

ML Advanced

Байесовское A/B-тестирование



Проверить, идет ли запись

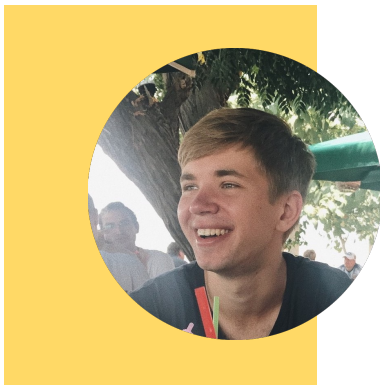
Напишите «+» в чат, если меня слышно и видно





Тема открытого урока

Байесовское A/B-тестирование



Бекетов Максим

Аспирант, стажер-исследователь ФКН ВШЭ;
ассистент кафедры дискр. математики МФТИ

Выпускник ФОПФ МФТИ, магистр
работал в анализе данных (appinteaair.com),
Python-разработке (Archeads Inc.),
преподаватель статистики, линейной алгебры в МФТИ и не только

преподаватель Байесовской статистики потоков ML-Advanced в ОТУС

Контакты: hse.ru/staff/bekemax, maxbeketov@outlook.com



Маршрут вебинара

1. Знакомство

2. Об OTUS

3. A/B-тестирование

4. Байесовское тестирование

5. Команда курса

6. О курсе, программа обучения

7. Бонус: карьерная информация

8. Рефлексия

Расскажите о себе

- Как вас зовут? Откуда вы?
- Ваш опыт работы в IT?
- С какой основной целью вы записались на занятие?



Правила вебинара



Активно
участвуем



Задаем вопрос
в чат



Вопросы вижу в чате,
могу ответить не сразу

Условные обозначения



Индивидуально



Документ



Ответьте себе или
задайте вопрос

06 OTUS

О компании



Сфера

ОТУС специализируется на обучении в IT. Наша фишка — продвинутые программы для специалистов с опытом и быстрый запуск курсов по новым набирающим популярность технологиям.



Клиенты

Наши партнеры современные технологичные компании. А обучение и открытые материалы привлекают специалистов разных грейдов: junior, middle, senior, lead.



Образование в ОТУС



Программы курсов

ОТУС имеет образовательную лицензию, поэтому наши курсы являются программами повышения квалификации и профессиональной переподготовки.

Направления курсов

Обучение специалистов разных грейдов: junior, middle, senior, lead



- Программирование
- Инфраструктура
- Тестирование
- Аналитика



- Data Science
- Управление
- GameDev
- Информационная безопасность

Мы в цифрах

130+

курсов для junior, middle, senior специалистов и менеджеров

600+

преподавателей делятся актуальными знаниями и реальными кейсами, востребованными в IT-индустрии



6

лет со дня основания компании

20 000+

выпускников уже прошли обучение по программам, адаптированным под запросы ведущих работодателей

430 000+

ИТ-специалистов в нашем сообществе, читают наши материалы, учатся и общаются на наших площадках



Напишите, пожалуйста, в чат подходящую цифру

- 1 - если уже учились у нас в компании
- 2 - если НЕ учились, но слышали о нас
- 3 - если впервые знакомитесь с OTUS



Байесовское А/В-тестирование

Цели вебинара

После занятия вы сможете



1. Использовать Байесовский вывод, чтобы встроить **априорные** представления в статистическую модель
2. Использовать Байесовское A/B-тестирование гипотез
3. Познакомиться с пакетом PyMC для Байесовского вывода

