



# ФЕДОСЕЙ СОФЬЯ

## ОБРАЗОВАНИЕ

ТГУ 2021 - 2026

Факультет инновационных технологий ("Робототехника и Искусственный интеллект")

## КУРСЫ

ТГУ 2022 - 2023

- Аналитика данных

### Stepic

- Web парсинг на Python
- SQL
- Статистика
- Глубокое обучение (Deep Learning)

## ОПЫТ РАБОТЫ

- Стажировка от Школы Анализа больших данных ТГУ
- Прошла стажировку в Школе анализа больших данных ТГУ с фокусом на аналитике и обработке данных —
- ML-инженер: 1 год
- Дата-инженер: 1 год

## КОНТАКТЫ

8-923-416-00-25

s.fedosey@bk.ru

## НАВЫКИ

### Языки программирования:

- Python
- Практический опыт в веб-скрапинге
- Работа с библиотеками Python: requests, BeautifulSoup, Selenium

### Искусственный интеллект и обработка естественного языка (NLP):

- Применение эмбедингов multilingual-e5 для семантического поиска по документам и ответов на вопросы
- Разработка методов разбиения текста (chunking), ранжирования и оптимизации prompt-ов для поддержки обработки длинного контекста в LLM
- еализация пайплайнов retrieval-augmented generation (RAG) с использованием модели Mistral 7B для генерации ответов на основе документов

### Data Engineering и автоматизация:

- Создание и сопровождение масштабируемых ETL-процессов с использованием Apache NiFi
- Автоматизация пайплайнов и уведомлений с помощью n8n, включая интеграции с внешними API и мессенджерами (например, Telegram)

### Базы данных и скрипты:

- Управление хранилищами данных и системами запросов на PostgreSQL, включая отслеживание пользовательских действий и обратной связи
- Разработка продвинутых скриптов на Python и Groovy для обработки данных, генерации эмбедингов и оркестрации моделей
- Визуализация данных:
- Создание дашбордов с использованием SupersetBI