



# MLOps

Практика

Формирование процессов команды



Проверить, идет ли запись

# Меня хорошо видно && слышно?



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



Тема вебинара

# MLOps

## Цели и метрики ML проектов



**Шиповников Георгий**

**Lead ML Engineer @Xstack**

**Опыт:**

Комерческий опыт работы на позиции ML Engineer более 6-ти лет в таких компаниях как Сбер, ВТБ, Открытие.

**Домены** - Classic ML, Deep Learning, Time Series, Anomaly Detection, Credit scoring, Uplift & Response, ReqSys & Dynamic pricing, MLOps

**Telegram: @Shipovnikovvv**



# Правила вебинара



Активно  
участвуем



Off-topic обсуждаем  
в учебной группе в tg



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

# Карта курса



# Маршрут вебинара

Постановка целей и выбор метрик

Подходы и команда

SMART

Домашнее задание

Заключение



# Цели вебинара

К концу занятия вы сможете

1. Выбирать бизнес и DS метрики под проект
  2. Формулировать задачи / гипотезы по SMART
-

# Смысл

## Зачем вам это знать

1. Закрепим информацию по всему модулю
  2. Для того, чтобы выполнить оцениваемое домашнее задание
-



# Постановка целей и выбор метрик

# Практика

1. Вы создаете рекомендательную систему для маркетплейса
2. Была организована встреча с бизнес-заказчиком
3. Вам нужно выбрать метрики для проекта
4. Вы можете задавать любые вопросы
5. Напишите ваши предложения в чат
6. Объясните эти метрики заказчику\*



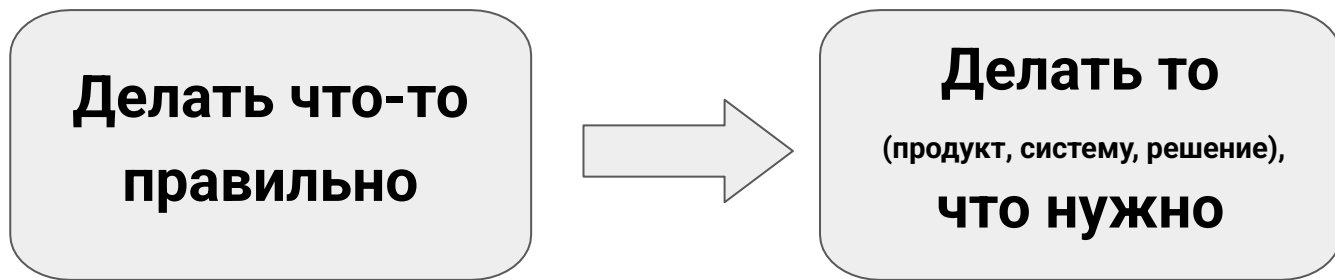
10 минут



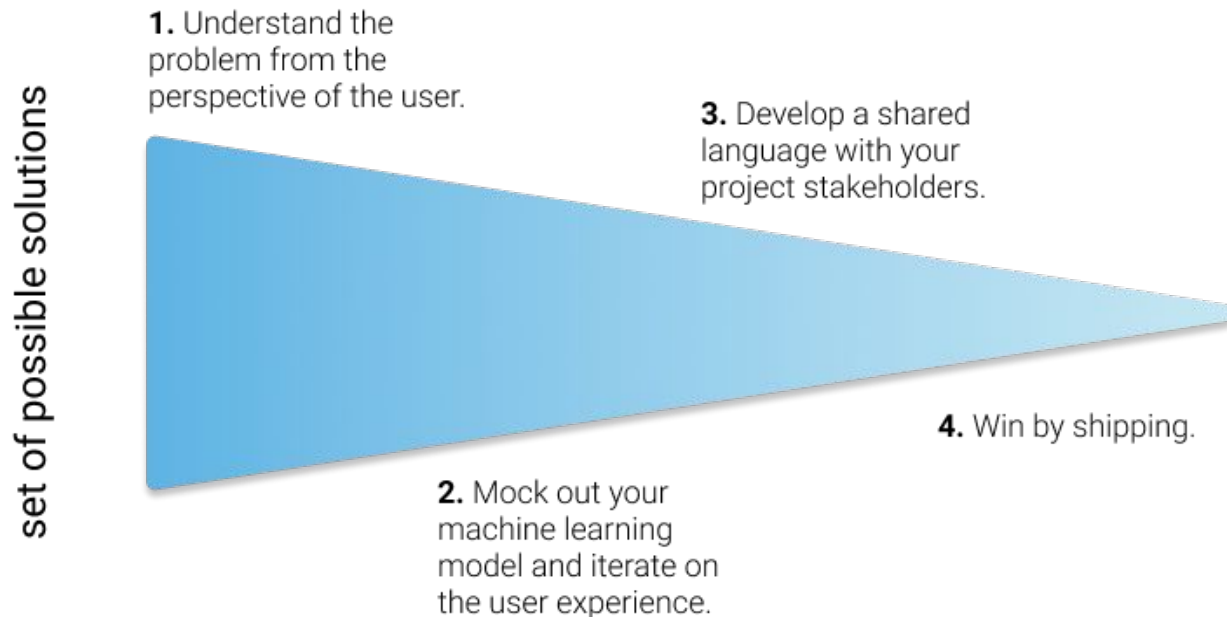
# Одна из ключевых причин провала ML-проектов / продуктов



