


ЕРҒАЗЫ ЕРСАЙЫН БЕКТҰРҒАНҰЛЫ

 [linkedin.com/in/ergazyersaiyn](https://www.linkedin.com/in/ergazyersaiyn)
 +7 702 629 50 88
 ergazyersaiyn@gmail.com
 github.com/yerssaintlaurent
 @yerssaintlaurent

ОБО МНЕ

Человек с практическим опытом в области ML и анализа данных. Работая над проектами в области CV и NLP, я стремлюсь к дальнейшему развитию своих навыков в динамичной и сложной среде.

ХАРД-СКИЛЛЫ

Core languages:

Python

Basic R

Basic MATLAB

Frameworks & tools:

NLP/CV

Transformers

SciPy

Matplotlib

DA skills:

Power BI/Tableau

P&S

A/B tests

SQL

PyTorch/TensorFlow

Git

Docker

basic FastAPI

Google Cloud/AWS

ОБРАЗОВАНИЕ

Bachelor's:

Computer Science
Satbayev University/Polytech
2019 - 2023

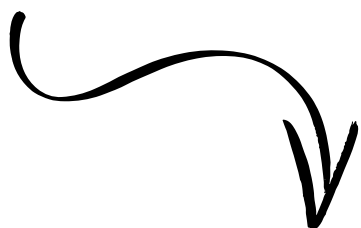
Master's:

Data Science
KBTU
2023-present

СОФТ-СКИЛЛЫ

- Широкий кругозор, умение учитывать мнение команды и открытость к фидбэку
- Умение объяснять технические материалы нетехническим способом
- Agile/Scrum
- Опыт волонтерской деятельности (особенно участие в организации крупных проектов)

ОПЫТ



ЯЗЫКИ

Казахский - C2
Русский - C2
Английский - B2-C1
Французский - B1

Data analysis:

● Анализ рынка аренды недвижимости

2024 | стек: Tableau, SQLite, Pandas, SciPy, Plotly и парсер

- Провёл комплексный анализ рынка аренды жилья в Казахстане с фокусом на города Алматы и Астана.
- Исследовал ценовые тренды, сегментировал предложения, визуализировал данные в Tableau и Plotly.
- Дополнительно построил real-time дашборд для бизнеса с основными метриками

● A/B-тестирование рекомендационной системы

2024 | стек: Pandas, SciPy, Seaborn, Matplotlib

- Провёл A/B-тест новой модели рекомендаций.
- Использовал статистические тесты (chi-square, тест Фишера, t-test и др.), визуализации, анализ показателей вовлечённости и др.

ML:

● Классификация возраста по изображениям лиц

2024 | стек: OpenCV, TensorFlow/Keras (VGG19), ONNX, Anvil, Matplotlib

- Построил и обучил сверточную нейросеть для классификации возрастных групп по изображениям.
- Реализована оптимизация обучения через EarlyStopping и ReduceLROnPlateau и оптимизаторы. Экспортировал модель в ONNX, деплой был через Anvil.

● Проект по распознаванию древней письменности

2024 (Отложен) | стек: TensorFlow, CNN/ReLU/VGG16, YOLO etc.

- Разработал систему распознавания не только отдельных символов, но и целых слов в рукописном тексте. Датасет полностью собран вручную и размечен через LabelStudio. Были эксперименты с искусственными повреждениями изображений (random noise, etc) для последующей реконструкции "потерянных" частей.

Проект может быть основой для цифровой реконструкции исторических документов.

● Интеллектуальный поиск недвижимости по текстовому запросу

2025 | стек: Transformers, Streamlit, Docker, FastAPI, Torch, RapidFuzz

- Разработал end-to-end систему для мультимодального поиска по объявлениям недвижимости с поддержкой неформальных запросов и с запросов с опечатками.
- Модель интерпретирует запросы (например, "снять трешку в Медеуском районе до 350 тыс тг с кондиционером в Алматы") и находит релевантные объекты. Если запрос содержит описание интерьера, то активируется поиск по изображениям. Подробнее в GitHub.
- В рамках проекта пробовал деплой через FastAPI и Docker, затем перешел на Streamlit.

Research-проекты:

● Сравнение LLM по решению математических задач

2024 (в процессе) | стек: OpenAI and others' API, Python, R (BERTScore, ROUGE), etc.

- Провели исследование способности крупных языковых моделей решать mwp на русском и английском языках (на основе собственного датасета).
- Провели глубокий анализ логической структуры решений, точность ответов и ошибки рассуждений. В процессе выдвинули несколько полезных гипотез. (any other info soon)

● CV в строительстве

2025 (в процессе) | стек: YOLO, Grounding DINO, Segment Anything, VLM etc.

- Исследуем ... в строительстве с помощью CV-моделей и дронов. (any other info soon)