



**ЦЕНТРАЛЬНЫЙ  
УНИВЕРСИТЕТ**

# Защита проекта

Production ML

28.05.2025

# Тебелев Георгий

Data Scientist at Сбер

---



Выпускник МГТУ им. Баумана



Магистр ЦУ второй волны



Создавал космическую технику



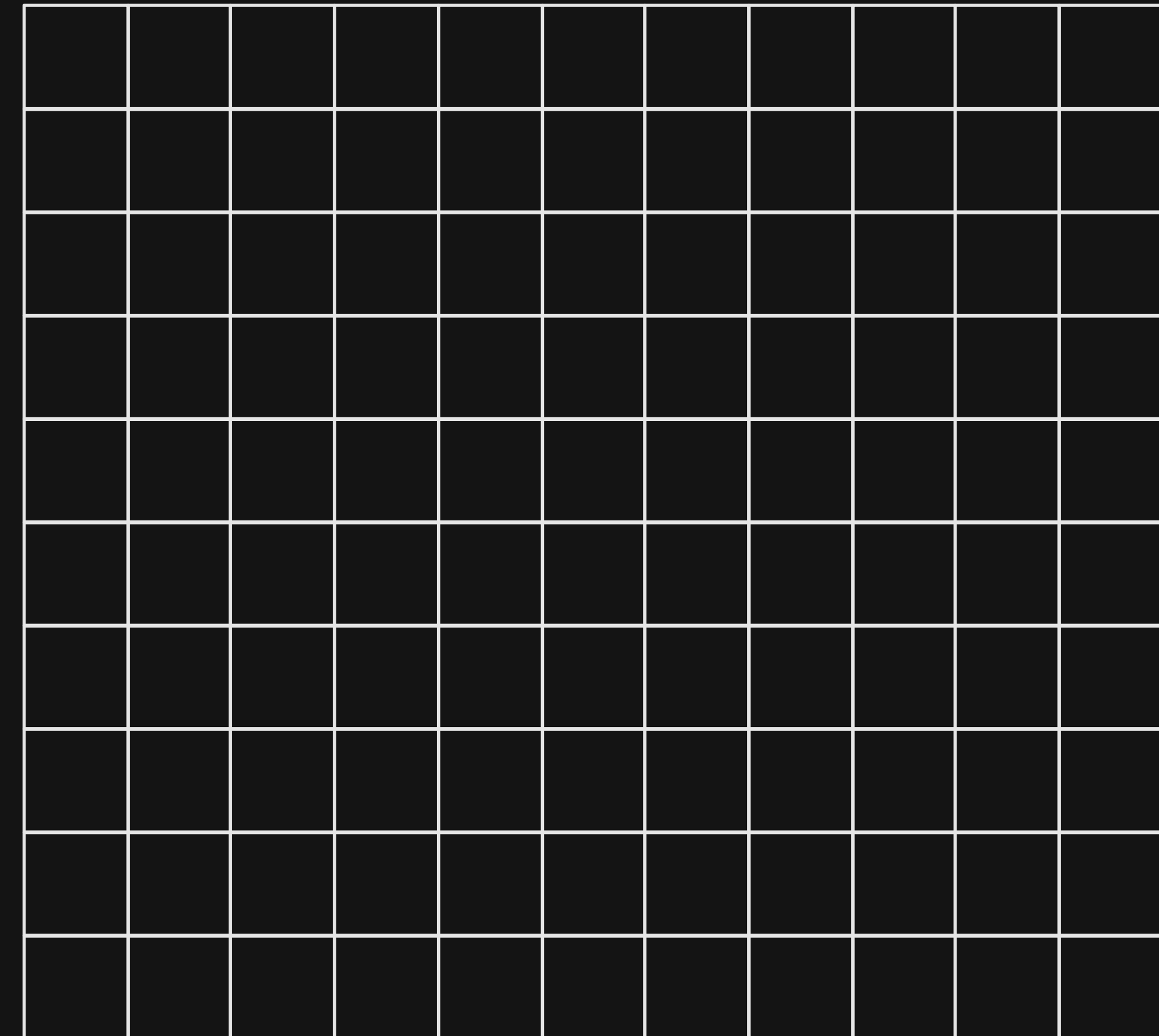
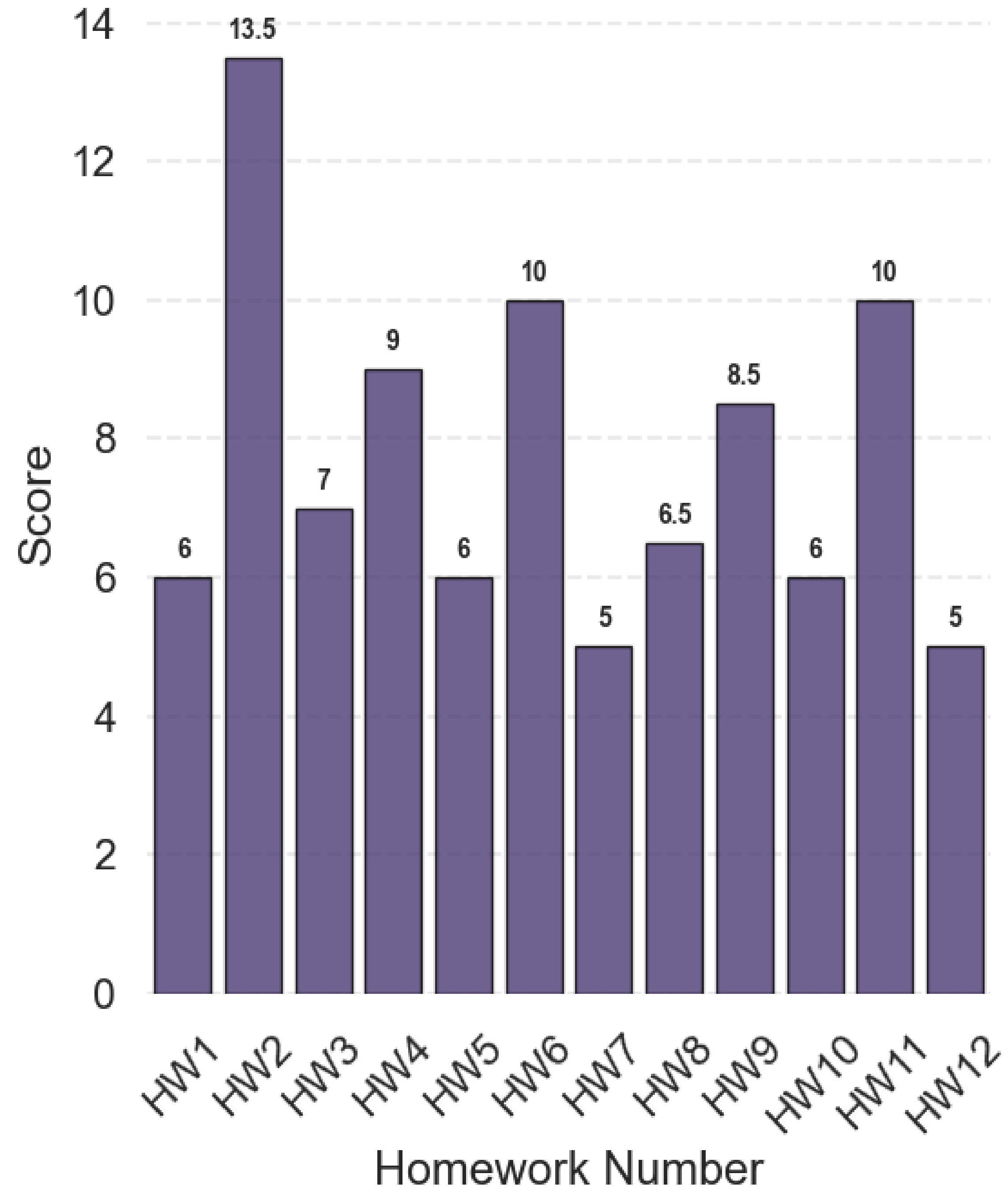
# My workflow



