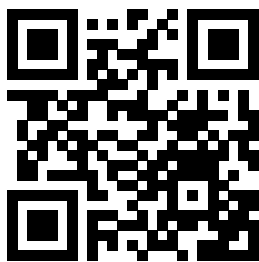
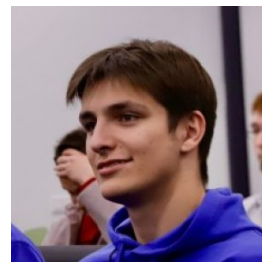


# Наиль Сираев — Data Scientist



Junior  
Россия, Казань  
120 000 - 180 000 ₽ (руб/мес)  
Полная занятость | Частичная занятость  
Удаленная работа  
Контакт: [geeklink.io/cv-113474](https://geeklink.io/cv-113474)



## Навыки

CV, Docker, FastAPI, HF, Langchain, NLP, Python, Pytorch, SQL.

## Опыт работы

### • Институт прикладной семиотики академии наук РТ

Май 2023 - Июль 2024 - ML engineer

- 1) Разработка модели распознавания татарской речи, обучение моделей семейства Whisper. Подготовка датасета, состоящего из 230 часов речи различной природы, фильтрация пустых записей. Тестирование моделей, сбор метри WER/CER. Перевод модели в ONNX формат, выведение в продакшен с помощью Docker, FastApi. Полученные метрики: WER = 0.1056, CER = 0.125
- 2) Разработка русско-татарского и татарско-русского переводчиков. Изучение различных RNN и Transformer моделей. Подготовка датасета, аугментация некоторых данных. Сбор метрик, выведение в продакшен с помощью Docker, FastApi. Полученные метрики: BLEU = 30.3
  - Обернул инференсы моделей в Docker
  - Перенес модель ASR в onnx формат

### • ГБУ Безопасность дорожного движения

Июнь 2022 - Май 2023 - ML engineer

- 1) Подбор модели детекции автомобильных номеров и определение скорости потока. Тестирование различных подходов. Улучшение трекинга автомобилей. Улучшение REid алгоритма - с mAP = 76 до mAP = 78.
- 2) Подбор методов для REid алгоритма, использующегося для нахождения людей, которые закрывают номера, в базе всех машин. Получение общей базы данных и выявление из нее автомобилей с закрытыми номерами. Тестирование несколько алгоритмов: REid, энкодер, обученный на автомобилях.
  - Улучшение mAP на таких автомобилях с 25 до 40
  - Ускорили работу алгоритма за счет распараллеливания работы методов с 0.8 секунд до 0.25 секунд

## Образование

### • Прикладная математика и информатика

КНИТУ-КАИ  
2021-2025

## Обо мне

Читаю статьи, участвую в хакатонах, люблю практически любые задачи (classic ML, CV, ASR, TTS, NLP, рекомендательные системы)

### КУРСЫ

- Deep Learning School Продвинутый поток  
Deep Learning School, Deep Learning
- Tinkoff Deep Learning  
Tinkoff, Deep Learning
- Основы статистики  
Stepik, Основы статистики
- Теория вероятностей  
Computer Science Center (CS центр)

### НАГРАДЫ

- Цифровой прорыв ИИ Межнар 2023- 9th CV.
- Цифровой прорыв ИИ СФО 2023- 3rd NLP.
- Tatar.By 2023 - 1st ASR TTS NLP.
- Letai Racing Kazan Digital Week 2023- 1st CV.
- Лидеры цифровой трансформации 2023-5th Time Series.