогнозы

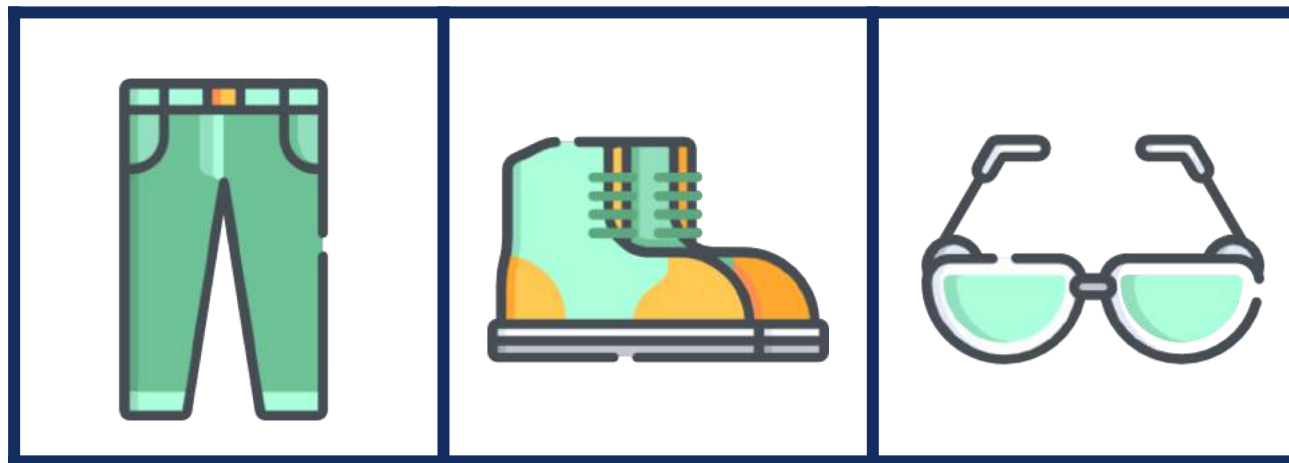
Выбор таргета

1. Понять, что вы хотите прогнозировать
2. Понять, как вы хотите использовать прогноз и почему именно так
3. Проверить, нет ли «побочных» эффектов у текущего выбора таргета (нужно строить много моделей или могут появляться странные прогнозы)
4. Проверить, что нет неоднозначности в определении таргета

Кейс: рекомендации товаров

Товар 1	Товар 2	Товар 3
---------	---------	---------

Максимизация количества покупок



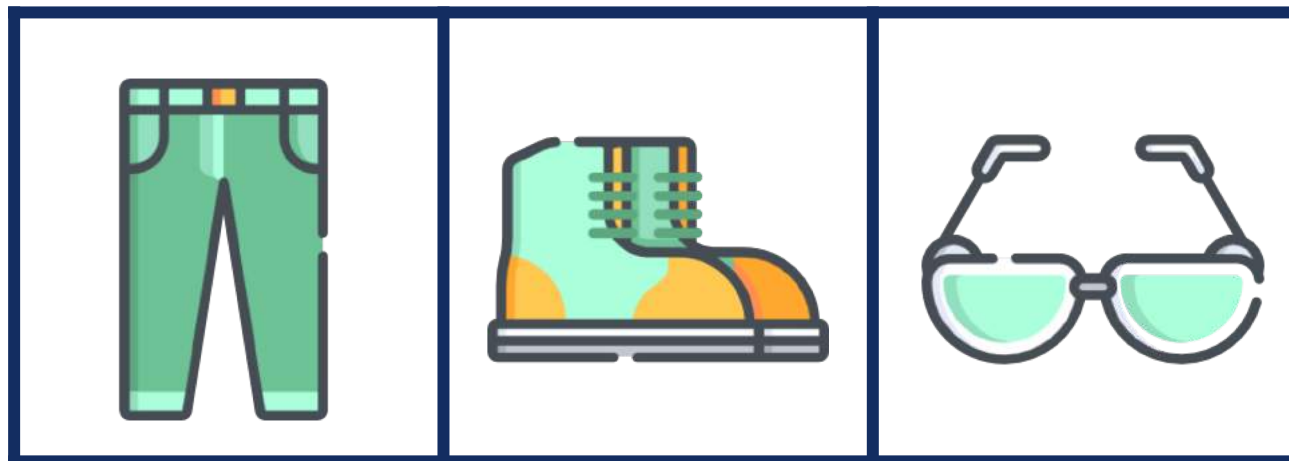
Вероятность

0.05

0.02

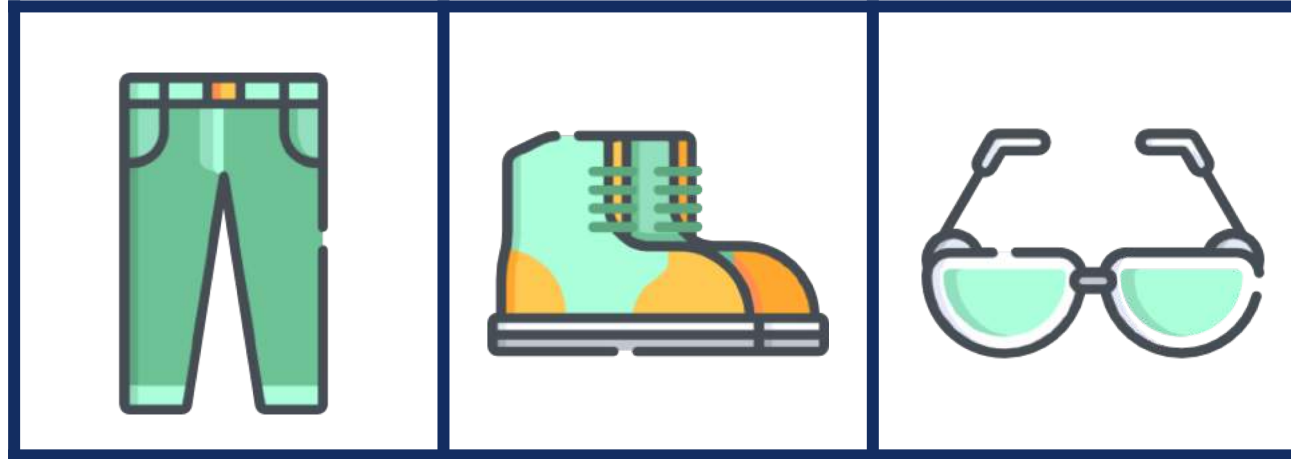
0.015

Максимизация дохода



Вероятность	0.05	0.02	0.015
Цена	7000	3000	900

Максимизация прибыли



Вероятность	0.05	0.02	0.015
Цена	7000	3000	900
Маржинальность	0.1	0.2	0.4

Прогнозирование вероятности

- Объекты: тройки (пользователь, товар, момент времени)

Прогнозирование вероятности

- Объекты: тройки (пользователь, товар, момент времени)
- Классы: 1 - товар будет куплен, 0 – товар не будет куплен

Прогнозирование вероятности

- Объекты: тройки (пользователь, товар, момент времени)
- Классы: 1 - товар будет куплен, 0 – товар не будет куплен
- Признаки: параметры пользователя, товара, момента времени и их «взаимодействие»

Кейс: подсказки в мобильном приложении



Предугадываем, что пользователь введет в приложении.

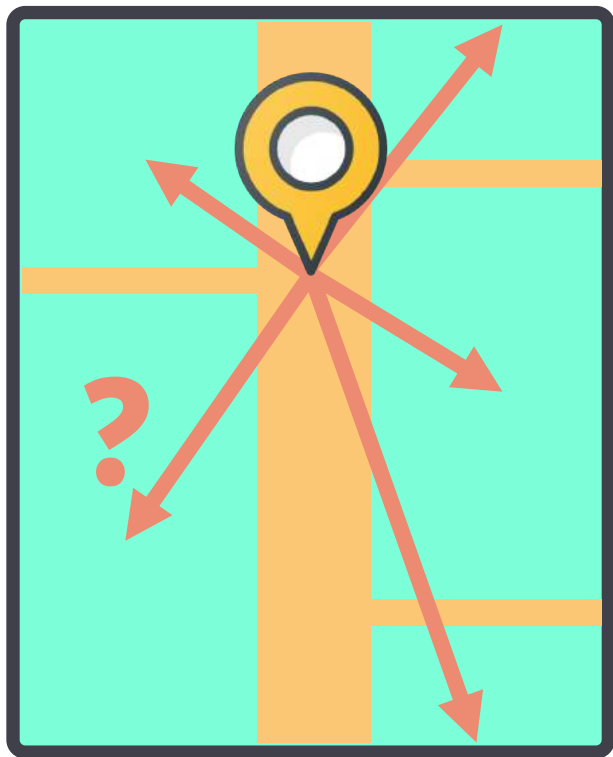
Пример: подсказки места назначения в приложении такси

Что делаем

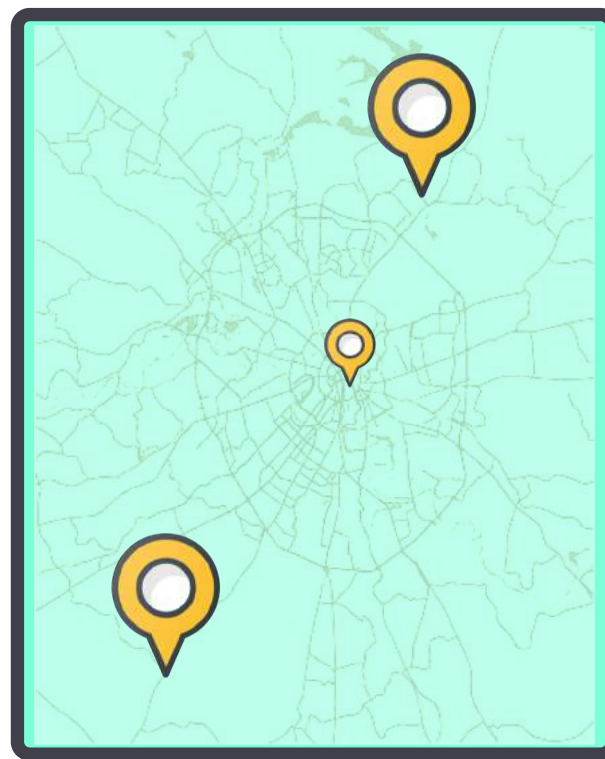
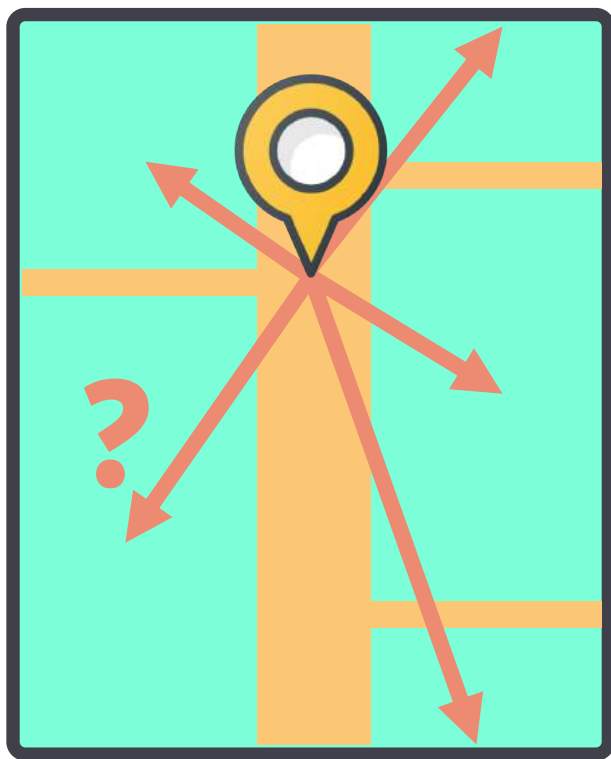
Какая постановка задачи должна быть?