# My workflow



Homework Scores Distribution



Это был тернистый путь



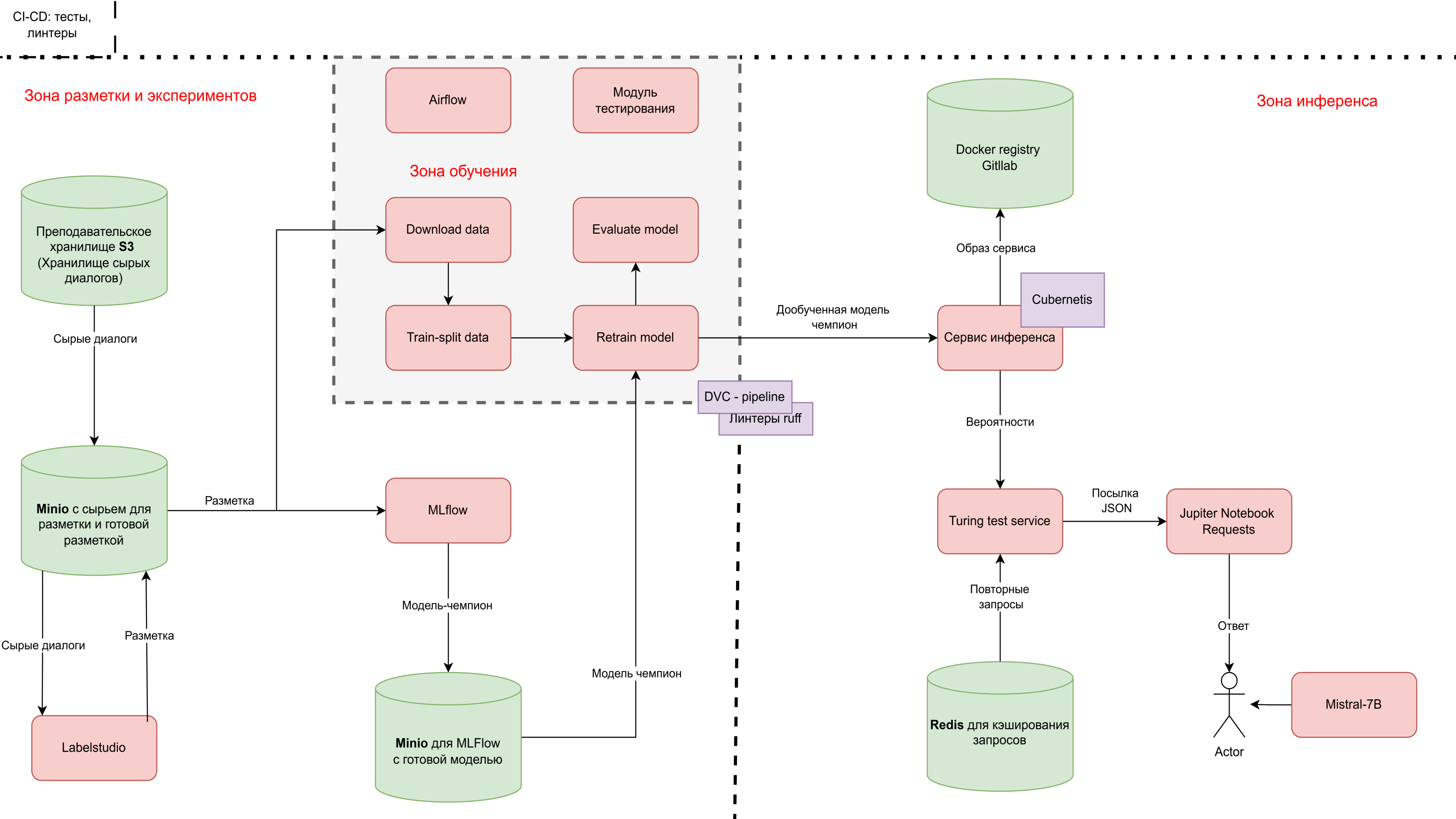
Что  
реализовано?

CI-CD: тесты,  
линтеры

Зона разметки и экспериментов

Зона обучения

Зона инференса





# В деталях



01

## Сервисы

12 Докер-контейнеров, 5 БД – все запускается одним docker-compose

02

## Airflow

Налажен пайп ежедневного дообучения модель чемпиона MLflow

03

## MLflow

В автоматическом режиме модель чемпион поступает в пайп дообучения и в модуль инференса

04

## Cubernetis

Сервис инференса поднят на подах кубернетиса для непрерывной работы сервиса во время обновления модели

05

## Бизнес логика отделена

Сервис инференса существует независимо от основного сервиса для удобства рефакторинга

06

## Тесты

90% кода покрыта тестами, налажены линтеры в dvc пайплайне и в CI-CD пайплайне



CONTAINERS

gtebelev

> gtebelev grafana grafana - Up 2 days

> prom/prometheus prometheus - Up 2 days

> prom/node-exporter node-exporter - Up 2 days

> heartexlabs/label-studio:latest labelstudio - Up 3 days

> minio/minio:latest minio-server-ls - Up 3 days

> postgres:latest postgres-server-ls - Up 3 days (healthy)

> gtebelev\_turing-test-service turing-test-service - Up ...