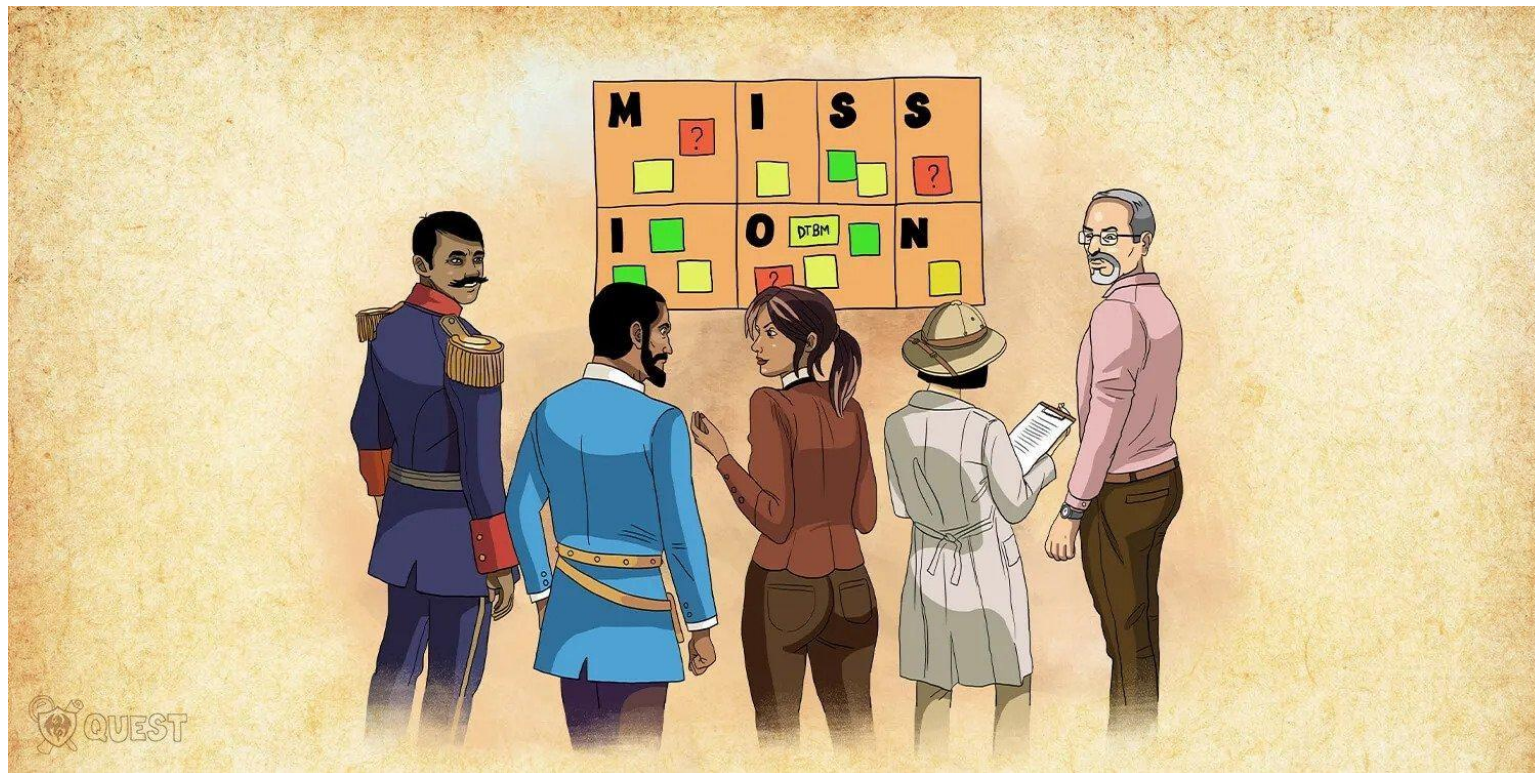
# На что обратить внимание



# Что может помочь



# Что может помочь













QUEST


[Статья с описанием](#)

# Пример: Machine Learning Canvas

**THE MACHINE LEARNING CANVAS (V1.0)** Designed for:  Designed by:  Date:  Iteration:

<b>PREDICTION TASK</b>  Type of task? Input object? Output: definition, parameters (e.g. prediction horizon), possible values?	<b>DECISIONS</b>  Process for turning predictions into proposed value for the end-user? Mention decision-making parameters.	<b>VALUE PROPOSITION</b>  Who is the end-user? What are their objectives? How will they benefit from the ML system? Mention workflow/interfaces.	<b>DATA COLLECTION</b>  Strategy for initial train set, and continuous update. Collection rate? Holdout on prod inputs? Output acquisition cost?	<b>DATA SOURCES</b>  Which raw data sources can we use (internal, external)? Mention databases and tables, or APIs and methods of interest.