1. Задача регрессии: объекты – пользователи, целевое значение – координаты точки В

Чем плохо угадывать координаты



Чем плохо угадывать координаты

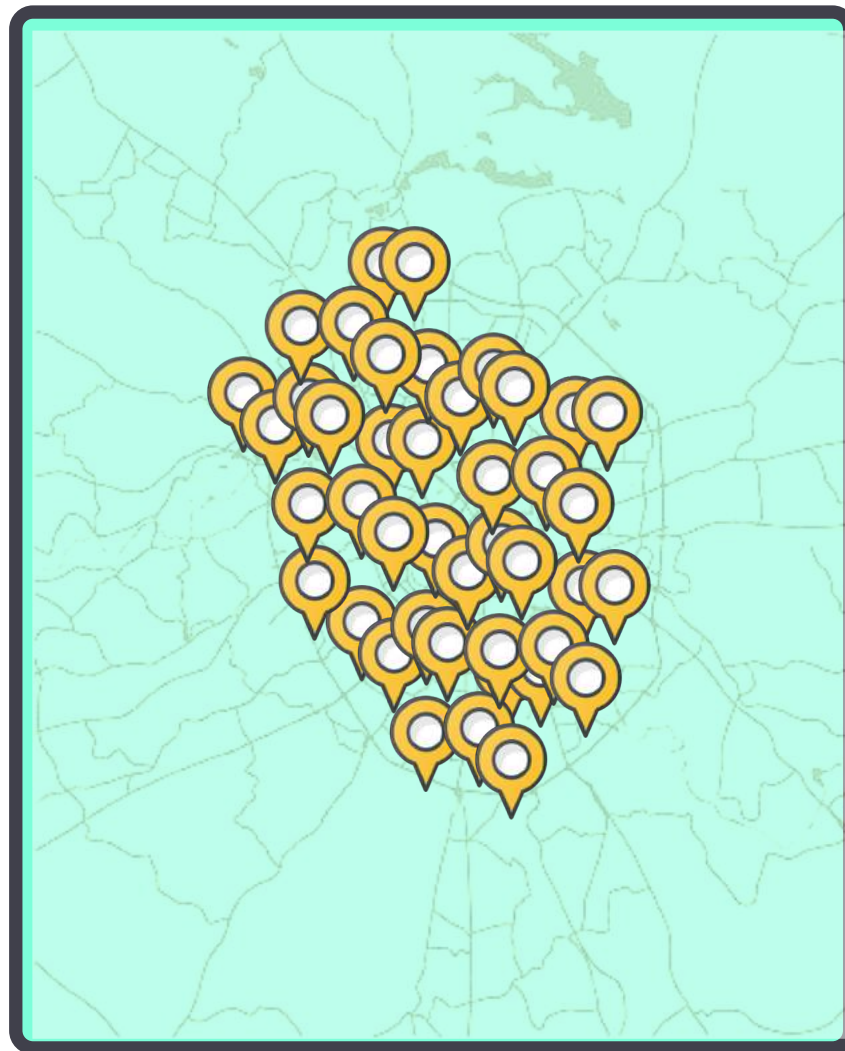


Что делаем

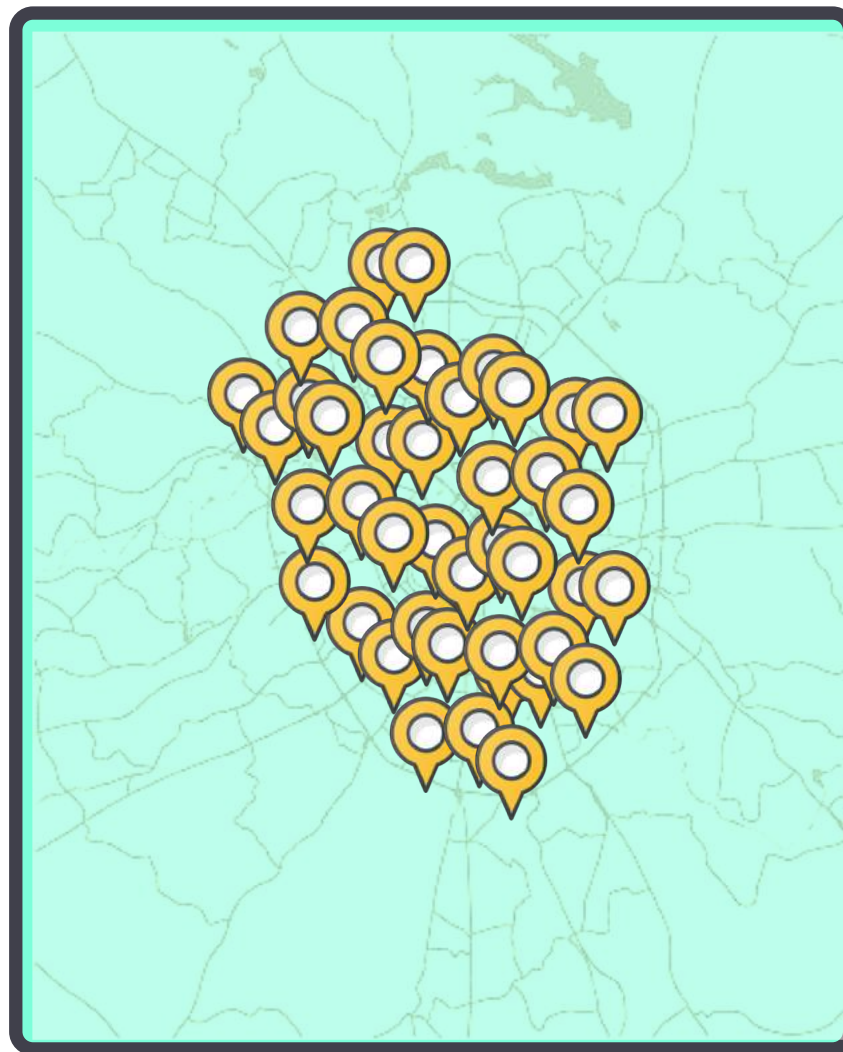
Какая постановка задачи должна быть?

1. Задача регрессии: объекты – пользователи, целевое значение – координаты точки В
2. Задача классификации: объекты - пары (пользователи, момент времени), классы – адреса мест, куда может поехать

Если классы –
все места в
городе

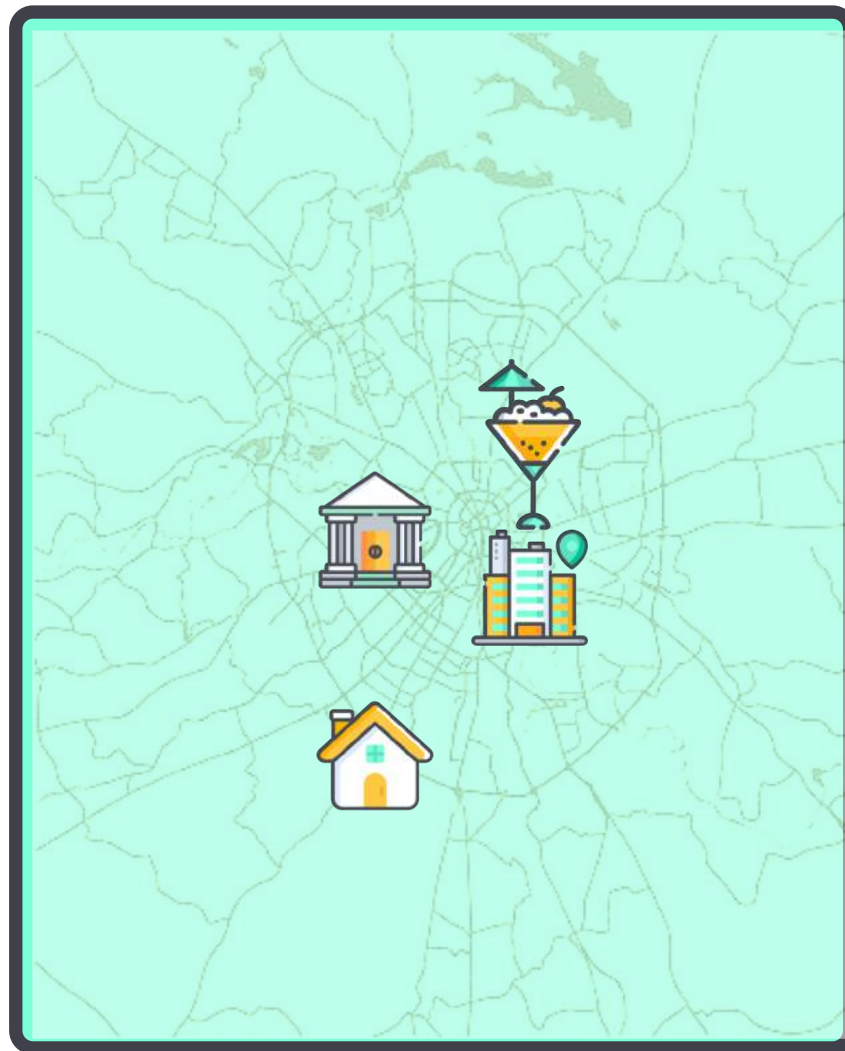


Если классы –
все места в
городе

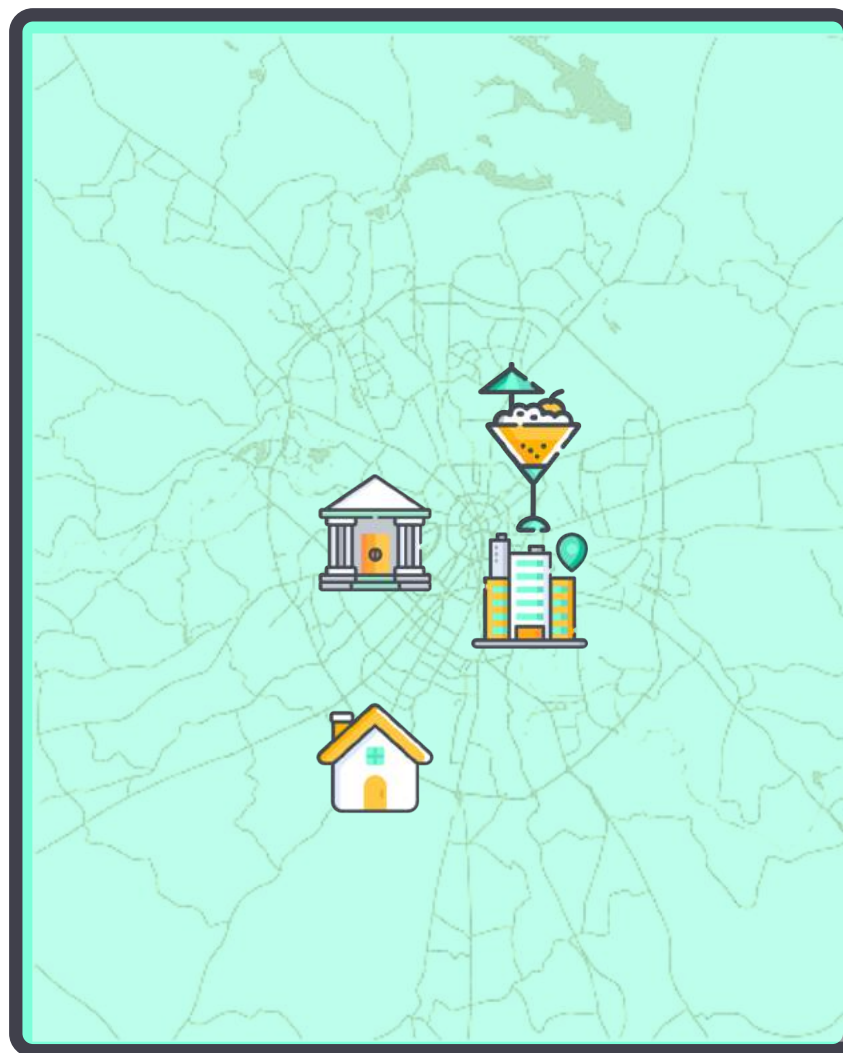


Слишком много классов!

