

Вадим Куликов – Data Scientist

✉ akarinokadze@yandex.ru
☎ +7 (916) 921-60-96
📍 Дубна, Россия, GMT+3

🔗 [Akarinokadze](#)
📱 [kulikov-vadim](#)
📌 [Akari no kadze](#)

О СЕБЕ

Data Scientist и Team Lead R&D в сфере защитной полиграфии.

За время учёбы в Яндекс.Практикум [освоил](#) Python, SQL и основные инструменты для машинного обучения. Дополнительно прошёл курсы по Python: [67](#) и [512](#) и по основам статистики: [76](#) от Института биоинформатики. Изучаю основы A/B тестов.

Работаю на парт тайм в стартапе приложения для контроля финансов *Nimble app* в роли Data Scientist'a. Готовлю NLP модели к внедрению в приложение в качестве микросервиса. Занимаюсь MLOps.

Пять лет опыта руководителем группы разработчиков-технологов. С азартом берусь за выполнение задач с неочевидным решением. Довёл от идеи до промышленной реализации ряд решений, в том числе для международного рынка, на два из них получил патенты.

Обладаю хорошей математической и инженерной подготовкой. Не боюсь глубокой проработки задач – для проектирования и строительства собственного дома получил строительное образование.

В настоящий момент ищу full-time работу в роли Data Scientist / ML engineer. Удалённо или гибридный график с 1 днём в Москве или север МО.

НАВЫКИ

- **Языки программирования:** Python, SQL
- **Фреймворки и инструменты:** Kedro, FastAPI, Scikit-learn, PyTorch, TensorFlow, Keras, Keras_nlp, transformers, sentence_transformers, SPACY, numpy, matplotlib, seaborn, plotly, pandas, ydata_profiling, LightGBM, CatBoost, SHAP, PIL
- **СУБД:** PostgreSQL
- **Распределённые вычисления:** PySpark, MLlib
- **Другое:** Git, bash, docker
- **Языки:** Русский (носитель), Английский (B2), Испанский (B1)
- **Soft skills:** системное мышление, ответственность, деловая интуиция

РАБОЧИЙ ОПЫТ

Data Scientist – Nimble app (приложение для контроля финансов) *part-time* (20 ч/нед.) авг. 2022 – по Н.В.

- Проанализировал текущее решение на Java (OpenNLP + WEKA), проинтервьюировал CEO, собрал требования. Разработал каталог категорий трат.
- Организовал сбор и разметку данных для задачи классификации на платформе toloka.ai с ограничениями по бюджету.
- Провёл анализ эффективности различных моделей категоризации и кластеризации, что позволило поднять точность предсказаний до 80%.
- На базе фреймворка Kedro написал микросервис предсказания категории по текстовому запросу в docker-контейнере, для работы через API. Занимаюсь MLOps.

Team Lead R&D – КРИПТЕН (голографические средства защиты для спец-документов) *full-time* май 2018 – по Н.В.

- Вместе с командой разработали и внедрили в производство ряд новых средств защиты спец-документов.
- Зпатентовали несколько уникальных технологических решений в области спец-полиграфии.
- Ведём постоянный анализ технологических процессов и занимаемся их совершенствованием.
- Занимаюсь внедрением систем автоматического контроля качества выпускаемой продукции, сбора статистики и верификации подлинности, в том числе с применением машинного обучения.
- Работа с внешними и внутренними заказчиками. Калибровка ожиданий, формирование КД и ТД. Декомпозиция проектов для команды.

Ведущий инженер-технолог – КРИПТЕН

мар. 2017 – май 2018

Инженер-технолог – КРИПТЕН

фев. 2012 – мар. 2017

ОБРАЗОВАНИЕ

Курс профессиональной подготовки DataScience+
онлайн

Яндекс.Практикум,
2022 – 2023

**Бакалавр по специальности Промышленное и
Гражданское Строительство** дистанционно

Московский Государственный Строительный Институт,
2014 – 2016

Референт-переводчик с испанского языка
очно, с отличием

Российский университет дружбы народов,
2004 – 2011

Магистр по специальности Химия
очно, с отличием

Российский университет дружбы народов,
2004 – 2011

ДОП. ИНФОРМАЦИЯ, ДОСТИЖЕНИЯ

Хакатон LinkedIn – *Принял участие в кросс-функциональном [хакатоне](#) по тематическому анализу публикаций в социальной сети LinkedIn, выявлению постов на тему наставничества и менторства:* июнь 2023

- создал краулер для сбора и подготовил исходные данные;
- провёл тематическое моделирование публикаций;
- проанализировал полученные результаты (выделение наиболее популярных тем, ключевых слов).

Мастерская PySpark – *Принял участие в мастерской по анализу количества заказов такси в Чикаго на Hadoop кластере развёрнутом в docker-контейнерах на локальной машине (PySpark + MLlib):* сен. 2023

- развернул Hadoop кластер на локальной машине в docker-контейнерах;
- провёл EDA средствами PySpark;
- провёл feature-engineering для time-series данных
- с помощью MLlib провёл обучение нескольких моделей машинного обучения и предсказал количество заказов на следующий час для каждого района Чикаго.