

# ЯШИН НИКОЛАЙ

## ML engineer

@ kol.yashin@yandex.ru

8 999 163-60-57

KolyaYashin

Nikolay

VK Коля Яшин

## ABOUT ME

- Студент 4 курса МГТУ им. Баумана.

Специальность - "Прикладная математика"

- Проходил курсы по машинному обучению: Deep Learning School, Karpon Courses "Симулятор DS". **Сертификат**

- Призёр и победитель олимпиад по математике. **Дипломы**

- Участник хакатонов:

Хакатон от Data Secrets совместно с Dodo Pizza

Призёр хакатона СЗФО 2023

## SKILLS

Python

Numpy

Pandas

nltk

Sklearn

MLflow

PyTorch

PostgreSQL

Greenplum

Tableau

Git

Docker

Ubuntu

## HACKATHON

### TopBlog Web Site

#### GitHub

- **Описание:** Бизнес задачей хакатона стояло создание продукта, который автоматизировал бы процесс считывание информации со скриншотов и занесение в таблицы. Нам надо было считывать основные KPI (кол-во друзей, подписчиков и т.д.) со скриншота личной страницы и вносить соответственно в таблицы.
- **Тех. описание:** На первом шаге работы модели использовалась модель классификации площадки, с которой сделан скриншот. Для парсинга скриншота мы использовали предварительно обученную OCR модель. Для того, чтобы находить нужную информацию, наша модель сначала находила ключевые слова (например, друзья, подписчики, просмотры, ...) в зависимости от площадки, потом обрезала скриншот в небольшой области возле найденного слова, и лишь в конце искала нужные цифры. В качестве MVP я сделал небольшой сайт на Flask и задеплоил на VPS. Сохранение таблиц мы сделали с помощью простых csv файлов.
- **Стек:** CV, Flask, Tesseract, Docker, Linux

## EDUCATION

### Бакалавриат

#### МГТУ им. Баумана

Sept 2021 – June 2025

## EXPERIENCE

### Junior Бизнес аналитик

#### Tinkoff

Октябрь 2023 – Апрель 2024 (6 мес.)

- В качестве основной задачи, я разработал ml модель, способную прогнозировать количество новых кандидатов на вакансию. В модели есть разделения по каналу привлечения кандидата. Для прогноза модель использовала исторические данные разбитые по месяцам и такие показатели, как сезонность, тренд, бюджет на канал. Усреднённая точность модели оказалась 7 процентов.
- **Значимость:** Благодаря прогнозам на следующий месяц, можно лучше планировать общую ситуацию по кандидатам. Также модель помогала более оптимально распределять бюджет между каналами.
- **Стек:** Time Series, Greenplum, Tableau, Jira

## PET PROJECT

### AI Vocabulary Bot

#### Github Telegram

- **Описание:** Бот, предназначенный для изучения английского языка. Позволяет людям сохранять новые слова, которые они узнают в процессе изучения языка, получать о них информацию, а также эффективно запоминать в контексте.
- **Тех. описание:** Для хранения данных я поднимал докер контейнер с базой данных PostgreSQL. Для извлечения информации о словах я использовал парсер html страницы с сайта Cambridge Dictionary. В боте встроена простенькая рекомендательная система, которая рекомендует пользователям новые слова, основываясь на их интересах. Также в боте есть возможность пообщаться с чат ботом, который будет в сообщениях использовать слова из вашего списка. Это реализовано с помощью подключения к Gemini API и составления промптов.
- **Стек:** NLP, Aiogram, Docker, PostgreSQL, BeautifulSoup, Gemini API, Linux