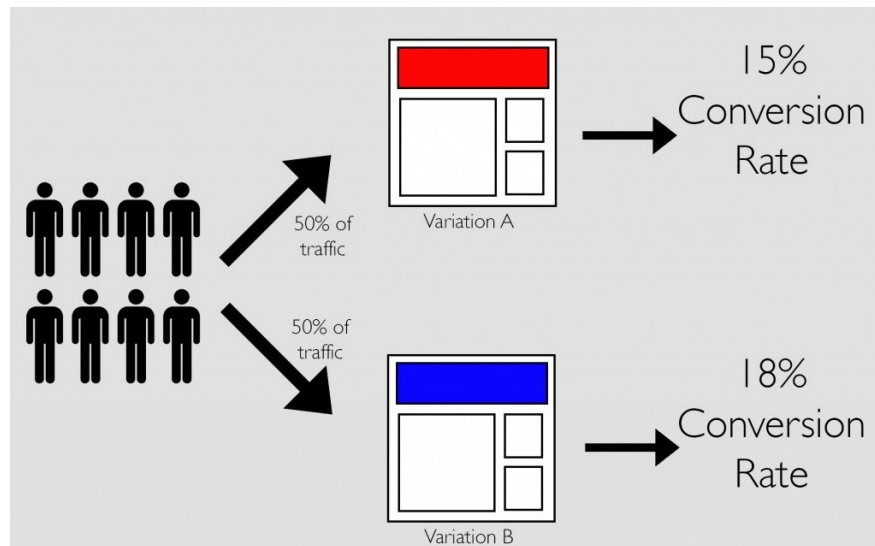


A/B-тестирование

A/B-тестирование

Разделение выборки на 2 подгруппы (подвыборки)

1. Измерение одного и того же параметра в обеих
2. Репрезентативность обеих подвыборок



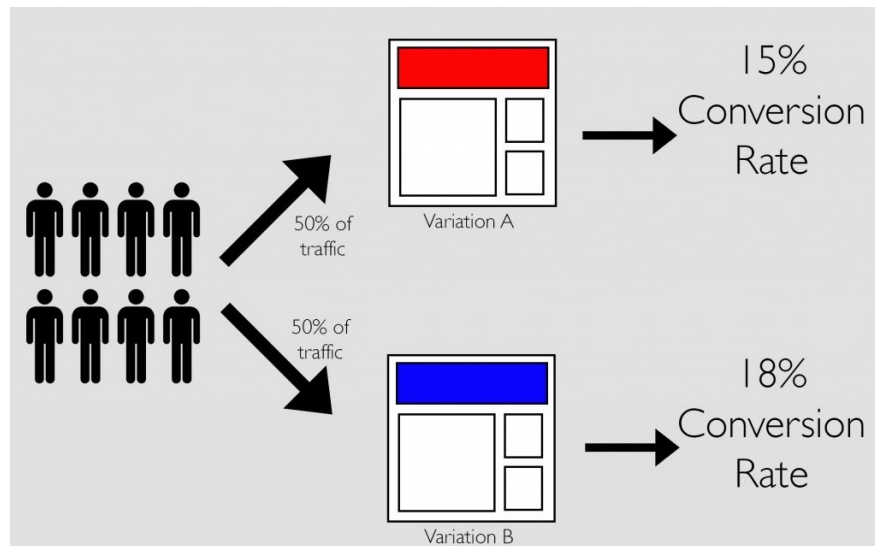
A/B-тестирование

Разделение выборки на 2 подгруппы (подвыборки)

Например, изменили элемент на сайте,
измеряем click-rate, θ_A и θ_B
 $\text{click} \sim \text{Bern}(\theta)$

Нуль-гипотеза $H_0: \theta_A = \theta_B$

Альтернатива $H_1: \theta_A \neq \theta_B$



Байесовская статистика

Байесовская статистика

В обычной статистике **оценка** $\hat{\theta}$ параметра θ –

это какая-то **функция** данных (выборки) X , $\hat{\theta}_n = f(X)$

Например, у нас

$$X_A = (\text{число кликов в группе } A) \sim \text{Binomial}(n_A, \theta_A)$$

Правдоподобие выборки: $p(X = k|\theta) = \binom{n}{k}\theta^k(1 - \theta)^{n-k}$

Оценка максимума правдоподобия (MLE): $\hat{\theta}_{MLE} = \arg \max_{\theta} p(X|\theta)$

Для биномиального правдоподобия просто: $\hat{\theta}_{MLE} = \frac{1}{n_A} \sum_{i=1}^n x_k$ – это **точечная** оценка
(мы пока не знаем ни её дисперсии, ни доверительного интервала)

Байесовская статистика

Формула Байеса:

$$p(\theta|X) = \frac{\overset{\text{likelihood}}{p(X|\theta)} \overset{\text{prior}}{p(\theta)}}{\underset{\text{evidence}}{p(X)}}$$