<b>OFFLINE EVALUATION</b>  Simulation of the impact of decisions/predictions? Which test data? Cost/gain values? Deployment criteria (min performance value, fairness)?	<b>MAKING PREDICTIONS</b>  When do we make real-time / batch pred? Time available for this + featurization + post-processing? Compute target?	<b>BUILDING MODELS</b>  How many prod models are needed? When would we update? Time available for this (including featurization and analysis)?		
<b>LIVE MONITORING</b>  Metrics to quantify value creation and measure the ML system's impact in production (on end-users and business)?		<b>FEATURES</b>  Input representations available at prediction time, extracted from raw data sources.		

machinelearningcanvas.com by Louis Dorard, Ph.D.

Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. 

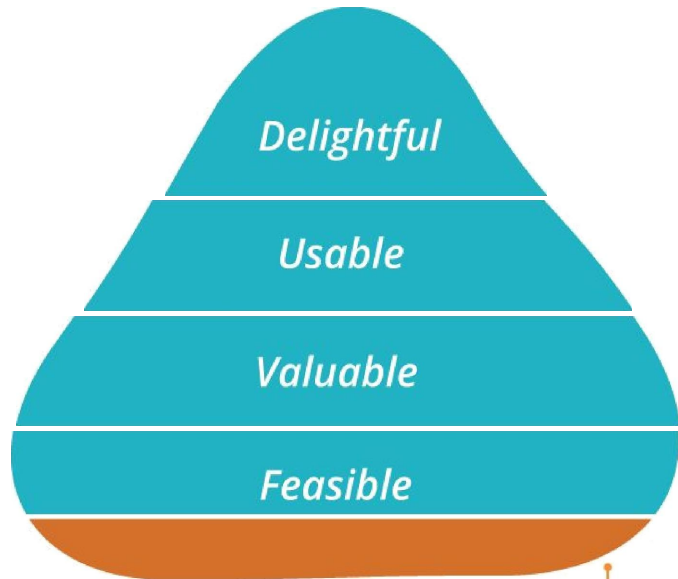
<https://www.ownml.co/machine-learning-canvas>