> redis:latest redis-cache - Up 3 days

> gtebelev\_inference\_service inference-service - Up 3 d...

> gtebelev\_mlflow gtebelev\_mlflow\_1 - Up 3 days

> minio/minio:latest gtebelev\_minio\_1 - Up 3 days

> postgres:15-alpine postgres-mlflow - Up 3 days

Task	Duration
download_raw_data	6d15:58:18
prep	6d15:58:18
train	13d07:56:36
eval	6d15:58:18

```
georgii-tebelev@georgii-tebelev-vm:~$ kubectl get pods
```

NAME	READY	STATUS	RESTARTS	AGE
inference-7c99795f65-62qdn	1/1	Running	1 (3d1h ago)	3d1h
inference-7c99795f65-7ks98	1/1	Running	1 (3d1h ago)	3d1h
nginx-544fd6dd95-65f5k	1/1	Running	2 (3d1h ago)	3d3h
nginx-544fd6dd95-6dmk2	1/1	Running	2 (3d1h ago)	3d3h
nginx-544fd6dd95-knszz	1/1	Running	2 (3d1h ago)	3d3h

	Run Name	Created
	legendary-boar-824	3 days ago
	suave-croc-587	3 days ago
	sincere-eel-482	3 days ago
	kindly-dove-948	3 days ago
	valuable-roo-588	3 days ago
	overjoyed-fox-817	3 days ago
	rare-bird-990	3 days ago
	sincere-newt-814	3 days ago
	rumbling-ape-733	3 days ago
	wistful-pig-867	3 days ago

