



Пичужкина Ольга Валерьевна

Женщина, 26 лет, родилась 17 октября 1998

+7 (926) 0112113

vyhuholl@protonmail.com — предпочитаемый способ связи

LinkedIn: <https://www.linkedin.com/in/vyuholl/>

Другой сайт: <https://t.me/olgap981>

Проживает: Санкт-Петербург

Гражданство: Россия, есть разрешение на работу: Россия

Не готова к переезду, не готова к командировкам

Желаемая должность и зарплата

Data Scientist

Специализации:

— Программист, разработчик

Занятость: полная занятость, частичная занятость, проектная работа

График работы: полный день, гибкий график, удаленная работа

Желательное время в пути до работы: не имеет значения

Опыт работы — 3 года 2 месяца

Июль 2023 —
Январь 2025
1 год 7 месяцев

ООО "БитРу"

Москва

Информационные технологии, системная интеграция, интернет

• Разработка программного обеспечения

Data Scientist

На последнем месте работы у нас было два продукта:

- новостной сайт о финансах, где все новости парсились из открытых источников и переписывались через нейронки. Изображения к новостям тоже генерились через нейронки;
- браузерная игра, где механики были основаны на чтении новостей.

Для обоих сайтов я:

- написала парсеры новостных сайтов и телеграм-каналов (технологии — Selenium и BeautifulSoup), которые находят и скачивают новости для дальнейшей обработки;
- реализовала дальнейшую обработку новостей — классификация текстов, определение языка, перевод (если необходимо), детекция спама, парафраз, саммаризация, генерация заголовков, генерация изображений или стилизация уже имеющихся. Сначала у нас были свои модели для каждой задачи (свёрточная нейронная сеть для детекции спама, CatBoost для классификации текстов, трансформеры для всего остального), потом перешли на API OpenAI и Gemini, а свою модель оставили только для стилизации изображений.
- реализовала загрузку обработанных новостей в BigQuery, откуда они потом брались для размещения на сайте.
- подняла всё это в гугл облаке (изначально мы использовали Kubernetes, потом перешли на связку Google Cloud Run + Apache Airflow).

Технологии: Python (pandas, numPy), PyTorch, torchvision, LLM (OpenAI API, Gemini), Computer Vision (python-opencv), Google Cloud Platform (Cloud Run, BigQuery), Apache Airflow, Docker, Selenium, BeautifulSoup.

Ноябрь 2022 —
Февраль 2023
4 месяца

Crowdtesting

crowdtesting.ru/

Информационные технологии, системная интеграция, интернет

- Разработка программного обеспечения
- Системная интеграция, автоматизации технологических и бизнес-процессов предприятия, ИТ-консалтинг

Программист Python

Компания занимается тестированием программного обеспечения, качественными и количественными исследованиями цифровых сервисов и продуктов.

Я разработала сервис для автоматизации тестирования голосовых ассистентов и придумала метрики для оценки качества работы сервиса. В работе частично использовала машинное обучение.

Июль 2022 —
Сентябрь 2022
3 месяца

ООО КИБЕРБОРЕЯ

Москва, hh.ru/employer/4835563

Информационные технологии, системная интеграция, интернет

- Разработка программного обеспечения

Data Scientist

Компания занимается разработкой программного обеспечения новейшего поколения и систем помощи в принятии решений для корпоративных клиентов.

Я занималась решением задач машинного обучения (препроцессингом и аугментацией текстовых данных, дообучением моделей для генерации текста, настройкой MLFlow) для целей компании, а также анализом эффективности и улучшением работы (рефакторинг, документация, тестирование, улучшение производительности) кода имеющихся сервисов компании.

Май 2022 —
Июнь 2022
2 месяца

Индивидуальное предпринимательство / частная практика / фриланс

Программист Python

Заказы:

- Topic modelling японских текстов.

Выкачала корпус текстов с сайтов японских газет, разбила по временным промежуткам и для каждого временного промежутка выполнила тематическое моделирование (автоматически выделила "темы" – часто встречающиеся вместе наборы слов).

- Модуль для чат-бота, собирающего новости из нескольких телеграм-каналов, который определяет (использовалась модель doc2vec), являются ли 2 новости одинаковыми.

- Парсеры для различных торговых площадок.

- Частеречный тэггер для языка пушту (https://github.com/vyuholl/pashto_corpus).

Январь 2021 —
Октябрь 2021
10 месяцев

IT ForYou

www.itfyou.com/

Информационные технологии, системная интеграция, интернет

- Разработка программного обеспечения

Программист Python

Компания разрабатывает приложения, специализирующиеся на процессинге аудиосигналов и разборе речи ("speech intelligibility and audio signal processing").

Я занималась решением задач машинного обучения — построением предиктивных моделей:

- Работа со статьями по глубокому обучению в области обработки аудиосигналов и разбора речи;

- Сбор данных для обучения моделей: искала, откуда можно взять данные, нормализовывала данные из разных датасетов, составила инструкцию для сборщиков данных, написала скрипт, автоматизирующий валидацию данных.

- Препроцессинг данных (преобразование аудиосигналов в tif-файлы, оптимизировала разработанную компанией функцию конвертации аудио в изображения);
- Построение архитектуры моделей — классификаторов типов человеческой речи (детский крик, храп и т.п.);
- Обучение классификаторов;
- Финальные модели предсказывали тип речи с accuracy не ниже 0.8.

Образование

Высшее

2022 **Национальный исследовательский университет "Высшая школа экономики", Москва**
Факультет гуманитарных наук, Фундаментальная и компьютерная лингвистика

Повышение квалификации, курсы

2021 **НИУ ВШЭ**
Факультет компьютерных наук, Майнор «Интеллектуальный анализ данных»

2021 **Нетология-групп**
Нетология-групп, Курс Математика для Data Science

2021 **Нетология-групп**
Нетология-групп, Курс Deep Learning: глубокое обучение от перцептрона до новейших концепций

2021 **Нетология-групп**
Нетология-групп, Курс Django: создание функциональных веб-приложений

2020 **Школа программирования 21**
ПАО Сбербанк

2020 **Нетология-групп**
Нетология-групп, Курс Основы HTML и CSS

2019 **НИУ ВШЭ**
Coursera, Курс Introduction to Deep Learning

2018 **deeplearning.ai**
Coursera, Курс Neural Networks and Deep Learning

Электронные сертификаты

2019 Introduction to Deep Learning

2018 Neural Networks and Deep Learning

Навыки

Знание языков Русский — Родной
Английский — C1 — Продвинутый

Навыки

Python SQL Git Linux HTML CSS Английский язык ООП
Django Framework Bash Data Science Machine Learning Deep Learning
Natural Language Processing Golang PostgreSQL Data Analysis
Data Mining Keras TensorFlow PyTorch NLTK Алгоритмы Flask
FastAPI MongoDB Docker

Дополнительная информация

Обо мне

Мой гитхаб: <https://github.com/vyuholl>

Мой профиль на Kaggle: <https://www.kaggle.com/vyuholl>

Постоянно совершенствую свои навыки, стремлюсь развиваться, готова изучать новое.

Быстро обучаюсь, умею разбираться в чужом коде и переводить задачу с языка программирования на человеческий язык и обратно. Ищу интересные проекты.