# Подходы и команда

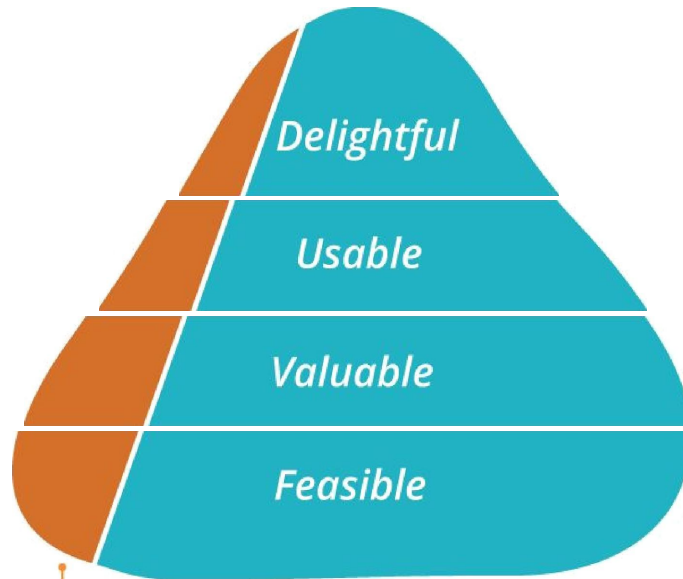


# Minimum Viable Product

**Минимально жизнеспособный продукт (MVP)** - продукт, обладающий минимальными, но достаточными для удовлетворения первых потребителей функциями.



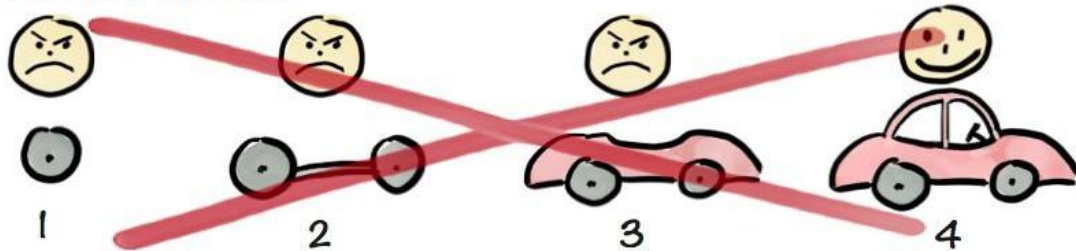
Not this



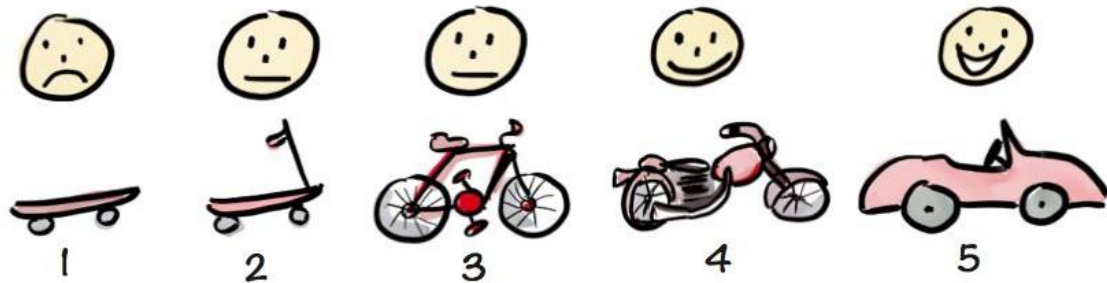
This

# Vertical Slicing

Not like this....



Like this!



# Vertical Slicing

*Как могут быть связаны Vertical Slicing и структура команды в ML проектах?*

# Vertical Slicing

- **Кросс-функциональные команды:** Вместо разделения команды по специализациям (например, отдельно data scientists, ML engineers, DevOps), формируются небольшие кросс-функциональные команды. Каждая такая команда способна выполнить полный "вертикальный срез" проекта.
- **Итеративная разработка:** Команды работают над небольшими, но полными функциональными возможностями. Например, вместо того чтобы одна группа долго собирала данные, другая создавала модель, а третья занималась развертыванием, команда может быстро пройти весь путь от простой модели до её развертывания.
- **Быстрое создание прототипов:** Vertical Slicing позволяет быстро создавать работающие прототипы ML-решений. Это может включать простую модель, базовую инфраструктуру для обучения и минимальный интерфейс для демонстрации результатов.