9

# В деталях



07

## Labelstudio

Настроен сервис разметки с удобным интерфейсом

08

## CI-CD

Команда `git push` инициализирует линтеры `ruff` и тесты

09

## DVC

Дообучение модели можно инициализировать через `dvc repro`

10

## README

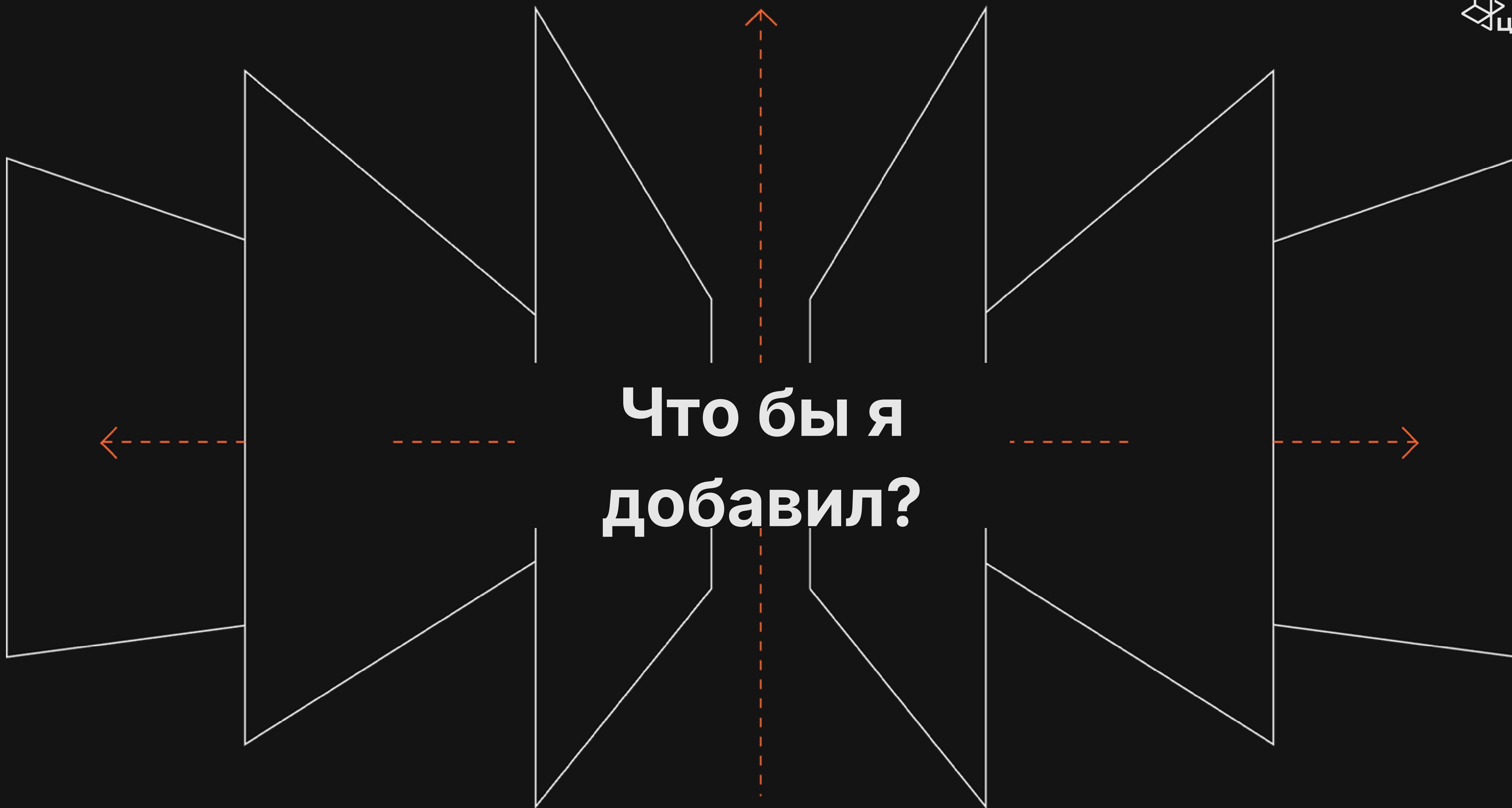
Подробное README с инфой по каждой домашке и аккуратный репозиторий

11

## Poetry

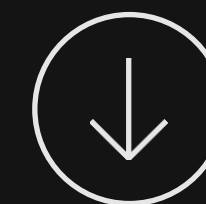
Все зависимости и либы удобно инициализированы в `poetry`

**Что бы я  
добавил?**



place	classifier_name	accuracy	roc_auc	log_loss	precision	recall
10	BotObormot	0.5256	0.3572	0.9704	0.0979	0.1894
11	Саша	0.5082	0.504	17.7264	0.4386	0.4717
12	Vadim	0.4961	0.4987	1.0009	0.4557	0.5039
13	МуслимНикита	0.4856	0.4898	0.9991	0.0624	0.4571
14	Илья Михеев	0.4846	0.5	0.7198	0.4846	1
15	Egor	0.4821	0.4865	0.6947	0.4845	0.9895
16	Уханов	0.4817	0.4677	1.0336	0.4411	0.4472
17	Даниил	0.4762	0.4817	1.6311	0.4581	0.7704
18	kiselevsemchikov	0.4607	0.4471	13.4425	0.4219	0.4181
19	Георгий	0.3148	0.3755	1.301	0.2609	0.2308

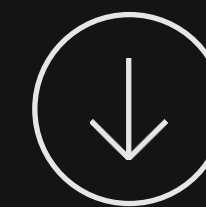
# Моя модель тупа



Так как 90% времени я возился с курсом в одиночестве, то совсем не осталось времени на модификацию модели

- Отказался бы от классической логистической регрессии и обратился бы к nlr подходам
- Возможно дообучил языковую модель для своей задачи, но это уже совсем другая история...

