

Сыймык Хамдамов

📍 Москва ✉ syimykhamdamov@gmail.com ☎ +7 926 959 57 92 🌐 semthedeв 📧 sk_khamdamov 📧 semthedeв

Технические навыки

- Языки программирования: Python, SQL, C#, C++, JavaScript
- Базы данных (СУБД): PostgreSQL, Greenplum, ClickHouse, Oracle, MySQL
- Инструменты & платформы: Apache Airflow, Docker, Git, Apache Spark, Power BI, Tableau, Apache Superset, Excel, Jupyter Notebook, Miro, Jira
- Фреймворки & библиотеки: pandas, NumPy, Matplotlib, Seaborn, SciPy, scikit-learn, CatBoost, XGBoost, LightGBM, PyTorch, TensorFlow, OpenCV, .NET, FastAPI
- Методологии & моделирование: CI/CD, BPMN 2.0, IDEF0/IDEF3, UML, ARIS, Unit-тестирование
- ML & Data Science: ML, бустинг, глубокое обучение, Computer Vision, Data Mining, NLP/LLM

Образование

- НИУ ВШЭ, Факультет компьютерных наук — Москва, Россия сентябрь 2025 — н.в.
 Магистратура по направлению «Прикладная математика и информатика», трек [Data Science](#) (ожидаемое окончание: август 2027)
 📄 (обучение на английском языке)
- Основные курсы: Methods of Artificial Intelligence in Decision Making; Mentor's Seminar; Ordered Sets in Data Analysis; Basics of Data Analysis; Машинное обучение; Research Seminar: "Number-Theoretic and Algebraic Methods in Data Analysis"; Modern Concepts of Dynamic Neural Networks; Natural Language Processing on Python.
- Российский Биотехнологический Университет (РОСБИОТЕХ) — Москва, Россия сентябрь 2021 — август 2025
 Бакалавриат по направлению «Информатика и вычислительная техника»
 Специализация: [искусственный интеллект для управления технологическими комплексами](#)
 📄 (красный диплом)
- Основные курсы: Введение в программирование (C++, Python); алгоритмы и структуры данных (C++, Python); математический анализ; линейная алгебра; дифференциальные уравнения; дискретная математика; математическая логика и теория алгоритмов; теория вероятностей и матстатистика; численные методы; архитектура компьютеров; операционные системы; вебпрограммирование; разработка вебприложений; основы теории управления; методы оптимизации и моделирование систем; корпоративные информационные системы; базы данных (MySQL); классическое ML; компьютерное зрение.

Опыт стажировок

- Стажёр Data Analyst / Data Scientist Москва, Россия
 Корпус Консалтинг декабрь 2024 — январь 2025
- Анализировал данные молочного завода и обрабатывал показания датчиков в Python (pandas, NumPy, Matplotlib, Seaborn) в Jupyter; формулировал гипотезы, находил аномалии и точки оптимизации расхода ингредиентов, что помогло снизить издержки. [Python, pandas, NumPy, Matplotlib, Seaborn, Jupyter, Miro]
 - Разработал модели машинного обучения (линейная регрессия, деревья решений, XGBoost, CatBoost, LightGBM) с помощью Scikit-learn, CatBoost, XGBoost, LightGBM и выбрал лучшую. [Scikitlearn, CatBoost, XGBoost, LightGBM]
 - Создал интерактивные отчеты в Superset и Power BI, визуализировав результаты анализа (деревья решений, графики важности признаков), что позволило задокументировать гипотезы и предоставить рекомендации по оптимизации производства и сокращению потерь ингредиентов. [Apache Superset, Power BI]
- Стажёр AR/VR Game Developer Москва, Россия январь 2022 — август 2022
 IT Pro Moscow
- Получил практический опыт программирования и закрепил принципы ООП; применял их при создании игр и AR/VR прототипов в Unity. [ООП, C#, Unity, JetBrains Rider]
 - Участвовал в VR проекте «Эндоскопия», моделирующем работу в операционной: пользователи отрабатывали виртуальные хирургические процедуры и эндоскопические навыки (логика взаимодействия, управление сценами, сценарии инструментов). [C#, Unity, JetBrains Rider, GitHub, GitLab, Jira]

Публикации

Разработка автоматизированной информационной системы для составления расписания в технических вузах на основе генетического алгоритма

ноябрь 2024

- III Всероссийская научная конференция «Искусственный интеллект в автоматизированных системах управления и обработки данных»