Если классы –
частые места
пользователя



Если классы –
частые места
пользователя



Обучающая выборка – история поездок **одного** пользователя!
Данных всегда будет мало, а моделей понадобится много.

Что делаем

Какая постановка задачи должна быть?

1. Задача регрессии: объекты – пользователи, целевое значение – координаты точки В
2. Задача классификации: объекты - пары (пользователи, момент времени), классы – адреса мест, куда может поехать
3. Задача классификации: объекты - тройки (пользователь, место, момент времени, классы – 1 и 0 (поехал или нет), 1 при точном попадании в адрес или в окрестность 20 метров

Кейс: прогнозирование оттока

- Отток - разрыв договора подключения к сервису
- Модель - бинарная классификации
- Горизонт прогнозирования – 2 недели
- Методика оценки – метрика AUC
- Применение – удержание топ-k% по вероятности
- Дизайн эксперимента – А/Б тестирование на 10% сегменте случайных пользователей
- Требования к модели: вероятностная модель

Возможные определения оттока

- Разрыв договора
- Отсутствие платных транзакций более 10/90 дней
- Отсутствие на сервисе более 14/28 дней

Доля уходящих в отток

- Разрыв договора – 0,1%
- Отсутствие платных транзакций более 10 дней – 22%
- Отсутствие платных транзакций более 90 дней – 5%
- Отсутствие в сервисе более 14 дней – 16%
- Отсутствие в сервисе более 28 дней – 9%

Доля возвращающихся

- Разрыв договора – 0%
- Отсутствие платных транзакций более 10 дней – 80%
- Отсутствие платных транзакций более 90 дней – 12%
- Отсутствие в сервисе более 14 дней – 90%
- Отсутствие в сервисе более 28 дней – 78%

Горизонт прогнозирования

- Как быстро мы можем связаться с пользователем?
- Какие методики удержания мы используем?
- Сколько времени занимает процесс удержания?

Выбор таргета

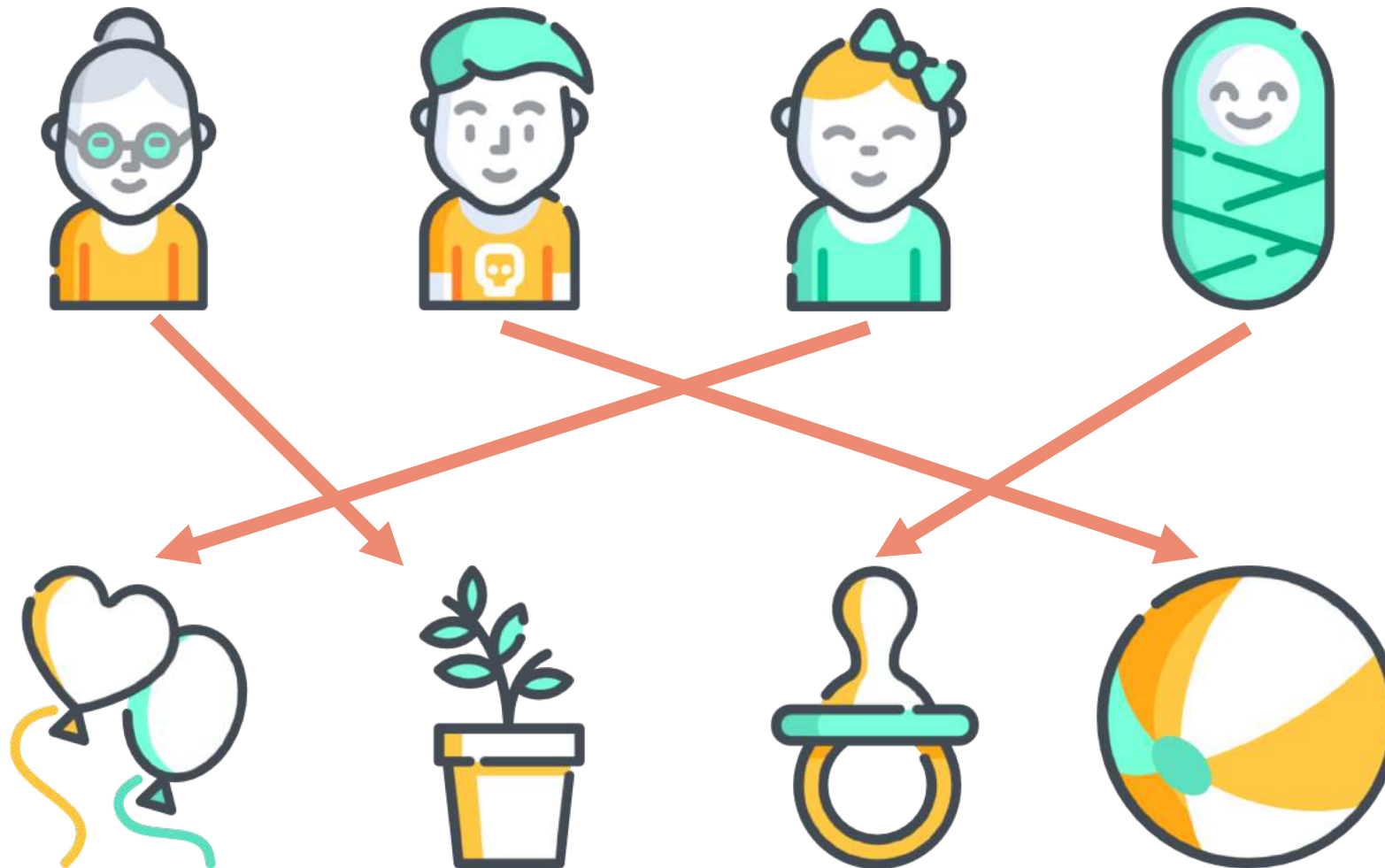
1. Понять, что вы хотите прогнозировать
2. Понять, как вы хотите использовать прогноз и почему именно так
3. Проверить, нет ли «побочных» эффектов у текущего выбора таргета (нужно строить много моделей или могут появляться странные прогнозы)
4. Проверить, что нет неоднозначности в определении таргета

Качество на исторических данных

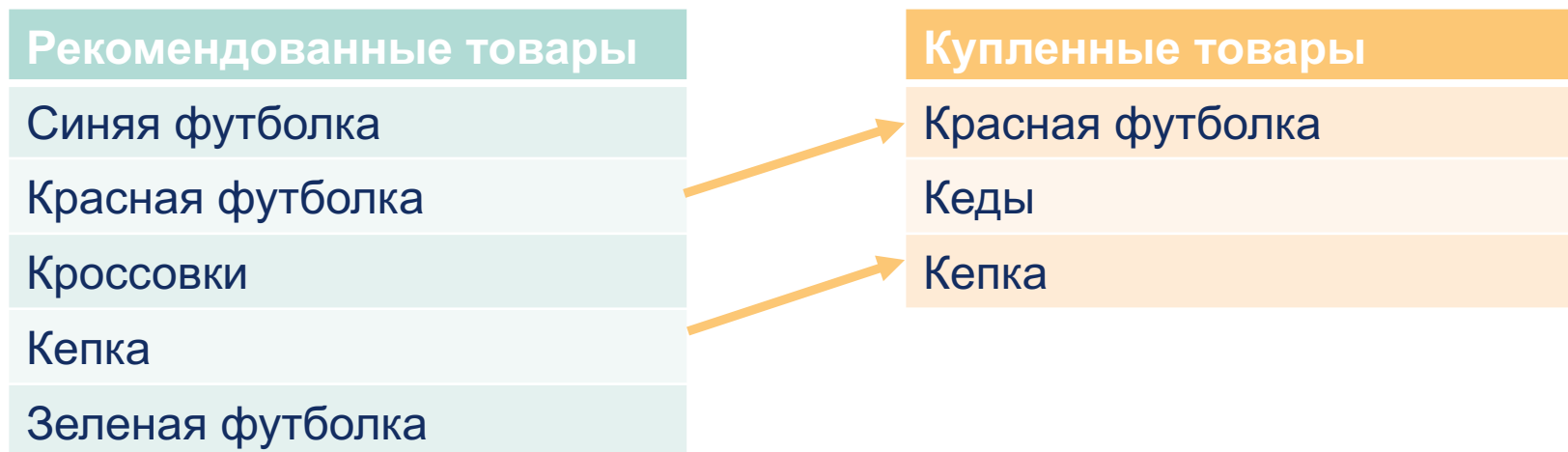
Выбор метрики качества

1. Исходя из того, как вы используете прогноз, что значит, что прогноз должен быть «хорошим»?
2. Если вам нужно сравнивать качество в разных экспериментах, позаботьтесь о том, чтобы оно было сравнимо
3. Проверьте, можно ли «сломать» метрику качества, зависив ее за счет очевидного «ухудшения» ответа
4. Постарайтесь связать качество с экономическим эффектом

Кейс: рекомендации товаров



Точность (Precision@k)

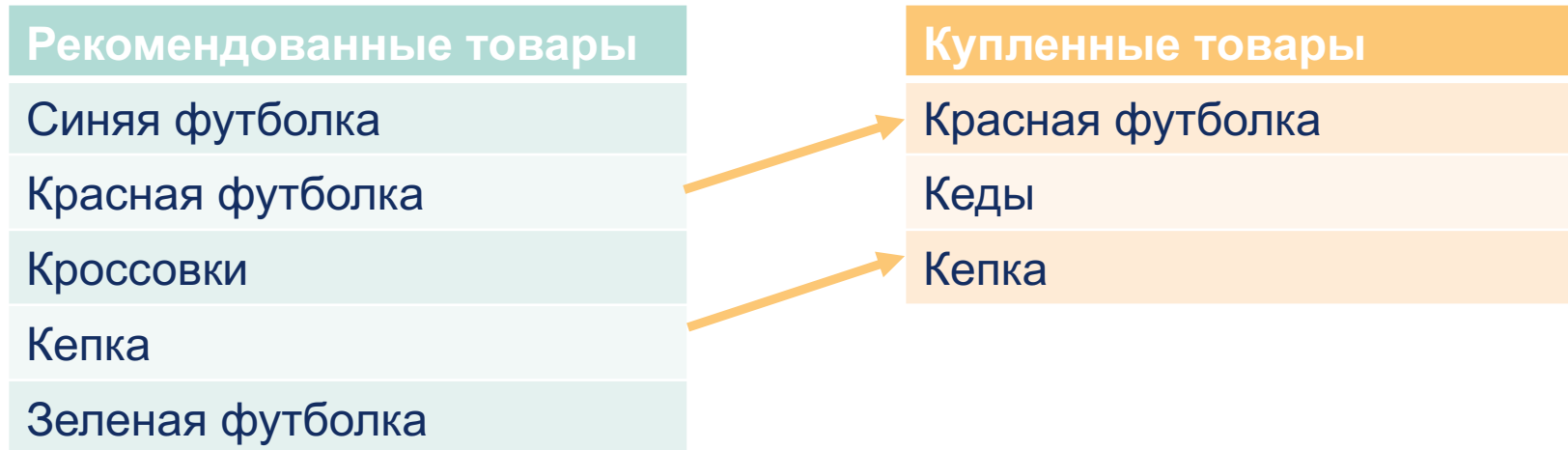


k – количество
рекомендаций

$$\text{Precision@}k = \frac{\text{купленное из рекомендованного}}{k}$$

AveragePrecision@k - усредненный по сессиям Precision@k

Полнота (Recall@k)



k – количество
рекомендаций

$$\text{Recall@k} = \frac{\text{купленное из рекомендованного}}{\text{количество покупок}}$$