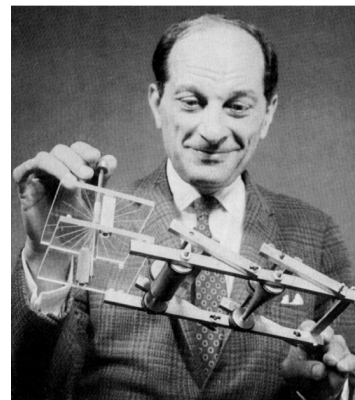
Наша цель – **posterior**, $p(\theta|X)$. Но evidence, $p(X) = \int p(X|\theta)p(\theta)d\theta$ – дается интегралом (/суммой) по всему пространству параметров θ , если оно многомерно – посчитать его непросто

Но наблюдение: $p(X)$ уже не зависит от θ ,
то есть $p(\theta|X) \propto p(X|\theta)p(\theta)$ – мы знаем **с точностью до множителя**

Тогда методами **Монте-Карло** сгенерируем выборку $\{\theta_1, \dots, \theta_M\}$ –
“гистограмма” этой выборки \approx желанный posterior



Томас Байес



Станислав Улам

Байесовское А/В-тестирование

Байесовское А/В-тестирование

Итак, у нас две группы, и количество кликов в них:

$$X_A \sim \text{Binomial}(\theta_A, n_A), \quad X_B \sim \text{Binomial}(\theta_B, n_B)$$

Нужно выбрать **априорное** распределение (prior), $p(\theta)$

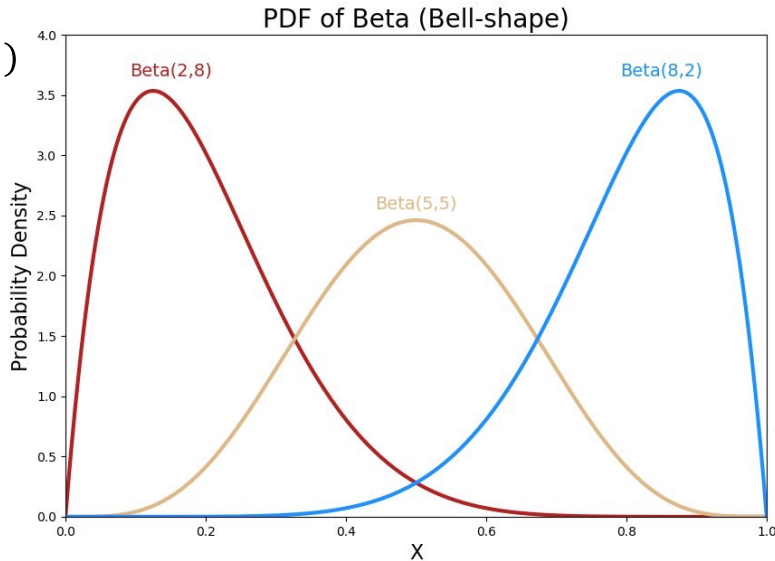
Можно выбрать **бета-распределение**:

$$p(\theta|\alpha, \beta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

Нормировочный множитель

$$\int_0^1 \theta^{\alpha-1}(1-\theta)^{\beta-1} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

– гамма-функция – обобщение факториала



Байесовское А/В-тестирование

На деле, бета распределение – **сопряженное (conjugate)** к биномиальному, то есть если:

$$\text{likelihood} \quad p(X|\theta) = \text{Binomial}(n, \theta) \quad \text{и} \quad \text{prior} \quad p(\theta) = \text{Beta}(\alpha, \beta), \text{ то}$$

$$\text{posterior} \quad p(\theta|X) = \text{Beta}(\alpha + \textcolor{red}{X}, \beta + (\textcolor{red}{n} - \textcolor{red}{X}))$$

– то posterior – **тоже** бета-распределение, просто с параметрами, **«уточненными/поправленными»** данными

– здесь X = число кликов («успехов») добавился к α , а $(n - X)$ – число не-кликов («неуспехов») добавилось к β ; потому α и β называют еще (априорными частотами) псевдо-успехов и неуспехов