# Vertical Slicing

- **Фокус на бизнес-ценности:** Каждый "срез" ориентирован на конкретную бизнес-задачу. Это помогает команде сосредоточиться на создании ценности, а не на технических аспектах.
- **Непрерывное улучшение:** После создания базовой версии, команда может итеративно улучшать каждый аспект решения - от качества данных до сложности модели и масштабируемости инфраструктуры.
- **Гибкость и адаптивность:** Такой подход позволяет команде быстро адаптироваться к изменениям требований или новым данным, так как каждый "срез" является законченным и функциональным.
- **Улучшенная коммуникация:** Работа над вертикальными срезами способствует лучшему пониманию всего процесса ML-разработки каждым членом команды, улучшая коммуникацию и обмен знаниями.

# Что лучше делать вначале? Чек Лист

1. Выяснить, какую проблему пытается решить бизнес-заказчик / пользователь. Каковы его критерии успешности решения?
2. Можно ли решить задачу без ML? Будет ли решение на базе ML эффективнее?
3. При общении с бизнес-заказчиком / стейкхолдерами обратить внимание не только на возможную ML-модель, но на систему / решение целиком.
4. Привлекать заказчика / стейкхолдеров для проверки своих гипотез / уточнения информации.
5. Использовать шаблон проекта, создать структуру.  
Подготовить входные документы.  
Readme / Contributing с ответами на основные вопросы.  
Объяснить цели / задачи и правила команде.  
Выявить возможные проблемные моменты / вопрос с ресурсами.
6. Использовать единый формат постановки целей / задач.

# SMART

# SMART

- **Specific** - конкретный
- **Measurable** - измеримый
- **Achievable** - достижимый
- **Relevant** - значимый
- **Time-bound** - ограниченный по времени



# Практика

1. Вы делаете пре-модерацию для доски объявлений
2. Вашей команде нужно сделать ML-модель для авто-категоризации товаров
3. Придумайте 2 гипотезы для такого решения
4. Сформулируйте их на основе SMART
5. Напишите ваши предложения в чат
6. Напишите в чате +, когда будете готовы



10 минут



# Еще один шаблон - в помощь

## Hypothesis canvas

<b>Business value</b> <i>What is the business value associated with this hypothesis?</i>	<b>Problem</b> <i>What is the problem we are trying to solve?</i>	<b>Customers</b> <i>Who is impacted by this problem?</i>
<b>Hypothesis</b> <i>What do we believe?</i> We believe that _____ will result in _____. We will know we've succeeded when _____.		<b>Metrics</b> <i>Record a baseline of key metrics for this hypothesis.</i>
<b>Solutions/Ideas</b> <i>How might we solve this problem?</i>	<b>Lessons learned</b> <i>Record our lessons learned.</i>	

# Домашнее задание

# Домашнее задание - сдается на оценку

1. Изучите описание задания в материалах вебинара
2. Создайте GitHub репозиторий
3. Все артефакты сохраните в репозитории
4. Используйте [GitHub markdown](#) для документов
5. Пришлите в чат с преподавателем ссылку на репозиторий

# Заключение

# Цели вебинара

К концу занятия вы сможете

1. Выбирать бизнес и DS метрики под проект
  2. Формулировать задачи / гипотезы по SMART
-

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

Спасибо за внимание!

# MLOps

## Цели и метрики ML проектов



**Шиповников Георгий**

**Lead ML Engineer @Xstack**

### Опыт:

Комерческий опыт работы на позиции ML Engineer более 6-ти лет в таких компаниях как Сбер, ВТБ, Открытие.

**Домены** - Classic ML, Deep Learning, Time Series, Anomaly Detection, Credit scoring, Uplift & Response, ReqSys & Dynamic pricing, MLOps

**Telegram: @Shipovnikovvv**