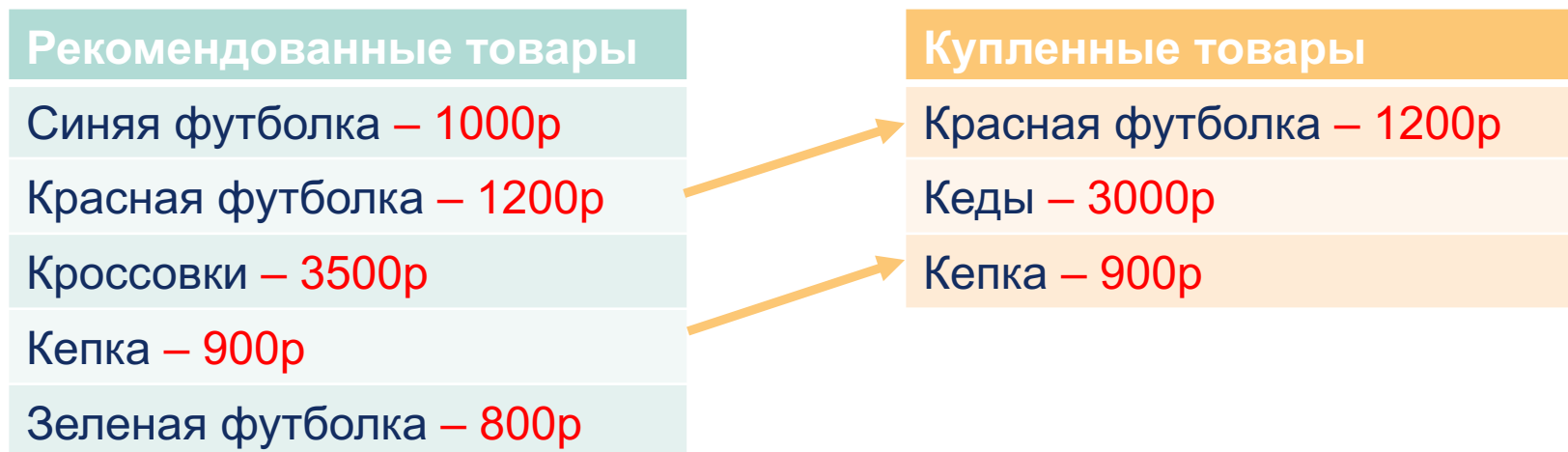
AverageRecall@k - усредненный по сессиям Recall@k

Качество классификации против качества рекомендаций

Пример – 2 решения для прогноза купит/не купит товар

	Алгоритм 1	Алгоритм 2
Точность классификации	0.52	0.85
Recall@5	0.72	0.71

Взвешенный ценами recall@k



$$\text{Взвешенный ценами Recall@k} = \frac{\text{стоимость купленного из рекомендованного}}{\text{стоимость покупок}}$$

AverageRecall@k - усредненный по сессиям Recall@k

Кейс: подсказки в приложении

Какая метрика более уместна?

1. Accuracy классификатора (доля угадываний поедет/не поедет)

Как оцениваем качество

Какая метрика более уместна?

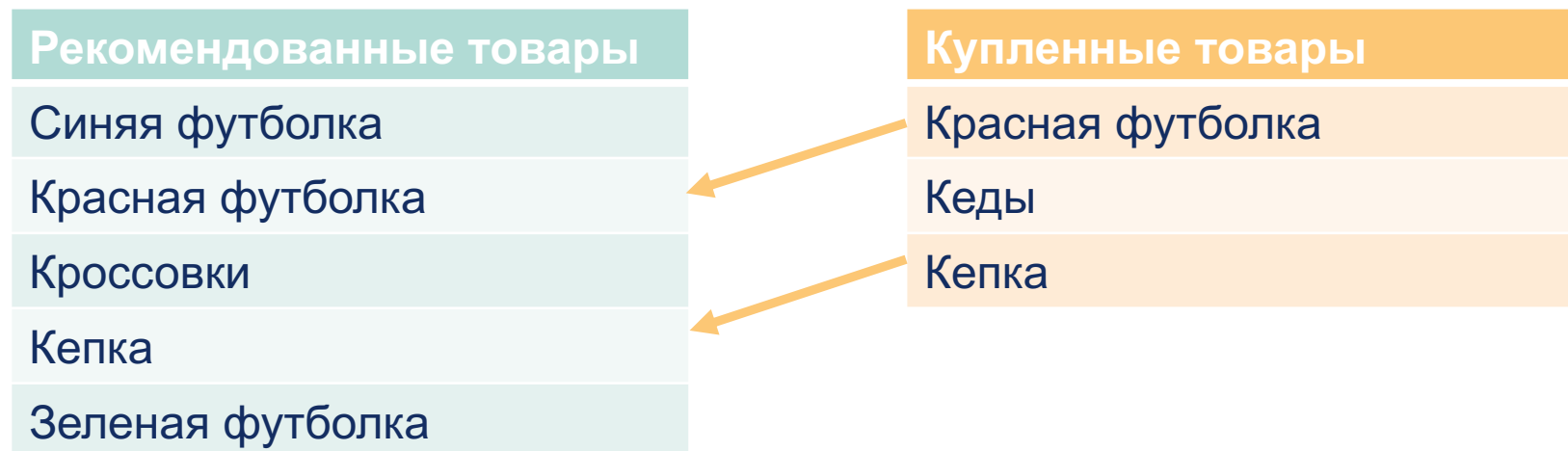
1. Accuracy классификатора (доля угадываний поедет/не поедет)
2. ROC-AUC

Как было в рекомендациях

Пример – 2 решения для прогноза купит/не купит товар

	Алгоритм 1	Алгоритм 2
ROC-AUC	0.52	0.85
Recall@5	0.72	0.71

Полнота (Recall@k)



k — количество
рекомендаций

$$\text{Recall@k} = \frac{\text{купленное из рекомендованного}}{\text{количество покупок}}$$

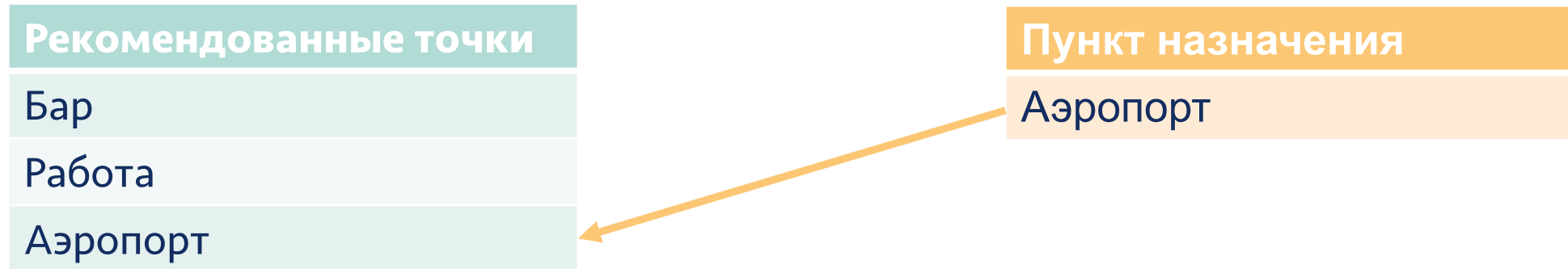
AverageRecall@k - усредненный по сессиям Recall@k

Как оцениваем качество

Какая метрика более уместна?

1. Accuracy классификатора (доля угадываний поедет/не поедет)
2. ROC-AUC
3. Recall@1, Recall@2

Полнота в нашем примере



k – количество
рекомендаций

$$\text{Recall@k} = \frac{\text{угаданный пункт назначения}}{1}$$

AverageRecall@k – как часто показываем правильную подсказку

Кейс: прогнозирование оттока

1. Удерживаем топ- $k\%$ по вероятности оттока
2. Логично смотреть на $\text{precision@}k\%$ или $\text{recall@}k\%$
3. Но не все так просто

Проблема

Доля отточников сама по себе очень сильно плавает на разных частях пользовательской базы и в разные моменты времени.

Проблема

Доля отточников сама по себе очень сильно плавает на разных частях пользовательской базы и в разные моменты времени.

Как следствие – precision@k очень нестабилен. Нужна метрика, не сильно зависящая от реальной доли отточников

Проблема

Доля отточников сама по себе очень сильно плавает на разных частях пользовательской базы и в разные моменты времени.

Как следствие – precision@k очень нестабилен. Нужна метрика, не сильно зависящая от реальной доли отточников

Вопрос 1: почему это – проблема?

Проблема

Доля отточников сама по себе очень сильно плавает на разных частях пользовательской базы и в разные моменты времени.

Как следствие – precision@k очень нестабилен. Нужна метрика, не сильно зависящая от реальной доли отточников

Вопрос 1: почему это – проблема?

Вопрос 2: как «подправить» метрику?

Lift-10%

$$\text{Lift-10\%} = \text{precision@10\%} / \text{precision@100\%}$$

Т.е. нормируем на долю отточников в выборке

Кейс: оценка качества продукта

Задача: классифицировать производимый на производстве продукт на 3 класса - пригодный к использованию, непригодный, нельзя понять по имеющимся замерам



Кейс: оценка качества продукта

Задача: классифицировать производимый на производстве продукт на 3 класса - пригодный к использованию, непригодный, нельзя понять по имеющимся замерам

Выход алгоритма: три вероятности (p_1 , p_2 , p_3)

Ответы в исторических данных: (1,0,0), (0, 1, 0), (0, 0, 1)

Выбор метрики качества:

Средний модуль по координатной разности



Кейс: оценка качества продукта

Задача: классифицировать производимый на производстве продукт на 3 класса - пригодный к использованию, непригодный, нельзя понять по имеющимся замерам

Выход алгоритма: три вероятности (p_1 , p_2 , p_3)

Ответы в исторических данных: (1,0,0), (0, 1, 0), (0, 0, 1)

Выбор метрики качества:

Средний модуль по координатной разности

Вопрос: в чем подвох?



Выбор метрики качества

1. Исходя из того, как вы используете прогноз, что значит, что прогноз должен быть «хорошим»?
2. Если вам нужно сравнивать качество в разных экспериментах, позаботьтесь о том, чтобы оно было сравнимо
3. Проверьте, можно ли «сломать» метрику качества, зависив ее за счет очевидного «ухудшения» ответа
4. Постарайтесь связать качество с экономическим эффектом

Экономический эффект

Оценка экономического эффекта

1. Как использование прогнозов модели позволяет сэкономить/заработать деньги? (или оптимизировать другой значимый показатель)
2. Какие затраты возникают?
3. Какие параметры в экономической модели не зависят от вашего алгоритма, а какие – зависят? Чем можно пренебречь?
4. Как можно уточнять экономическую модель? Как это скажется на постановке задачи? Будет ли это действительно уточнением, а не огрублением?