Байесовское А/В-тестирование

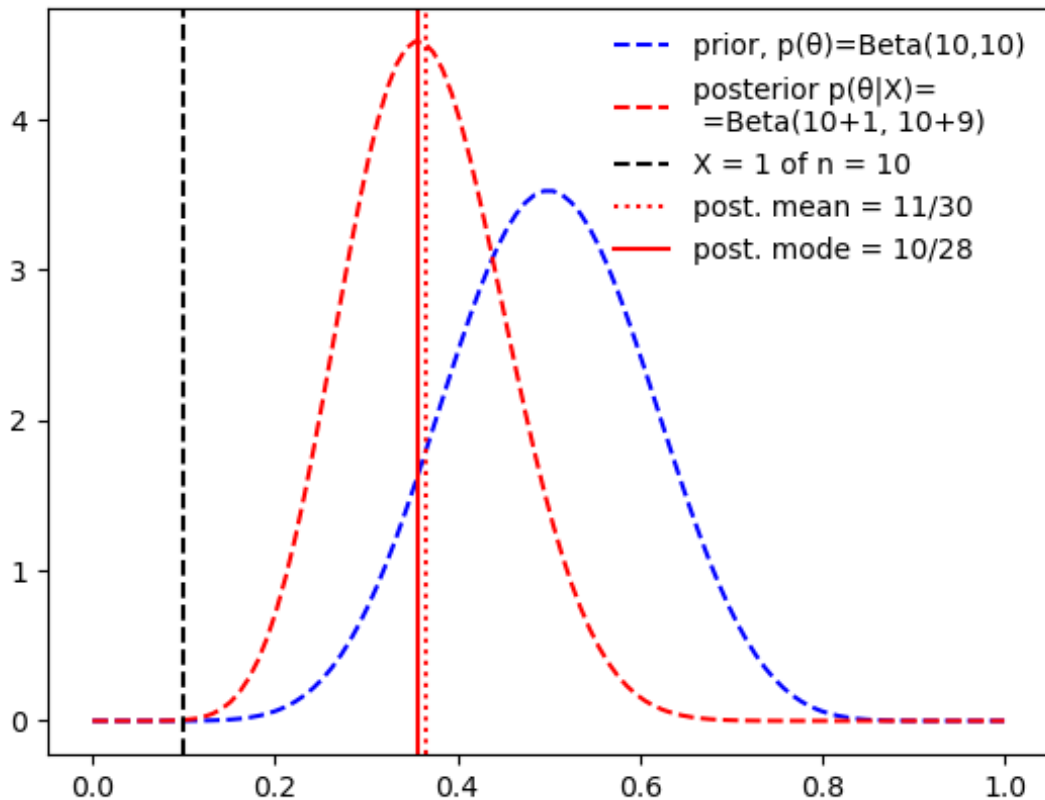
У нашего постериорного распределения *posterior*
 $p(\theta|X) = \text{Beta}(\alpha + X, \beta + (n - X))$

Можно взять, например, **матожидание** $\mathbb{E}(\theta|X) = \frac{\alpha + X}{\alpha + \beta + n}$

или **моду** $\arg \max p(\theta|X) = \frac{\alpha + X - 1}{\alpha + \beta + n - 2}$ – в качестве точечных оценок $\hat{\theta}(X)$

Известна и дисперсия. Но все же самое ценное – это что из данных, X , мы получаем **целое распределение** $p(\theta|X)$ возможных значений параметра θ – при условии данных X и с учетом наших априорных знаний $p(\theta)$ о его возможном распределении (безотносительно данных).

Байесовское А/В-тестирование



Байесовское А/В-тестирование

Мы с вами обсудили т.н. **бета-биномиальную модель**. Это редкий случай когда есть аналитическое решение.

Но пока мы знаем только распределение $\theta \sim p(\theta|X)$

А если нас интересует (относительное) увеличение click rate $= \frac{\theta_A}{\theta_B} - 1$?

Или, интереснее: каждый кликнувший пользователь оставляет случайное (например, экспоненциально распределенное $\sim \text{Exp}(\lambda)$) количество денег?

И нас интересует распределение $\text{Revenue} \sim \text{Bern}(\theta) \cdot \text{Exp}(\lambda)$?

Тогда уже не обойтись без Монте-Карло семплирования!

Байесовское A/B-тестирование

Посмотрим код!

