

МОРОЗОВ ДЕНИС

Москва · гибрид | ✉ Kor4yz@yandex.ru | 📠 Телеграмм | 🌐 GitHub с работами | 📞 Whatsapp |

ОБРАЗОВАНИЕ

Высшая Школа Экономики

Россия, Москва

Бакалавриат, Прикладная математика, 3 курс

2023 – 2027

- **Ключевые дисциплины:** Численные методы анализа данных · Теория вероятностей и математическая статистика · Обработка и визуализация статистических данных · A/B Week (ШАД Яндекс) · Интерактивный тренажёр по SQL · DataLens: анализ и визуализация данных

DATA-ПРОДУКТ

MVP сервиса «Определение марки автомобиля по фото» | *Python, Docker, Selenium* 2023-2025

- Разработал end-to-end ML-продукт, распознающий марку автомобиля по изображению: от парсинга данных до онлайн-инференса.
- Создал распределённый парсер (Avito, Auto.ru, Яндекс.Картинки) с обходом капчи и фильтрацией изображений по марке.
- Собрал и разметил датасет (50 тыс. изображений, > 800 брендов), реализовал пайплайн предобработки (OpenCV + Albumentations).
- Обучил нейросеть (Best val accuracy 95 %), оптимизировал вес модели для inference.
- Развернул API на FastAPI + Docker и веб-интерфейс Streamlit для тестирования модели в реальном времени.
- Настроил автоматическое дообучение модели: еженедельный cron-пайплайн увеличивает датасет на 10 новых марок, повторно обучает и валидирует модель без потери предыдущих знаний.
- Собрал пользовательские метрики для последующего MVP-дашборда с аналитикой использования.

ПРОДУКТОВАЯ АНАЛИТИКА

Анализ пользовательских видеосессий VK Видео | *Python, Pandas, Numpy, SQL, DataLens* 2025

- Очистил и проанализировал 4.5 тыс. сессий, оставив 1 932 валидных (фильтрация аномалий > p99).
- Исследовал длительность просмотра, платформы и источники трафика; определил топ-видео по удержанию и просмотрам.
- Сформулировал гипотезы о поведении пользователей и метрики успеха; предложил гипотезы для дальнейших экспериментов и A/B-тестов.
- Собрал BI-дашборд в DataLens с ключевыми графиками и показателями вовлечённости.
- Результаты анализа легли в основу гипотез по повышению удержания и улучшению рекомендаций контента.

Анализ продаж и прогноз выручки e-commerce | *Python, SQL, DataLens, ClickHouse* 2024-2025

- Провёл анализ 1.2 млн транзакций, построил витрину данных и рассчитал ключевые продуктовые метрики — DAU, MAU, Retention, LTV, ARPU.
- Разработал модель прогнозирования выручки на 3 месяца вперёд (градиентный бустинг, MAE < 5 %), что позволило повысить точность планирования продаж и оптимизировать маркетинговые кампании и прогнозирование выручки.
- Визуализировал динамику метрик в интерактивном дашборде DataLens; подготовил рекомендации по росту повторных покупок и удержанию клиентов.
- BI-дашборд используется для еженедельного мониторинга продуктовых метрик и планирования промо-активностей.

T-Bank: Аналитика путешествий (отели и авиаперелёты) | *Python/Jupyter, SQL, DataLens* 2024

- Проанализировал транзакции travel-направления, выявил сезонность, пики продаж и ключевые клиентские сегменты (личные/рабочие поездки).
- Рассчитал метрики среднего чека и конверсии в покупку; выделил топ-направления и возрастные группы с наибольшей активностью.
- Собрал BI-дашборд в DataLens с визуализацией выручки, сезонных трендов и структуры заказов.
- Подготовил рекомендации: продвижение раннего бронирования (30+ дней), персонализированные офферы и cashback-кампании для ядра 25–34 лет.

ИССЛЕДОВАТЕЛЬСКИЕ ПРОЕКТЫ

Web-Japan: интерактивный сайт экскурсионных туров | *HTML, CSS, JavaScript, Leaflet.js* 2024

- Разработал адаптивный веб-сайт с поиском, фильтрацией, сортировкой и интерактивной картой (Leaflet + OSM).
- Реализовал корзину (localStorage), мультиязычность (RU/EN/JP), UI-оптимизацию под мобильные устройства.

Классификация данных с помощью нейросетей | *Python, TensorFlow, PyTorch, OpenCV* 2023-2024

- Изображения: обучил CNN для распознавания номеров домов (SVHN dataset, accuracy > 98 %) и классификации марок автомобилей (accuracy > 90 %).
- Тексты: реализовал бинарную (спам/не спам) и многоклассовую классификацию писем (accuracy > 97 %) на Keras с TF-IDF и токенизацией.
- Графы: применил Graph Neural Networks (GNN) для предсказания гамильтоновости графов (accuracy > 96 %).

ОЛИМПИАДЫ

Межрегиональная олимпиада школьников на базе ведомственных образовательных организаций — «Математика» 2022

Призер II степени

ТЕХНИЧЕСКИЕ НАВЫКИ

Инструменты анализа: Python, SQL (PostgreSQL, MySQL), ClickHouse, Excel

Аналитика и метрики: DAU/MAU, Retention (когорты), LTV/ARPU, Воронки, A/B-тесты, статистика

BI/визуализация: Power BI, Yandex DataLens

ML и DS: Scikit-learn, XGBoost/LightGBM; PyTorch, TensorFlow/Keras, PyTorch Geometric, NLP(NLTK), CV (OpenCV, Albumentations)

Инженерия и DevOps: Git, Docker, Jupyter/Colab, LaTeX, Excel

Дополнительно: базово Amplitude/Mixpanel

 Подробнее о проектах