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег
- p – доля настоящих отточников среди удерживаемых

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег
- p – доля настоящих отточников среди удерживаемых

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег
- p – доля настоящих отточников среди удерживаемых
- Np – удержанных пользователей, если удерживаем со 100% успехом

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег
- p – доля настоящих отточников среди удерживаемых
- Np – удержанных пользователей, если удерживаем со 100% успехом
- Npr – удержанных пользователей, если удержание успешно с вероятностью r

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег
- p – доля настоящих отточников среди удерживаемых
- Np – удержанных пользователей, если удерживаем со 100% успехом
- Npr – удержанных пользователей, если удержание успешно с вероятностью r
- ARPU (average revenue per user) тоже известна

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег
- p – доля настоящих отточников среди удерживаемых
- Npr – удержанных пользователей, если удержание успешно с вероятностью r

Экономический эффект: ?

Кейс: прогнозирование оттока

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег
- p – доля настоящих отточников среди удерживаемых
- Npr – удержанных пользователей, если удержание успешно с вероятностью r

Экономический эффект: $ARPU * N * p * r - C * N$

Вернемся к экономике

- Пусть мы удерживаем N пользователей, наиболее вероятно уходящих в отток по прогнозу нашей модели
- Тратим на удержание каждого C денег
- p – доля настоящих отточников среди удерживаемых
- Npr – удержанных пользователей, если удержание успешно с вероятностью r

Экономический эффект: $ARPU * N * p * r - C * N$

Вывод: нужно максимизировать p Precision@k

Упражнение 1

Попробовать усложнить экономическую модель:
что еще можно добавить, как поменяется
выражение для экономического эффекта?

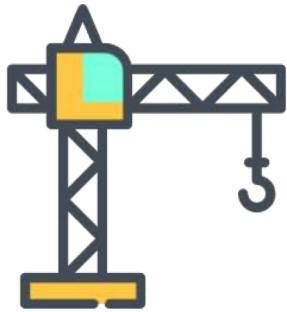
Упражнение 2

Как все поменяется, если мы хотим учесть, что у каждого пользователя свое LTV (lifetime value)?
Что теперь потребуется прогнозировать и как оценивать экономический эффект?

Упражнение 3

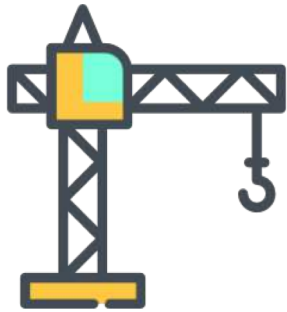
Разрабатывается модель машинного обучения для оценки надежности клиента, желающего арендовать автомобиль. Цель использования модели – избежать потерь на недобросовестных клиентах. Как можно сформулировать задачу исходя из экономических соображений?

Поучительный пример про удержание



- Решалась задача прогноза оттока абонентов
- LIFT-10% получился высоким
- Но оказалось, что модель просто научилась хорошо детектировать трудовых мигрантов

Поучительный пример про удержание



- Решалась задача прогноза оттока абонентов
- LIFT-10% получился высоким
- Но оказалось, что модель просто научилась хорошо детектировать трудовых мигрантов

Что это значит для экономического эффекта?

Кейс: приоритезация обращений в поддержку



В отдел поддержки постоянно поступают запросы от клиентов организации

- 2016 год: 1000 обращений в день, успевают посмотреть все
- 2017 год: 5000 обращений в день, успевают посмотреть 2000 (фильтр по ключевым словам)
- 2018 год: как выжить с 20000 обращений?

Кейс: приоритезация обращений в поддержку



Где здесь деньги:

По простым фильтрам за день выделяется 2000 важных обращений, из них действительно срочных – 800

Кейс: приоритезация обращений в поддержку



Где здесь деньги:

По простым фильтрам за день выделяется 2000 важных обращений, из них действительно срочных – 800

Итого – поддержка работает на 40%

Кейс: приоритезация обращений в поддержку



Где здесь деньги:

По простым фильтрам за день выделяется 2000 важных обращений, из них действительно срочных – 800

Итого – поддержка работает на 40%

Как обеспечить то же качество на большем объеме обращений без расширения штата?
(Или повысить качество)

Упражнение 5



Что делаем: классифицируем обращения на более срочные и менее срочные, на выходе модели – вероятность «срочности»

Вопрос: как оценить экономический эффект? Как он связан с качеством модели?

Упражнение 6



Как сформулировать
экономический эффект от
добавления подсказок в
приложение?

(Можно предложить несколько
вариантов)

Оценка экономического эффекта

1. Как использование прогнозов модели позволяет сэкономить/заработать деньги? (или оптимизировать другой значимый показатель)
2. Какие затраты возникают?
3. Какие параметры в экономической модели не зависят от вашего алгоритма, а какие – зависят? Чем можно пренебречь?
4. Как можно уточнять экономическую модель? Как это скажется на постановке задачи? Будет ли это действительно уточнением, а не огрублением?

Оценка после внедрения

Оценка после внедрения

1. Решить, какие показатели важны после внедрения модели
2. Придумать, как оценить их изменение после внедрения (дизайн эксперимента)
3. Заранее подумать о том, как проверить, что измеренный эффект – не случайный (оценка стат. значимости) **Вопрос: до какой степени это возможно?**
4. Оценить необходимую длительность эксперимента
5. Проверить, что нет «нечестных» способов зависить результаты (может означать неудачный выбор KPI)

Онлайновая оценка качества

Допустим, на исторических данных качество алгоритма высокое, а будет ли оно высоким в реальности?

Онлайновая оценка качества

Допустим, на исторических данных качество алгоритма высокое, а будет ли оно высоким в реальности?

Идеи:

1. A/B тест
2. Оценка статистической значимости результата

A/B тест

1. Случайным образом делим пользователей на равные группы
2. Измеряем целевые метрики (например, конверсию, количество заказов или доход) в каждой группе за длительный период времени
3. Получаем какое-то число для каждой группы
4. Что дальше?