PyMC.io: [What is \(Bayesian\) A/B testing?](#)

Байесовский t-test

Байесовский t-test

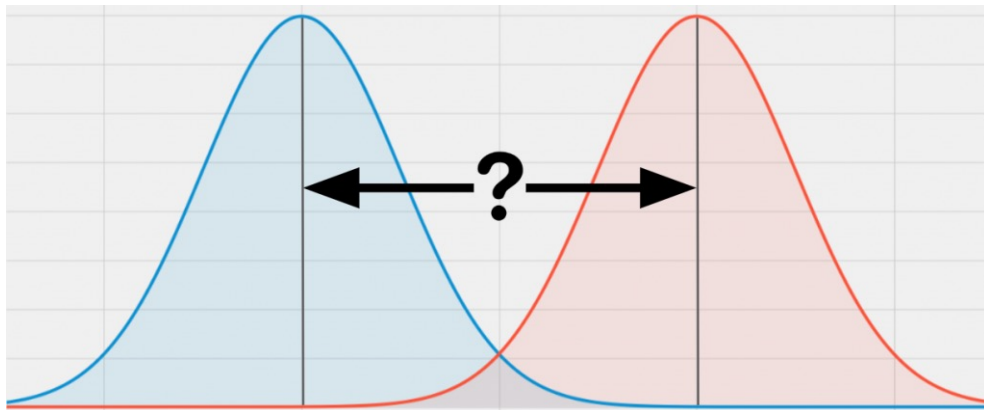
А если случайная величина не дискретная (клики), а непрерывная?

Например, мы испытываем лекарство для гипертоников – людей с повышенным артериальным давлением.

Для большой группы людей этот показатель будет почти нормально распределен

$$\text{pressure} \sim \mathcal{N}(\mu_A, \sigma)$$

Хотим проверить, есть ли разница между группами: $\mu_A - \mu_B = 0$?



Байесовский t-test

Уильям Госсет показал, что если

$$X_n = \{x_1, \dots, x_n\} \sim \mathcal{N}(\mu, \sigma)$$

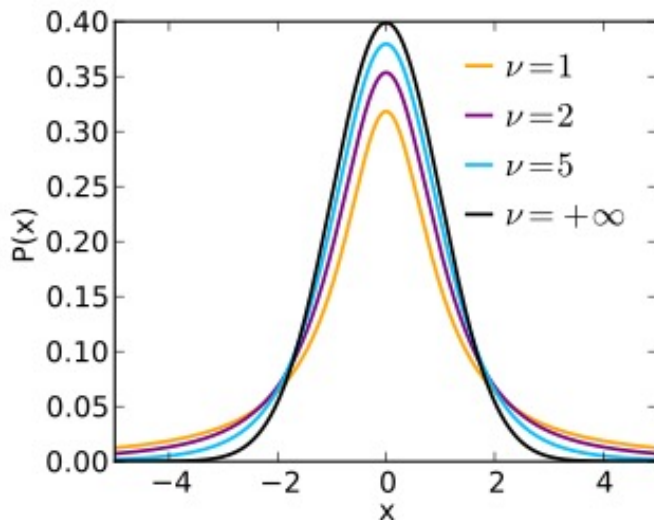
То t-статистика: $t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$ (где s – sample std.)

имеет особое распределение

$$t \sim \text{Student}(n - 1)$$

– с $\nu = n - 1$

«степеней свободы»



Уильям Госсет
(ака Стьюдент)

Байесовский t-test

Теперь для наших групп $X_{A,B} \sim \mathcal{N}(\mu_{A,B}, \sigma_{A,B})$ можно протестировать:

$$H_0: \mu_A = \mu_B \quad \text{против} \quad H_1: \mu_A \neq \mu_B$$

т.к. t-статистика $t = \frac{\bar{X}_A - \bar{X}_B}{s_p \sqrt{2/n}}$ где $s_p = \sqrt{(s_A^2 + s_B^2)/2}$

(при справедливости H_0) имеет t-распределение с $n_A + n_B - 2$ степенями свободы

Байесовская модификация этого теста позволяет нам **использовать априорные знания** о распределении $\mu_{A,B}$ и $\sigma_{A,B}$ – необязательна большая выборка испытуемых, если мы уверены в априорных знаниях!

Ключевые тезисы

1. Мы научились **«встраивать» априорные знания** в модель – если выборка небольшая, это поможет нам сделать вывод
2. Самое важное – постериор – не всегда легко получить, но нам на помощь приходят Монте-Карло и симуляции (семплирование) на мощном компьютере
3. Языки/библиотеки типа PyMC позволяют строить такие модели байесовского вывода «из кирпичиков»

Вопросы?



