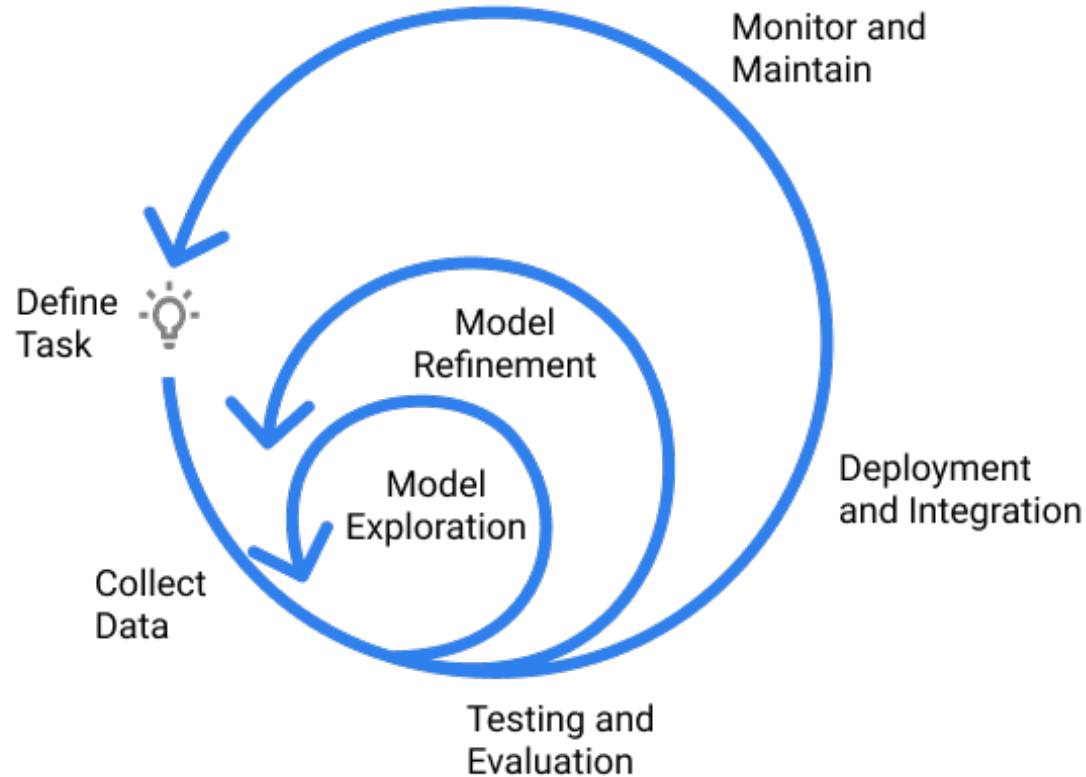


Обзор курса Production ML

Гущин Александр, весна 2021

Machine Learning Development Lifecycle




<https://www.jeremyjordan.me/ml-projects-guide/>

О чем наш курс

- Как превратить готовую ML модель в сервис и задеплоить
- Как настроить CI/CD пайплайн, который будет выкатывать изменения в сервисе в продакшн
- Как написать ML пайплайн, переобучающий модель, и добавить это в CI/CD
- Как организовать мониторинг моделей в проде и работать с логам
- Как трекать проведенные эксперименты, сохранять обученные модели и принимать решение об обновлении
- Как обеспечить стабильность и надежность ML-систем с помощью тестирования

Входные требования

- Базовые практические знания ML
- Знание Python на уровне младшего Python-разработчика
- Свободное время: 3 часа вебинаров в неделю и 5-15 часов самостоятельных занятий
-  Будет легче, если у вас есть опыт работы с *nix-системами, вы знаете Docker, сети, web-сервисы (Flask, REST API)

Что будет являться результатом прохождения

- Познакомимся с основными принципами DevOps и узнаем о специфике MLOps
- Поработаем с актуальными продакшн-инструментами, используемыми в ML
- Обсудим выбор подходящих инструментов в рамках конкретной задачи
- Выполним проект, о котором можно рассказать на собеседовании:



Контент

- Типичная тема состоит из лекции, семинара и домашнего задания.
- Мы начинаем курс с участия в соревновании про сложность паролей - с ним мы выполним несколько первых домашних заданий.
- После этого мы займемся основным проектом нашего курса - ботом для Шляпы. Мы также устроим соревнование между ботами, в котором определяющую роль помимо самой логики алгоритма будет играть доступность сервиса, его надежность и скорость работы.

Проект - пароли

DMIA sport intro: How frequent is this password?

Kaggle competition: [Link](#)

Predict





Password

{{ password }}

Prediction

{{ prediction }}

Проект - шляпа



-  Бот для игры в Шляпу
-  ML Pipeline для переобучения модели
-  CI/CD
-  Monitoring & Alerting

Структура курса по неделям

1. Обзор отрасли. Введение в MLOps. Практики написания кода.
2. Системное администрирование. Менеджеры пакетов и виртуальных окружений Python. REST API и Flask.
3. Git, версионирование кода и артефактов. Docker и контейнеризация сервисов. Деплой ML моделей.
4. Основы хранения и версионирования данных. Сервера в интернете и облачные вычисления. Тестирование ML моделей и сервисов.
5. Архитектура ML-систем. Создание ML пайплайнов.
6. Воспроизводимость экспериментов. Мониторинг сервисов и ML моделей.
7. Построение CI/CD, автоматизация тестирования, сборки и деплоя пайплайнов и сервисов.
8. Подведение итогов курса. Обзор дальнейшего развития, задач и целей для ML Engineer.
9. Финал шляпы.

Рейтинг

Рейтингом будут служить итоговые лидерборды Шляпы:

-  Соревновательный
-  Технический

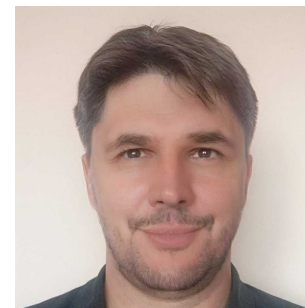
Авторы



Александр Гущин,
Tech Lead в
[Mechanica.ai](#)



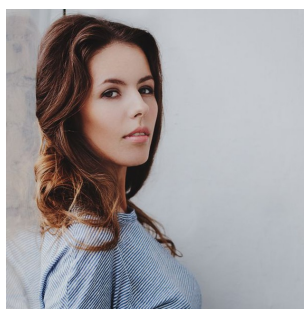
Михаил Трофимов, ML
Product Owner в Praxis



Глеб Ерофеев, Lead
ML Engineer в
Сбермаркет



Виталий Белов, DS в
[Mechanica.ai](#)



Ольга Филиппова, DS
Team Lead в банке
Открытие



Илья Ирхин, Chief DS
Yandex.Taxi