Пример: А/В тесты и статистика



Пример: А/В тесты и статистика

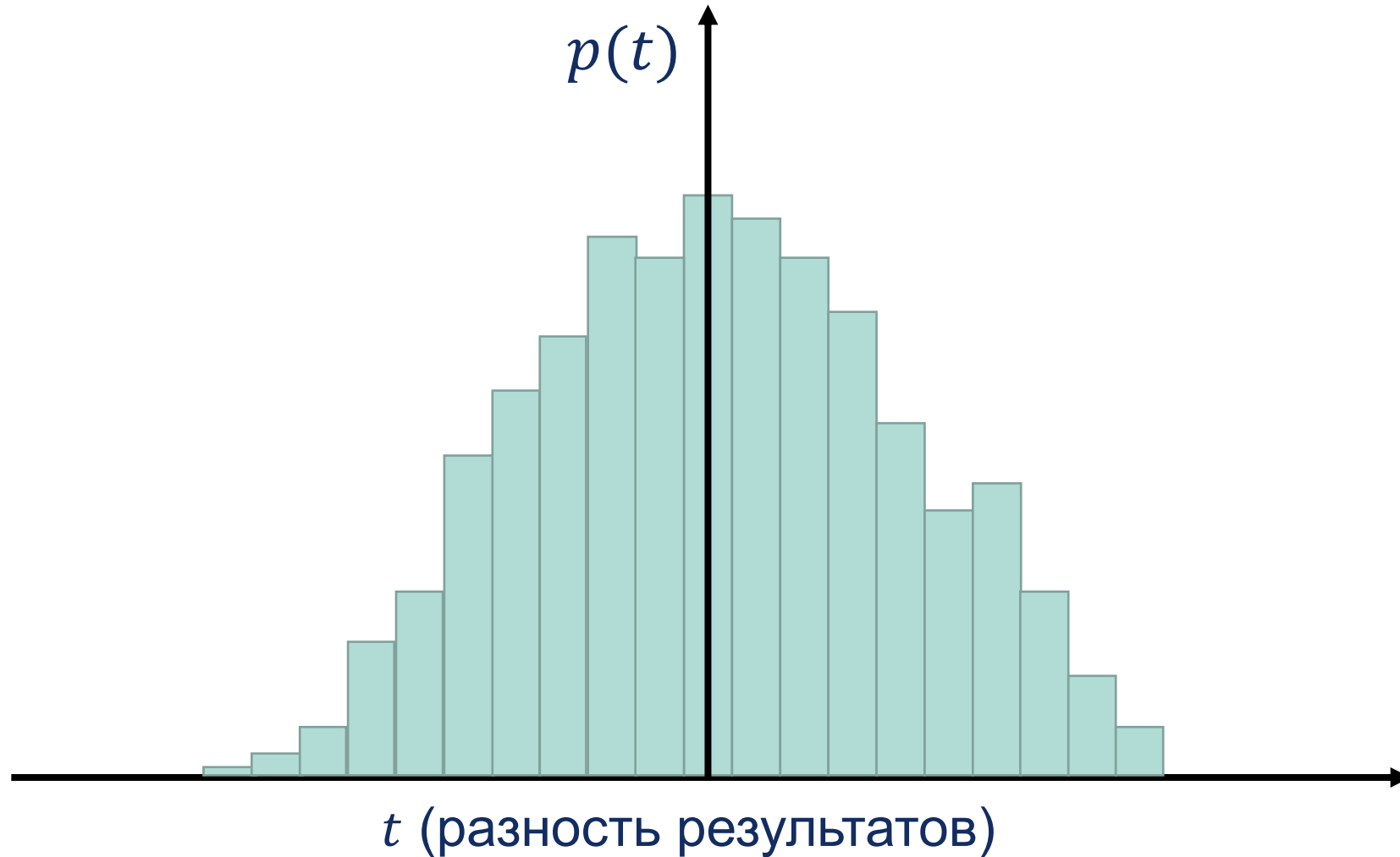


Одна кривая отличается от других на 10%
Но разбиение на самом деле – случайное

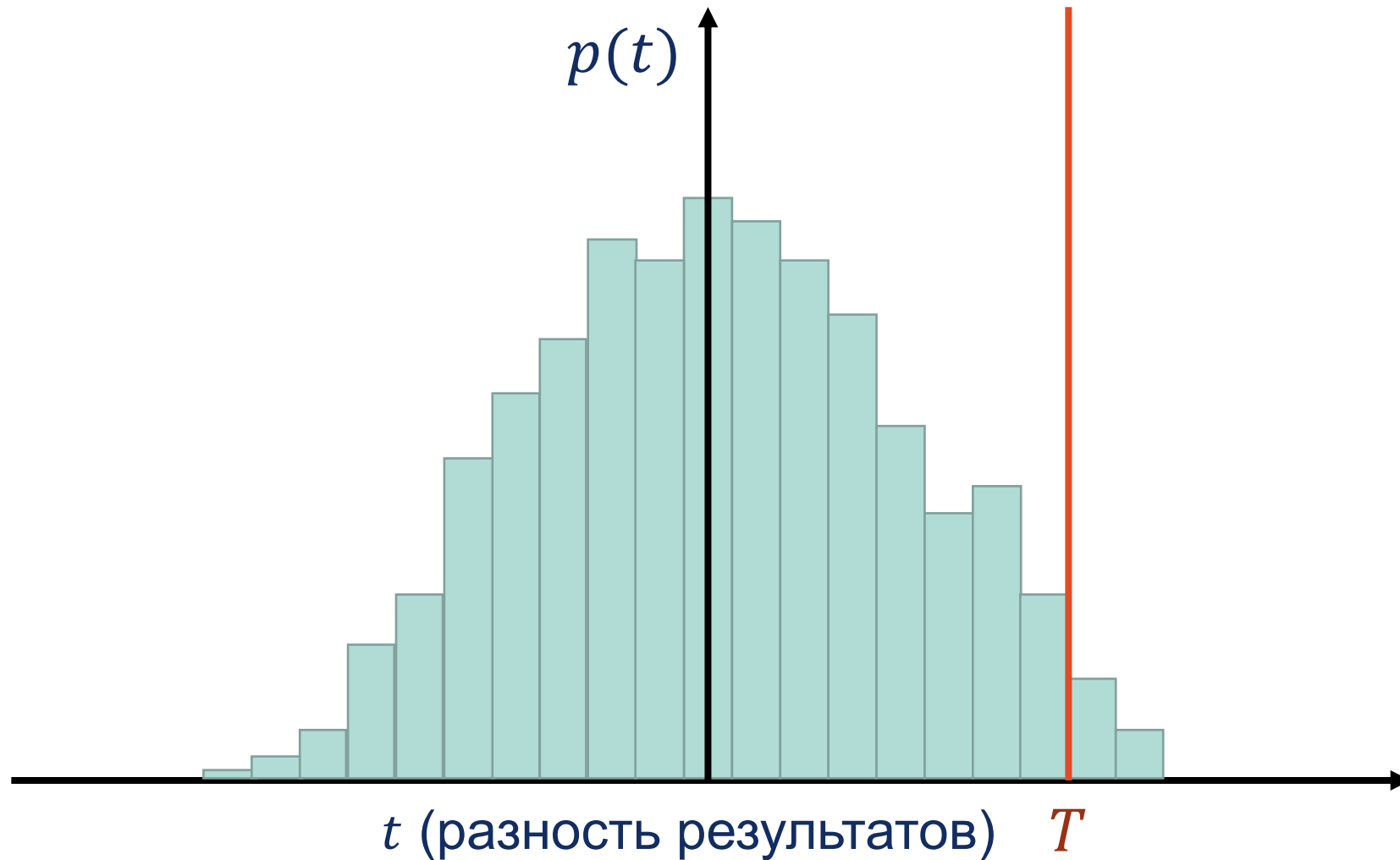
Самый простой тест: бутстреп

1. Имитируем А/А тест на исторических данных, N раз случайно разбив на две группы и посчитав результаты в каждой
2. Строим распределение разности результатов в группах
3. По этому распределению оцениваем вероятность получить в А/А тесте такую же разность как в А/В