Задаем
вопросы в чат



Ставим “–”,
если вопросов нет

Знакомство с командой и программой курса

Процесс обучения



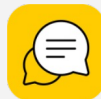
Обучение выстроено в формате вебинаров (онлайн). Онлайн-вебинары проводятся по вечерам или в выходные дни



Все записи занятий и материалы, предоставляемые преподавателями, сохраняются в личном кабинете и остаются доступны даже после окончания обучения



Домашние задания позволят Вам применить на практике полученные во время вебинаров знания. По каждому домашнему заданию преподаватель дает развернутый фидбек



В процессе обучения Вы можете задавать преподавателю вопросы по материалам лекций и домашних заданий, уточнять моменты, которые были непонятны на уроке



Время на обучение: от 4 ак. часов на занятия и 4-8 часов на домашнюю работу в неделю



Программа обучения на курсах обновляется каждый запуск в зависимости от актуальных запросов в сфере IT-технологий

Карьерная информация



Анализ позиции Data Scientist

156

Вакансий Data Scientist для
соискателей уровня Junior+ в Москве.

289

Data Scientist для соискателей
уровня Junior+ в России.



*вакансии для аудитории с опытом работы от года до трех лет. Источник — hh

Аналитика зарплатных предложений

(Junior - Senior)

Медиана

185 000



Источник — hh

Требования работодателей к соискателям



Навыки:

- Отличное знание классических алгоритмов Machine Learning, хороший кругозор в этой области;
- Умение разбираться в математике, стоящей за алгоритмами машинного обучения;
- Опыт работы с базами данных, SQL;
- Знание математической статистики (умение проверять статистические гипотезы);
- Есть опыт вывода ml-решений в продакшн.

Вакансии из telegram каналов

Middle/Senior Machine Learning Engineer

от 200 000 руб. до вычета налогов

Требуемый опыт работы: 3–6 лет

Полная занятость, полный день

Требования:

- Уверенное знание математической статистики;
- Глубокие знания в областях Machine Learning, Deep Learning;
- Опыт в сферах Computer Vision, Image Recognition;
- Опыт решения задач классификации, детектирования и сегментации с помощью CNN, знание и опыт применения архитектур SSD, YOLO и т.д.;
- Опыт разработки алгоритмов пост- и пред- процессинга изображений, знание методов openCV;
- Понимание метрик эффективности моделей машинного обучения;
- Знание и опыт использования Keras, tensorflow;
- Знание Python: multithreading, numpy, scikit, pandas, matplotlib и др.;
- Уверенное владение linux/bash;
- Опыт работы с git;
- Английский язык на уровне чтения технической литературы.

Плюсом являются:

- Опыт разработки коммерческих ML продуктов;
- Знание и опыт использования TFLite, техник оптимизации ML моделей;
- Ученая степень физико-математических или технических наук;
- Наличие научных статей по машинному обучению;
- Наличие сертификатов пройденных курсов по машинному обучению и анализу данных от ВШЭ, МФТИ, Яндекса;
- Fluent English B2+.

Вакансии из telegram каналов

Город и адрес офиса: Питер, м. Чкаловская/Горьковская или удалёнка;

Формат работы: гибрид

Занятость: полная

Зарплатная вилка: 2000\$ до 3000\$ net

Сейчас мы в поисках middle/senior **Data Scientist** для развития нашей рекламной платформы.

Чем предстоит заниматься?

- Построить эффективную ротацию рекламы на основе ML;
- Планировать, проводить и оценивать результаты A/B-тестов, выдвигать и проверять гипотезы по улучшению;
- Оптимизировать скорость и эффективность полученных моделей;
- Взаимодействовать с командой разработки.

Что мы ожидаем от тебя?

- Разбираешься в теории вероятностей и статистике;
- Знаешь, как работают традиционные алгоритмы, использующиеся для ML. Применял их в production'e;
- Уверенное знание Python для анализа данных;
- Опыт работы с TensorFlow;
- SQL (написание запросов, понимание принципов организации БД).

Рефлексия

Список материалов для изучения

1. [Байесовская статистика](#) – Википедия
2. [About PyMC](#), и вообще про т.н. Probabilistic programming (Stan и пр.)
3. Chris Stucchio, «[Bayesian A/B Testing at VWO](#)»
4. Материалы Сергея Николенко



0 курсе



Machine Learning. Advanced



старт обучения: 27.10.2023



otus.ru/lessons/advanced-ml



Заполните, пожалуйста, опрос о занятии

Важно! Пройти опрос могут только залогиненные пользователи платформы OTUS



otus.pw/IMZW/



Спасибо за внимание!