ИИАСУ'24 — «Искусственный интеллект в автоматизированных системах управления и обработки данных». Сборник трудов III Всероссийской научной конференции (Москва, 30 октября — 1 ноября 2024), т. 2, стр. 379–387 [↗](#)

Проекты

Автоматизированная система составления расписания для технических университетов

- Разработал модуль расписания на основе генетического алгоритма: анализ предметной области, математическое моделирование, предобработка данных, реализация.
- Стек: Python (pandas, NumPy, Matplotlib, Seaborn, SciPy, DEAP (библиотеки для генетического алгоритма)), Google Colab
- [Код](#) [↗](#)

А/В-тест для оптимизации лендингов в E-Commerce

- Провёл А/В-тест для оценки влияния новой лендинг-страницы на конверсию пользователей интернет-магазина.
- Выполнил статистический анализ (z-тест, p-value), подтвердив отсутствие значимых улучшений. Подготовил отчёт с рекомендацией оставить текущую версию страницы.
- Использовал Python (Pandas, Matplotlib), Scipy для проверки гипотез, оформил результаты в Jupyter Notebook.
- [Код](#) [↗](#)

Анализ влияния активности и сна на изменение веса (Exploratory + Predictive Analysis)

- Исследовал влияние уровня физической активности и качества сна на изменение веса участников. Выполнил очистку и объединение датасетов, визуализацию распределений, корреляционный и кластерный анализ.
- Построил регрессионную модель для прогнозирования итогового веса участников; определил наиболее значимые факторы (текущий вес и изменение веса), влияющие на финальный результат.
- Реализованы различные методы визуализации: тепловые карты, boxplot-графики, barplot'ы, line-графики. Применял кластеризацию KMeans и интерпретировал поведение кластеров на основе уровня стресса, продолжительности программы и профицита/дефицита калорий.
- Стек: Python (pandas, seaborn, matplotlib, scikit-learn), Jupyter Notebook
- [Код](#) [↗](#)

Пайплайн аналитики продаж (Greenplum–Airflow–Superset–ClickHouse)

- Спроектировал и реализовал учебный ETL-пайплайн для обработки данных о продажах: загрузка из CSV и PostgreSQL через gpfdist и PXF, партиционирование, агрегирование в витрину продаж с множеством метрик.
- Для оркестрации использовал Apache Airflow (DAG с задачами сбора временных таблиц и загрузкой витрины в Clickhouse), для визуализации дэшборда – Superset, для хранения – Greenplum и Clickhouse с настройкой распределения данных по нодам.
- Greenplum, Clickhouse, Apache Airflow, Apache Superset, Python.

Компьютерное зрение: классификация и детекция

- Классификация изображений: обучил модели для MNIST (MLP, PyTorch) и Fashion-MNIST (Dense/Dropout, Keras); подготовил полный пайплайн предобработки (нормализация, разбиение, DataLoader/генераторы), настроил обучение (Adam/SGD), валидацию и логирование метрик.
- Распознавание объектов на CIFAR-10: спроектировал и обучил CNN (Conv2D + MaxPooling, Dropout, Flatten, Dense), реализовал one-hot кодирование меток и автоматическую оценку качества на тестовом наборе.
- Детекция лиц: интегрировал MediaPipe Tasks (BlazeFace, .tflite) и OpenCV; реализовал преобразование нормализованных координат в пиксельные, отрисовку рамок и ключевых точек.
- Обработка изображений: собрал утилиты для resize/rotate/translate, конвертации цветовых пространств и визуализации (до/после); исследовал пространственную фильтрацию (медианный и лапласов фильтры) для шумоподавления и повышения резкости.
- Стек: Python, PyTorch, TensorFlow/Keras, OpenCV, MediaPipe, NumPy, Matplotlib
- [Код](#) [↗](#)

Онлайн курсы

- [Foundations of Statistics \(Stepik Certificate\)](#)
- [Курс SQL от SQL Academy](#)
- [Введение в теорию вероятностей](#) [SQL 50 \(LeetCode\)](#) [SQLзадачи \(SQLEx\)](#)
- [Программирование на Python](#) [Основы Python \(Hexlet\)](#) [CodeRun](#)
- [Тренировки. Забег по алгоритмам 27-место 40/40 задач](#)
- [Power BI](#)
- [Курсы edX](#) [Курсы Coursera](#)
- [Deep Learning School \(Школа глубокого обучения ФПМИ МФТИ\)](#)
- [YSDA NLP by Elena Voita](#)