# Разобрался бы с triton

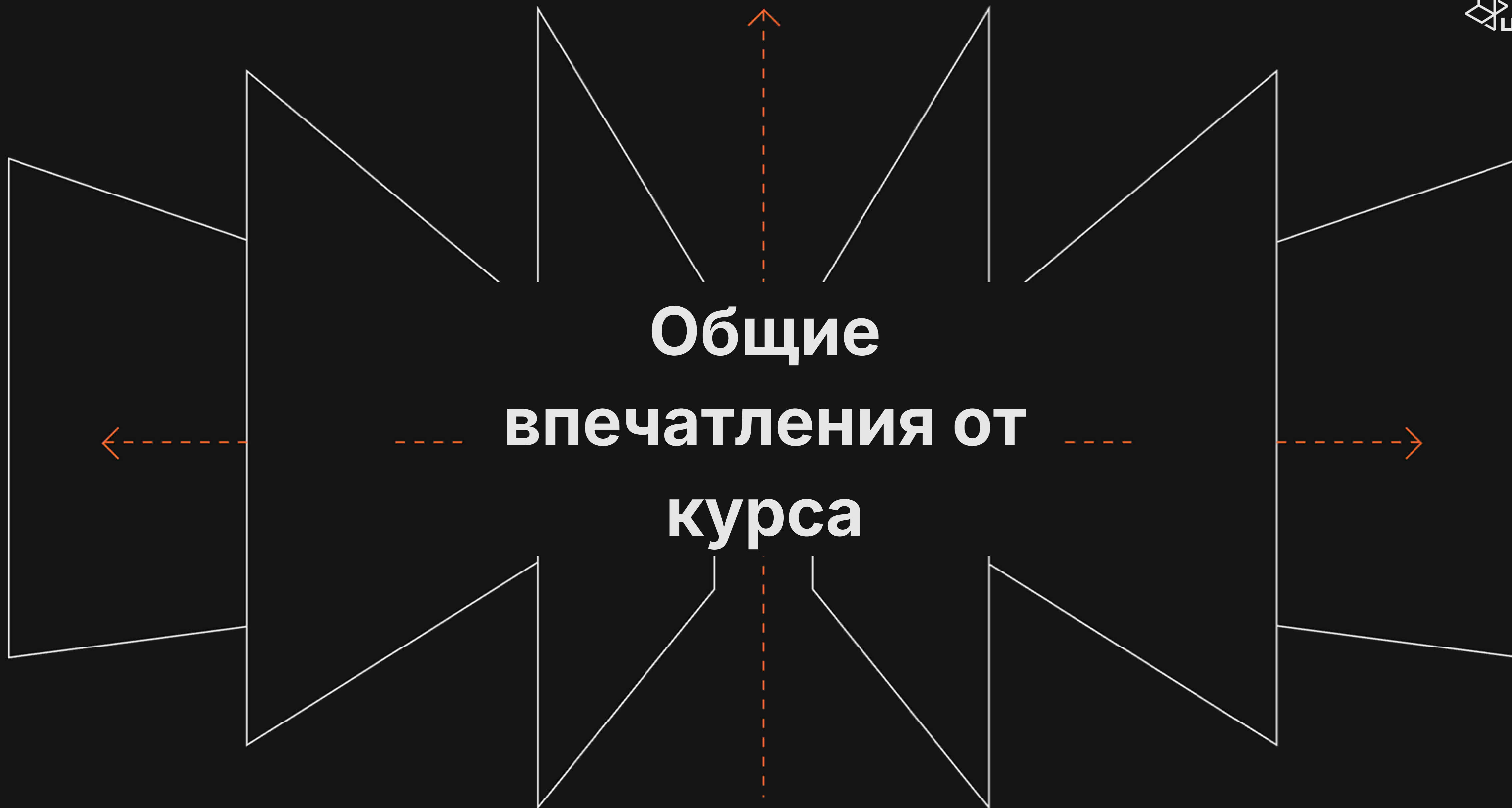


Какая-то неведомо сложная сущность, которую смогли реализовать лишь избранные команды

- Общее название группы животных отряда хвостатых земноводных.
- Реализую штуку в дипломе, просто потому что все вокруг в ужасе с неё.



# Общие впечатления от курса



# Шедевр! – лучший курс



Ассистенты: 10/10 – Отдельное спасибо Максиму Новикову, прости что писал тебе каждый день



Грейдер: 10/10 – Удобно, особенно после введения комментариев



Структура: 10/10 – Спикеры топ, материал топ, все топ



Применимость на работе: 10/10 – поднял несколько полезных сервисов для команды и Сбера



Сложность: 10e7/10 – в одиночку преодолеть тяжело



**Спасибо  
за внимание**

Тебелев Георгий

