

Крохина Софья Петровна

Общая информация:

Дата рождения: 1 октября 2002 (22 года)

Гражданство: Российская Федерация

Город проживания: Воронеж

Контакты:

Номер телефона: +7 (980) 340-74-86

Эл. почта: sonyakrohina@yandex.ru

Telegram: @sopfiee (предпочитаемый способ связи)

GitHub: <https://github.com/Sopfiee>

Желаемая должность:

Data Scientist (ML Engineer)

Образование:

- Высшее, бакалавриат (2024)
Воронежский государственный университет
Факультет прикладной математики, информатики и механики
Прикладная математика и информатика
- Высшее, магистратура (2026)
Воронежский государственный университет
Факультет прикладной математики, информатики и механики
Математическое обеспечение и администрирование информационных систем

Повышение квалификации, курсы:

- Специалист в области искусственного интеллекта и машинного обучения (2025)
Национальный исследовательский Томский государственный университет
Цифровые кафедры
Полученные навыки: классический ML и анализ данных
- Курс теоретической и практической подготовки «Тренировки по машинному обучению 3.0» (2025)
Яндекс
Полученные навыки: Computer Vision (CV)
- Deep Learning School (1 и 2 семестры) (2025)
МФТИ
Полученные навыки: классический ML, CV, NLP

Навыки:

Языки программирования: **Python, SQL, C++**

Библиотеки: **Pandas, NumPy, Scikit-learn, XGBoost, CatBoost, LightGBM, PyTorch, Transformers, Matplotlib, Seaborn**

Инструменты: **Git, Docker, Linux**

Языки: Русский (родной), Английский (B1)

Опыт работы/стажировка/практика:

1. ООО “ВейвАксесс Сервис” июнь-июль 2025 (1 месяц)

- Изучала и применяла базовые методы предобработки данных, классификации и анализа временных рядов
- Проводила анализ данных и сравнивала качество ответов различных LLM моделей
- Работала в команде над проектом рекомендательных систем (устранение багов и тестирование)

2. NearCrowd февраль 2022 - январь 2023 (1 год)

- Описывала изображения
- Дополняла текстовую информацию для LLM моделей
- Писала код на C++ для решения алгоритмических задач

Обо мне:

- 2+ года практического опыта на Python (пет-проекты, курсы, ML-задачи).
- Уверенное знание классического ML (регрессия, классификация, кластеризация и т. д.), EDA и корреляционного анализа.
- Базовые навыки в Deep Learning – работала с NLP (transformer, RNN) и CV (CNN) задачами.
- Есть pet-проекты на GitHub, некоторые из них:

1. Smile Generator with Autoencoders – добавление лицам на изображениях улыбок с помощью кастомного автоэнкодера и латентных векторов изображений улыбающихся и грустных лиц. Использовался PyTorch, Pandas, Matplotlib ([GitHub](#)).
2. Stellar Classification – классификация небесных тел на основе их числовых/категориальных признаков. Использовались и сравнивались результаты различных моделей: RandomForest, XGBoost, LightGBM ([GitHub](#)).
3. Diamond Analysis - корреляционный и регрессионный анализ данных датасета о характеристиках бриллиантов. Построение регрессионной модели, предсказывающей цену за бриллиант в зависимости от значений его характеристик. Использовались scipy, pandas, statsmodels, numpy ([GitHub](#)).