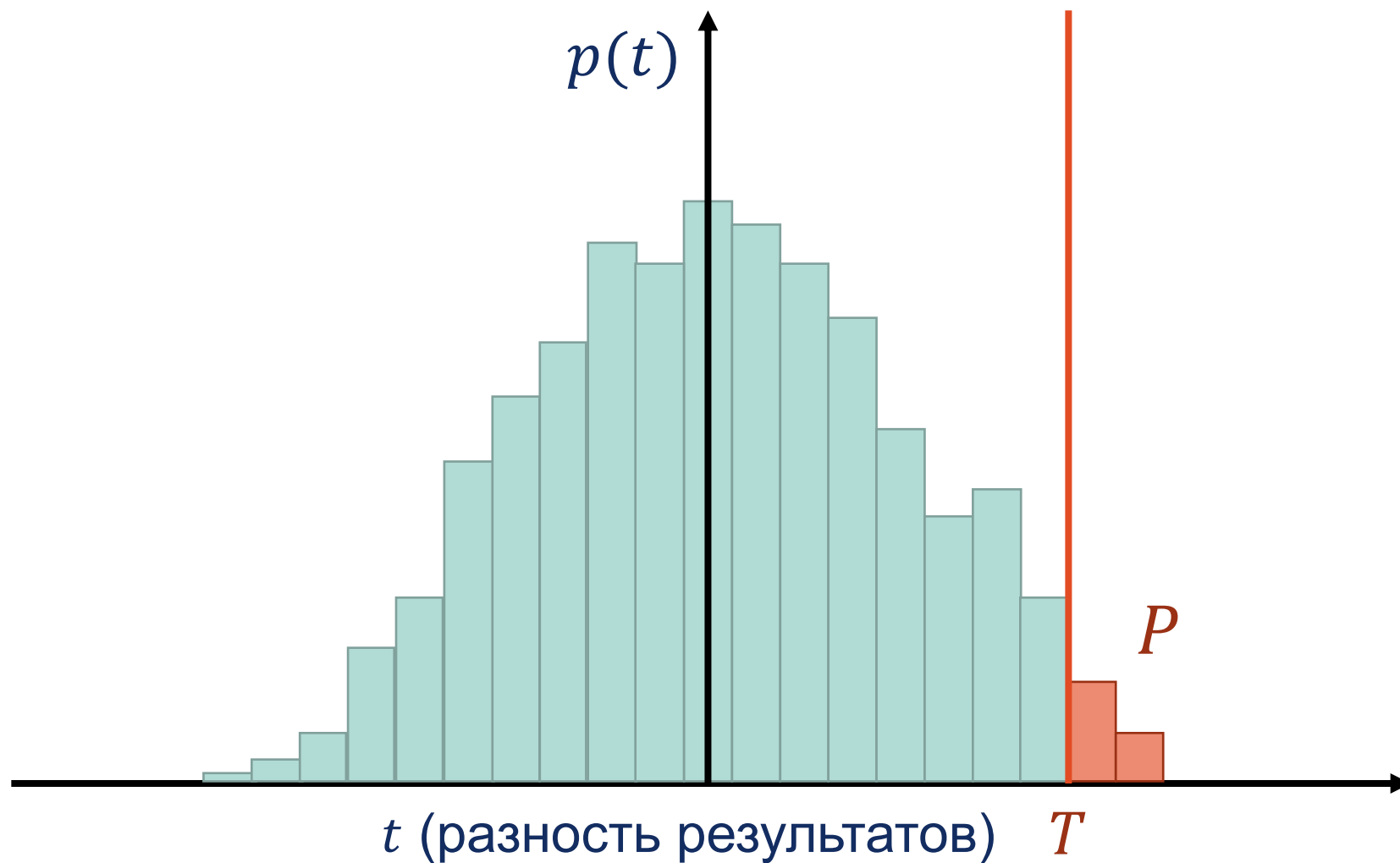
Гистограмма распределения из А/А тестов



Разность из А/В теста на гистограмме

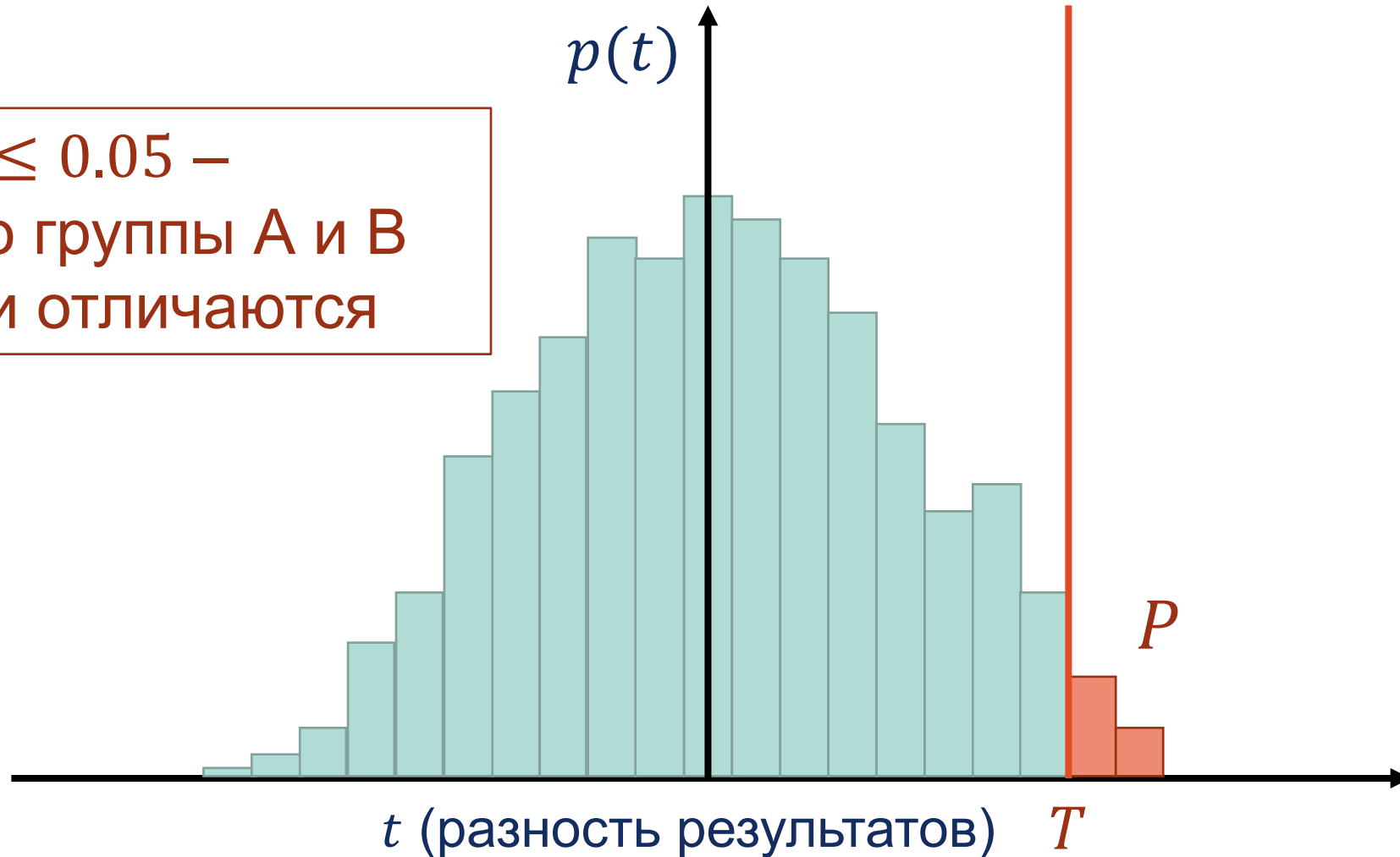


Вероятность отклонения в А/А



Вероятность отклонения в А/А

Если $P \leq 0.05$ –
наверно группы А и В
все-таки отличаются



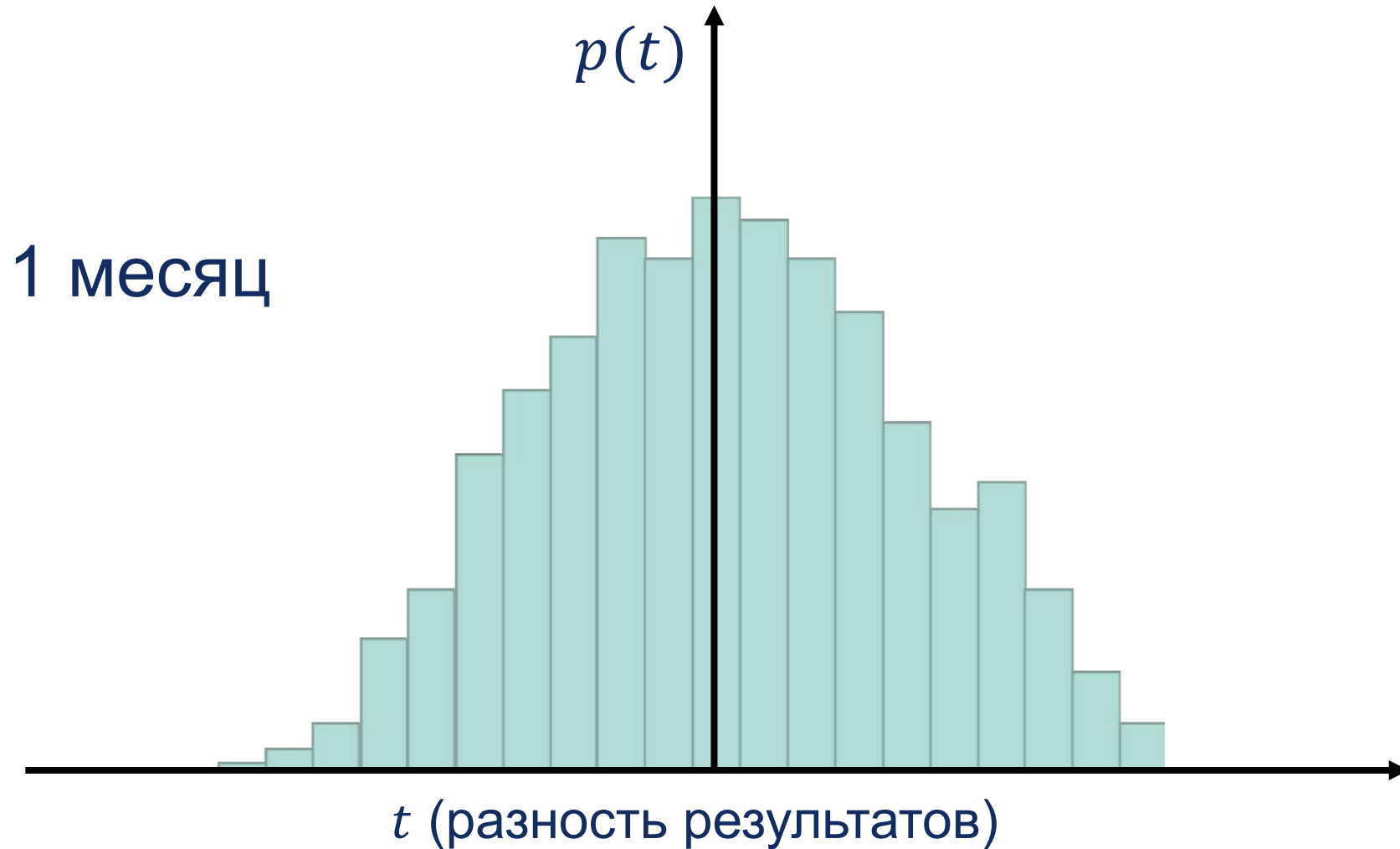
Планирование A/B теста

1. Решаете, насколько большое (по величине) улучшение метрики детектируете

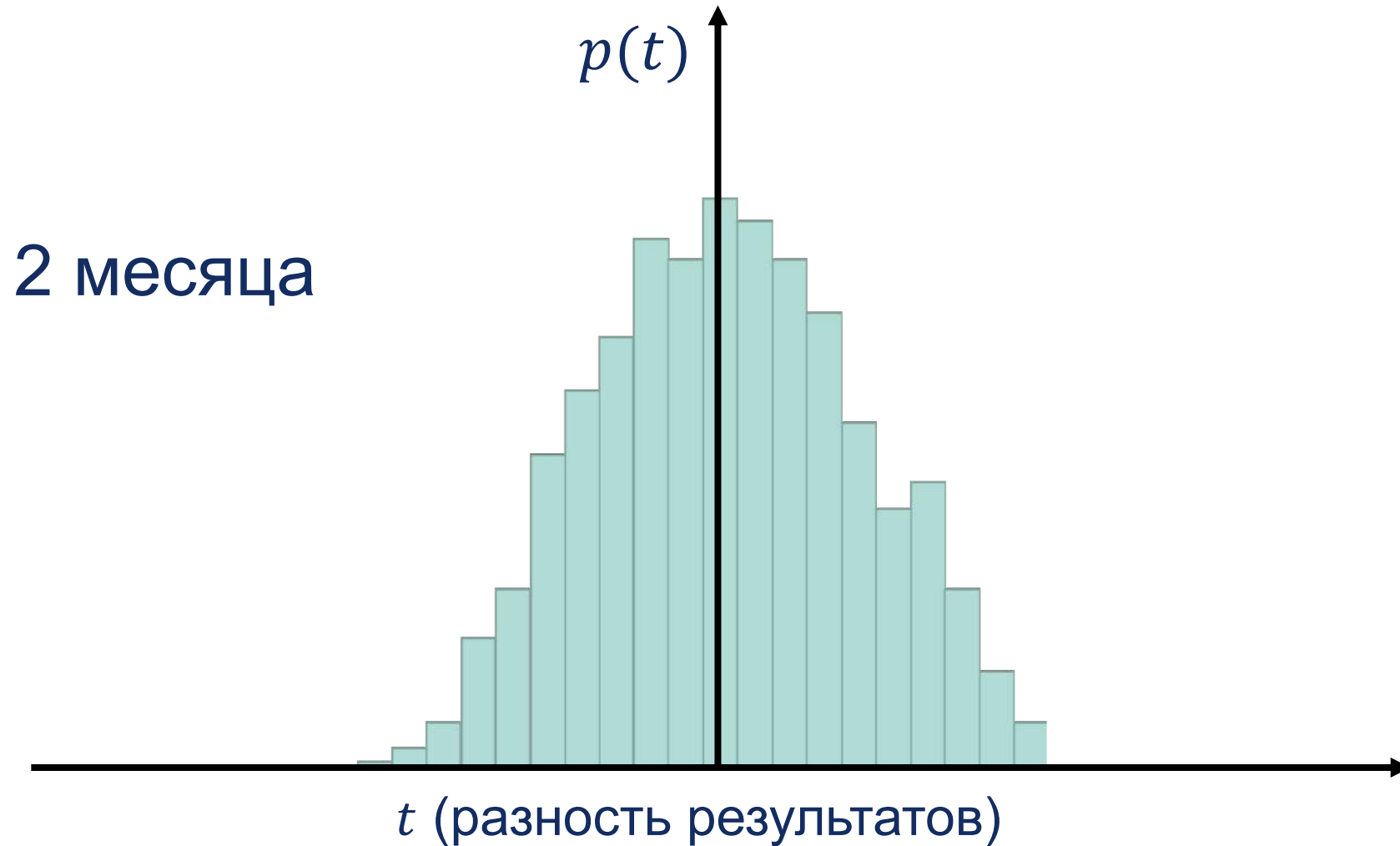
Планирование A/B теста

1. Решаете, насколько большое (по величине) улучшение метрики детектируете
2. По оценкам статзначимости (например, бутстрепом) на исторических данных понимаете, сколько данных нужно (=какая продолжительность у A/B теста), чтобы это улучшение алгоритма было статзначимым на нужном уровне значимости

