

ПАВЛОВ ИГОРЬ

☎ +7-925-369-20-48 ✉ ivladipavlov@gmail.com 🌐 [boomboom](https://boomboom.com)

ОПЫТ РАБОТЫ

Sber AI Research

Август 2020 – Настоящее время

Machine Learning Engineer, MLOps

Москва

- Разрабатывал ETL пайплайны сбора и обработки мультимодальных данных для обучения text2image, text2video и других моделей.
- Разработал фреймворк для унифицированной обработки и фильтрации данных. Данный фреймворк используется различными командами в Сбере для фильтрации данных. [Код](#). [Статья](#)
- Руководил Data Engineering направлением, отвечал за коммуникацию и взаимодействие с research командой. Организовывал процессы сбора и обработки данных.
- Документировал процессы разработки, писал техническую документацию, ставил цели в команде. Настроил CI в проектах с использованием black, pytest, mypy, isort.
- Собрал датасет пар текст-изображение, с помощью которого было улучшено качество генераций модели Kandinsky.
- Проектировал Data Lake для хранения мультимодальных датасетов.
- Проектировал и разрабатывал инференс моделей Kandinsky, Kandinsky Video. Деплоил эти модели в production.
- Оптимизировал модели на PyTorch. Ускорил инференс диффузионной модели на 15%.
- Разрабатывал процессы для воспроизводимой оценки text2image моделей.
- Обучил модель улучшения качества изображений Real-ESRGAN. Модель затем использовалась в пайплайне генерации изображений и в других внутренних проектах. [Код](#).
- Обучил модель детекции водяных знаков на изображениях. Модель имеет точность 93,44%, в то время как лучшая open-source модель имела точность 77,86%. Модель использовалась для фильтрации датасетов. [Код](#).
- Разработал модель для удаления водяных знаков с изображений. Модель используется для обработки обучающих данных.

ОБРАЗОВАНИЕ

МГТУ им. Баумана, кафедра ИУ9 (ПМИ)

Сентябрь 2020 – Настоящее время

Теоретическая информатика и компьютерные технологии. Неоконченное высшее.

Москва

Karpov Courses. ML system design course

Октябрь 2023 – Декабрь 2023

System Design

Москва

Прошёл курс по проектированию высоконагруженных систем. [Сертификат](#).

ТЕХНИЧЕСКИЕ НАВЫКИ

Языки программирования: Python

Обработка данных: Python, Pandas, Scipy, Numpy, Matplotlib

Машинное обучение: PyTorch, Transformers, Diffusers, CatBoost

Python: asyncio, multiprocessing, threading, boto3, fsspec, fastapi, pytest, mypy

Математика: Теория вероятностей и математическая статистика, Линейная алгебра, Математический анализ

Computer Science: Алгоритмы и структуры данных, Дискретная математика, Параллельное программирование

Прочее: Linux, Docker, AirFlow, S3, Bash, Git

Языки: English B2, Russian Native

ПРОЕКТЫ

SRGAN и Real-ESRGAN | Python, PyTorch

- Написал tutorial, в котором рассказал как работает SRGAN и как его обучать. [Код](#)

Style Transfer | Python, PyTorch

- Реализовал классический алгоритм переноса стиля с одного изображения на другое, но со множеством дополнительных настраиваемых параметров переноса. [Код](#)

Optimal Style Transfer | Python, PyTorch

- Адаптировал код статьи Optimal Style Transfer под PyTorch. [Код](#)

СТАТЬИ

MineralImage5k: A benchmark for zero-shot raw mineral visual recognition and description

- Подготавливал данные и разрабатывал бейзлайн для задачи zero-shot детекции и классификации минералов. [Статья](#), [Код](#)

Kandinsky: an Improved Text-to-Image Synthesis with Image Prior and Latent Diffusion

- Собираю и подготавливаю данные для обучения. Деплою модель в production. [Статья](#), [Код](#)

eco2AI: Carbon Emissions Tracking of Machine Learning Models as the First Step Towards Sustainable AI

- Тестирую эффективность применения fewbit функций активации для ускорения обучения модели Kandinsky. [Статья](#), [Код](#)

ru-CLIP

- Собираю и подготавливаю данные для обучения. [Статья](#), [Код](#)

ОЛИМПИАДЫ, ХАКАТОНЫ, ДОСТИЖЕНИЯ

AI Talent Hackathon 2023

- Занял призовое место в треке "Open Source". Был разработан бенчмарк для воспроизводимого подсчёта метрик моделей генерации изображений по тексту. [Решение](#).

Свидетельство о признании Президента Сбербанка

- Свидетельство о признании за достижение высоких показателей в работе и личный вклад в реализацию стратегии развития Сбера.

Олимпиада НТИ 2019-2020

- Победитель олимпиады НТИ по профилю "Искусственный интеллект". [Список победителей](#).