

Даниил Матвиенко

Phone: +7 (953) 939-11-03 **Email:** matvienko.d2003@yandex.ru **GitHub:** <https://github.com/SaylesMand>

SUMMARY

Я ML-инженер с опытом работы с англоязычными научными статьями, генерации датасетов и разработки пайплайнов обучения. В основном работал с архитектурами глубоких нейронных сетей (CNN, LSTM, Transformer), но имеется опыт применения классических алгоритмов МО (линейные модели, SVM, ансамбли). Я разработал решения для детекции ИИ-сгенерированного текста, классификации текста и изображений, а также решение для распознавания рукописного текста (HTR).

SKILLS

Programming:

Python, SQL.

Complete ML & DL stack:

NumPy, Pandas, matplotlib, sklearn, PyTorch, OpenCV, Albumentations, Ultralytics, Transformers, Accelerate, PEFT, TensorBoard, Docker, Git, Linux, Maths.

ML fields:

- Classic ML (Price prediction, Personality Type Prediction, Image Classification, AI-Generated Text Detection).
- Computer Vision (Object Detection, Image Segmentation, Image Classification, Image-to-Text (HTR)).
- NLP (Text Classification, Text Generation, Sentence Similarity).

Languages:

Русский (родной), Английский (B1).

EDUCATION

САФУ имени М.В. Ломоносова | 2021-2025

Прикладная математика и информатика

Средний балл: 4.79 / 5.

KEY ACHIEVEMENTS

- Разработанная в рамках *ВКР* модель распознавания материалов диалектологических экспедиций будет интегрирована в информационную систему “Электронный корпус архангельских говоров” и использоваться студентами и преподавателями в учебных и научных целях.
- *Решение* по доставке лекарств роботом с использованием компьютерного зрения заняло второе место на Международных соревнованиях “Кубок РТК Высшая Лига: Доставка лекарств” [[Сертификат](#)].

COURSES & CERTIFICATES

Stepik

- Введение в Data Science и машинное обучение [[Сертификат](#)].
- Нейронные сети и компьютерное зрение [[Сертификат](#)].

Coursera

Supervised Machine Learning: Regression and Classification [[Сертификат](#)].

PROJECTS

RuDialect-HTR [[Github](#)]

Описание: разработка моделей для детекции и распознавания рукописного русского текста с диакритическими знаками.

Результат: дообучены модели YOLO11 (детекция текста) и TrOCR (распознавание) на доступных и сгенерированных датасетах.

Detect AI Generated Text [[Kaggle](#)]

Описание: бинарная классификация: является ли автор сообщения человеком или ИИ на основе диалогового контекста.

Результат: обучен XGBClassifier на эмбедингах, полученных с помощью SentenceTransformer.

Tourist RAG [[Kaggle](#)]

Описание: разработка туристической системы RAG с предварительным анализом данных (EDA).

Результат: использованы Chroma для семантического поиска и Phi-4-mini-instruct для генерации ответов.