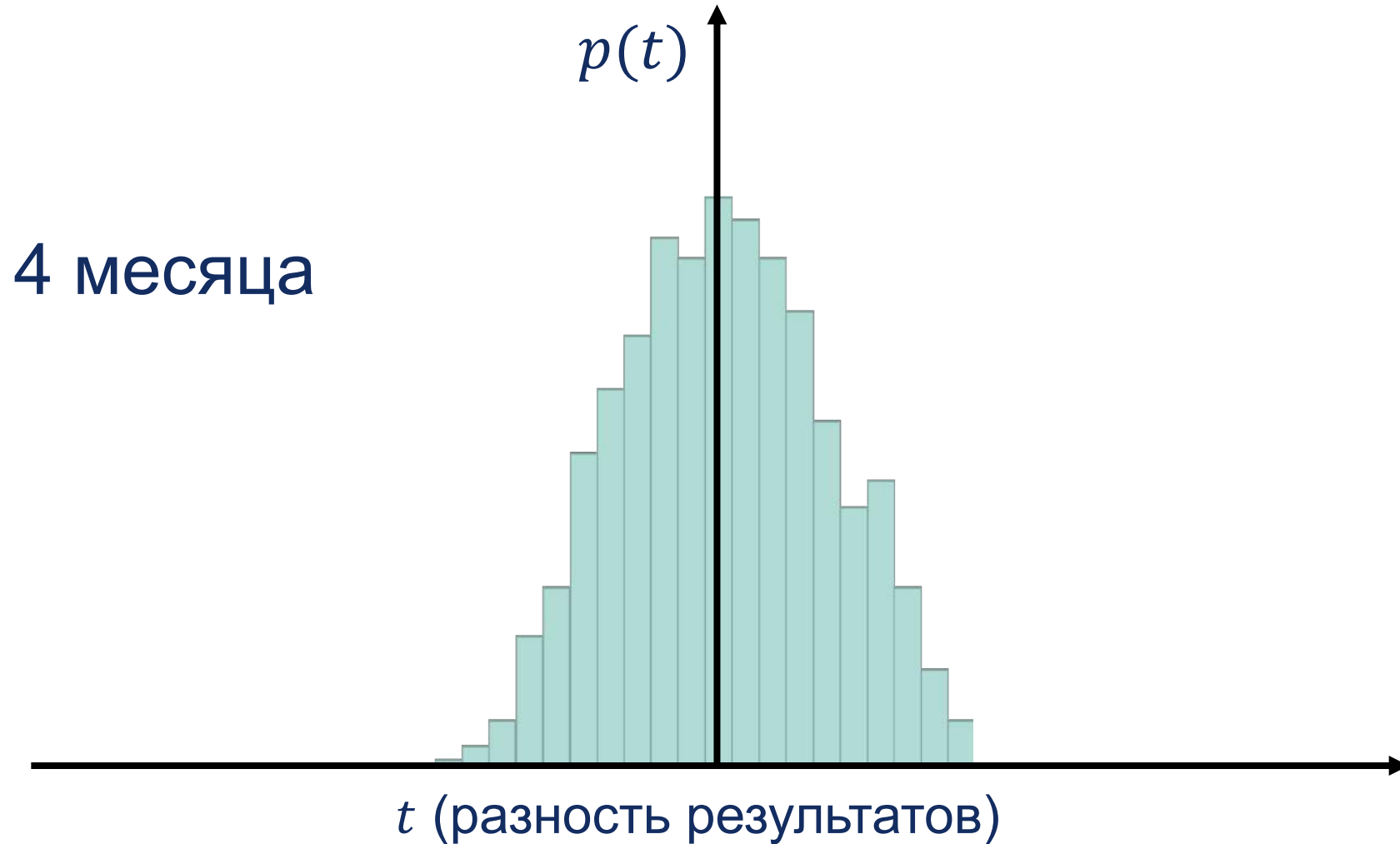
Как подбирать длительность А/В теста



Как подбирать длительность А/В теста

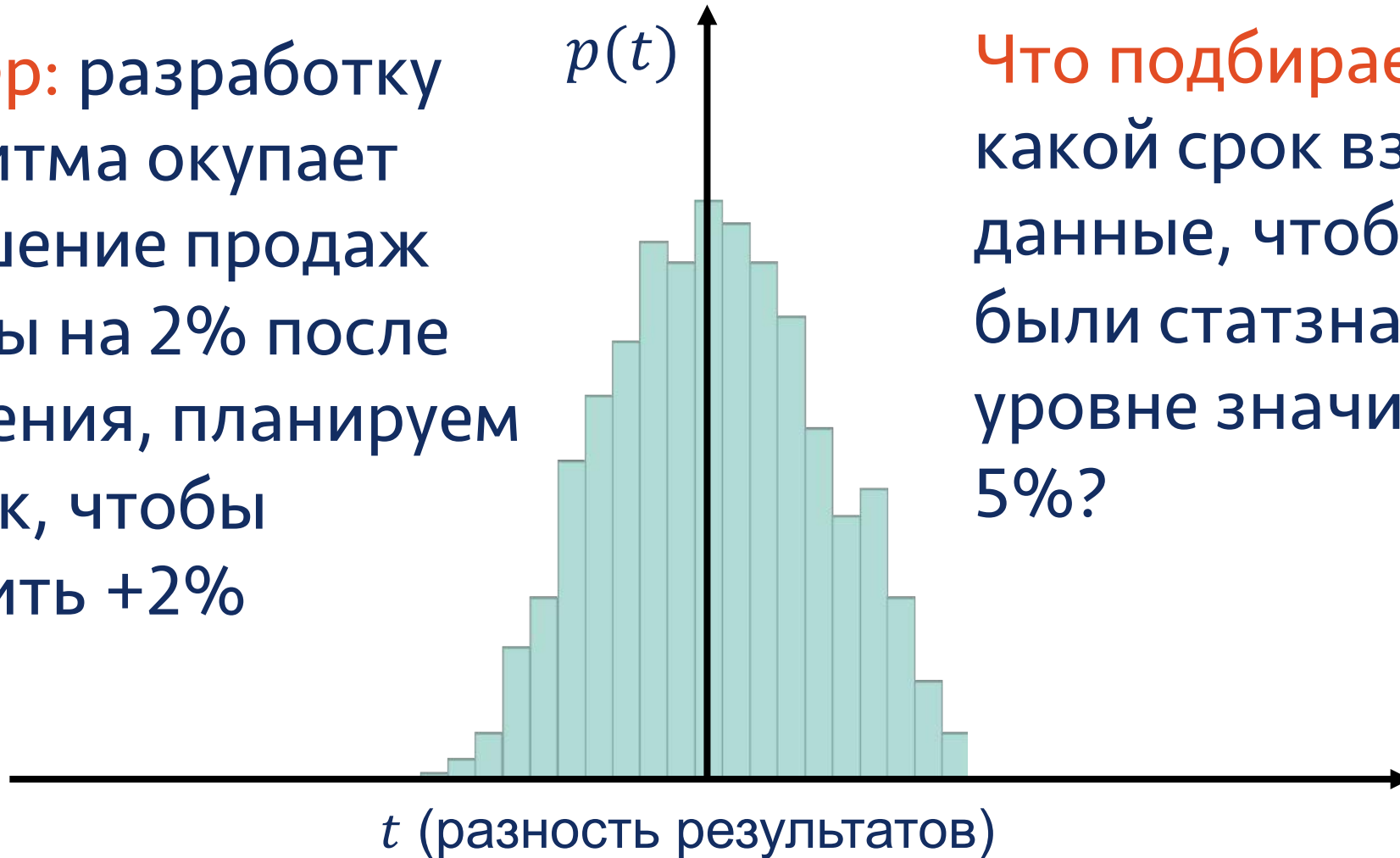


Как подбирать длительность А/В теста



Как подбирать длительность A/B теста

Пример: разработку алгоритма окупает повышение продаж хотя бы на 2% после внедрения, планируем A/B так, чтобы заметить +2%



Что подбираем: за какой срок взять данные, чтобы +2% были статзначимы на уровне значимости 5%?

Планирование A/B теста

1. Решаете, насколько большое (по величине) улучшение метрики детектируете
2. По оценкам статзначимости (например, бутстрепом) на исторических данных понимаете, сколько данных нужно (=какая продолжительность у A/B теста), чтобы это улучшение алгоритма было статзначимым на нужном уровне значимости
3. Помните про сезонность, округляете продолжительность теста хотя бы до недель

Планирование A/B теста

1. Решаете, насколько большое (по величине) улучшение метрики детектируете
2. По оценкам статзначимости (например, бутстрепом) на исторических данных понимаете, сколько данных нужно (=какая продолжительность у A/B теста), чтобы это улучшение алгоритма было статзначимым на нужном уровне значимости
3. Помните про сезонность, округляете продолжительность теста хотя бы до недель
4. Перед A/B тестом полезно провести A/A, чтобы проверить, настолько ли похожи результаты в группах, как на исторических данных, а возможно – даже проверить, не срабатывают ли ваши критерии в A/A тесте

Планирование A/B теста

1. Решаете, насколько большое (по величине) улучшение метрики детектируете
2. По оценкам статзначимости (например, бутстрепом) на исторических данных понимаете, сколько данных нужно (=какая продолжительность у A/B теста), чтобы это улучшение алгоритма было статзначимым на нужном уровне значимости
3. Помните про сезонность, округляете продолжительность теста хотя бы до недель
4. Перед A/B тестом полезно провести A/A, чтобы проверить, настолько ли похожи результаты в группах, как на исторических данных, а возможно – даже проверить, не сбываются ли ваши критерии в A/A тесте

С учетом перезапусков из-за ошибок – фактические сроки могут быть еще в 2-3 раза больше

Кейс: last-click и рекомендации

Задача: построить рекомендательную систему для магазина одежды



Кейс: last-click и рекомендации

Задача: построить рекомендательную систему для магазина одежды

Онлайн-метрика: количество товаров, купленных «с блока рекомендаций»



Кейс: last-click и рекомендации

Задача: построить рекомендательную систему для магазина одежды

Онлайн-метрика: количество товаров, купленных «с блока рекомендаций»

Атрибуция: last-click



Кейс: last-click и рекомендации

Задача: построить рекомендательную систему для магазина одежды

Онлайн-метрика: количество товаров, купленных «с блока рекомендаций»

Атрибуция: last-click

Вопрос с подвохом:
показывать ли в рекомендациях
на странице товара этот же товар?



На какие метрики смотрят в рекомендациях

- **Доход (прибыль, если известна)** – но обычно «не значима»

На какие метрики смотрят в рекомендациях

- **Доход (прибыль, если известна)** – но обычно «не значима»
- **Количество покупок / конверсия в покупку** – и про покупки и бывает статзначимой

На какие метрики смотрят в рекомендациях

- **Доход (прибыль, если известна)** – но обычно «не значима»
- **Количество покупок / конверсия в покупку** – и про покупки, и бывает статзначимой
- **Конверсия в клик** – очень быстро «прокрашивается», но от денег дальше остальных метрик

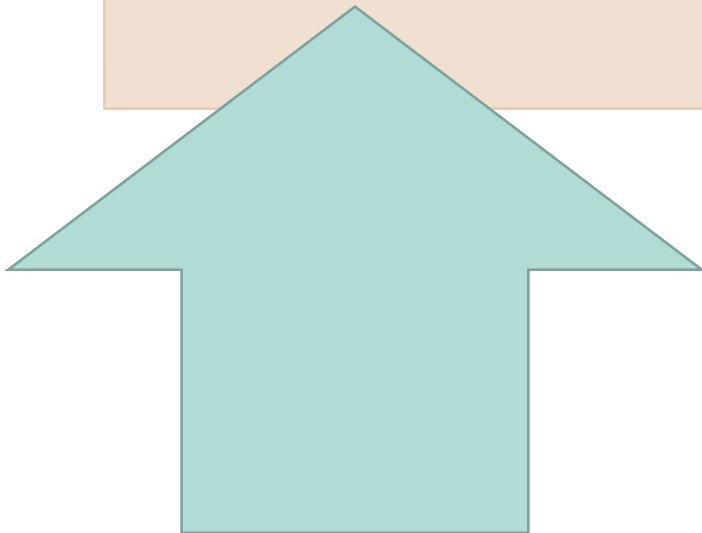
На какие метрики смотрят в рекомендациях

- **Доход (прибыль, если известна)** – но обычно «не значима»
- **Количество покупок / конверсия в покупку** – и про покупки и бывает статзначимой
- **Конверсия в клик** – очень быстро «прокрашивается», но от денег дальше остальных метрик
- **Last click / first click атрибуция** – быстро прокрашиваются, но не учитывают каннибализацию

Каннибализация трафика

Каталог:

- Верхняя одежда
- Трикотаж
- Обувь
- Аксессуары
- Нижнее белье
- Для дома

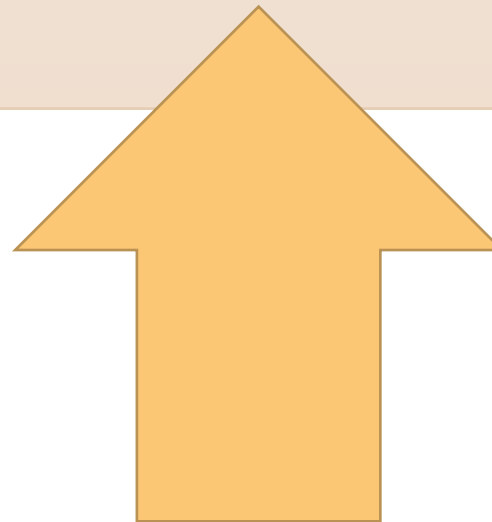
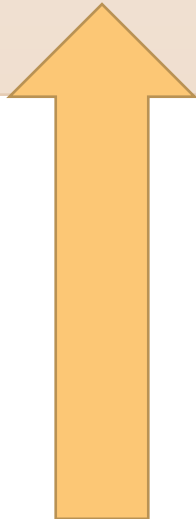


Каннибализация трафика

Каталог:

- Верхняя одежда
- Трикотаж
- Обувь
- Аксессуары
- Нижнее белье
- Для дома

Персональные рекомендации:



Оценка после внедрения

1. Решить, какие показатели важны после внедрения модели
2. Придумать, как оценить их изменение после внедрения (дизайн эксперимента)
3. Заранее подумать о том, как проверить, что измеренный эффект – не случайный (оценка стат. значимости) **Вопрос: до какой степени это возможно?**
4. Оценить необходимую длительность эксперимента
5. Проверить, что нет «нечестных» способов зависить результаты (может означать неудачный выбор KPI)

План

Таргет и прогнозы

Качество на исторических данных

Экономический эффект

Оценка после внедрения

Что входит в постановку задачи



Какое
целевое
значение



Как
используются
прогнозы



Как
оценивается
качество



Доп.
требования

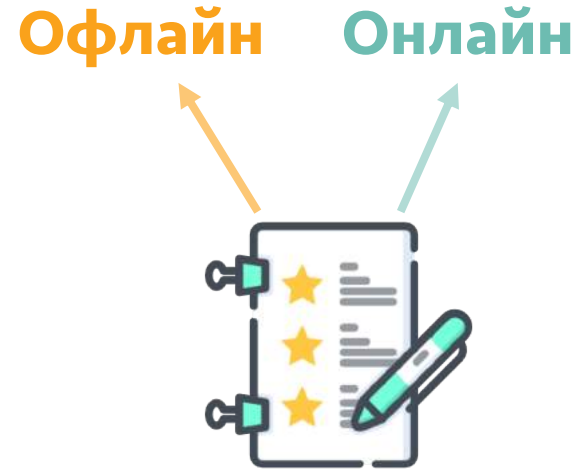
Что входит в постановку задачи



Какое
целевое
значение



Как
используются
прогнозы



Как
оценивается
качество



Доп.
требования

Что входит в постановку задачи



Что входит в постановку задачи



Что входит в постановку задачи



Что входит в постановку задачи



Домашнее задание

1. Ответить на вопросы по постановкам задач:
<https://goo.gl/forms/Aoi0027LZUwGLke02>
2. Придумать модель для оценки экономического эффекта в одной из рабочих задач, связать с метрикой качества модели машинного обучения (обсуждаем результаты 6 февраля)
3. Составить чек-лист для старта проекта по прогнозированию оттока. В результате прохождения по чек-листу должно становиться ясно, нужно ли начинать работать над проектом. (Обсуждаем 6 февраля)

Домашнее задание

4. Выполнить задания в Jupyter Notebook и прислать ФИО и блокнот либо ссылку на свое решение, выложенное на GitHub

Ссылка на Jupyter Notebook: https://github.com/utd-ai/workshops2019/blob/master/Workshop1_TaskFormulation_HWpart4.ipynb

Результаты присылать на почту: viktor.kantor@phystech.edu

В следующий раз

6 февраля:

1. Разбираем домашнее задание
2. Обсуждаем планирование проектов

Следующий воркшоп – 15 февраля (инструменты и организация работы с кодом)