

Мороз Роман Алексеевич

Мужчина, 20 лет, родился 24 мая 2005

+7 (930) 0871782 — предпочитаемый способ связи
morozdef24@gmail.com

Проживает: Белгород

Гражданство: Россия, есть разрешение на работу: Россия

Готов к переезду, не готов к командировкам

Желаемая должность и зарплата

Data Scientist / ML-engineer

Специализации:

- Аналитик
- Дата-сайентист

Занятость: полная занятость

График работы: полный день

Желательное время в пути до работы: не имеет значения

Опыт работы — 1 год 6 месяцев

Январь 2025 —
настоящее время
8 месяцев

Collaba

Data Scientist

Сгенерировал синтетических пользователей с помощью LLM-моделей и анализировал данные в Pandas, смоделировал реальное поведение аудитории и создал обучающий датасет.

Построил алгоритмы на основе графа знаний для разных типов пользователей, улучшил релевантность выдачи и персонализацию.

Разработал алгоритмы и модели рекомендательной системы с использованием библиотек Surprise, LightFM, а также PyTorch, CatBoost, scikit-learn, повысил качество рекомендаций (CTR +11% на тестовой группе).

Настроил инфраструктуру для ML-экспериментов: развернул MLflow, организовал трекинг моделей и управление экспериментами.

Разработал и интегрировал сервис рекомендаций на FastAPI с использованием Kafka, обеспечил связь между backend и сервисом рекомендаций.

Снизил latency выдачи рекомендаций на ~20%, ускорил отклик системы в реальном времени и повысил удобство пользователей.

Май 2024 —
настоящее время
1 год 4 месяца

Школа программирования 21

Белгород

Студент

Data Science Bootcamp

Очищал, агрегировал и визуализировал данные с помощью NumPy, Pandas, Matplotlib, Seaborn, Plotly.

Проводил анализ пропущенных значений и выбросов, выполнял EDA для выявления

закономерностей и подготовки признаков.

Разрабатывал модели классификации и регрессии с использованием scikit-learn, XGBoost, CatBoost.

Оценивал модели по ключевым метрикам: Accuracy, Precision, Recall, F1-score, ROC-AUC, MAE, RMSE, R².

Подбирал гиперпараметры с помощью GridSearchCV и Optuna.

SQL Bootcamp (PostgreSQL)

Писал и оптимизировал сложные SQL-запросы с использованием JOIN, GROUP BY, HAVING, оконных функций.

Оптимизировал производительность с помощью индексов и анализа плана выполнения (EXPLAIN, EXPLAIN ANALYZE).

Проектировал схемы баз данных и настраивал связи, обеспечивал целостность данных.

Работал с транзакциями и триггерами, автоматизировал обработку данных и логику бизнес-процессов.

Март 2024 —
настоящее время
1 год 6 месяцев

Индивидуальное предпринимательство / частная практика / фриланс

Python Developer / Telegram Bot Developer

Технологии: Python, aiogram 3, PostgreSQL, SQLite, MongoDB, REST API, Docker, Transformers, OpenAI API

Разрабатывал Telegram-ботов с нуля на основе aiogram 3 — от простых чат-ботов до комплексных многофункциональных систем.

Работал с базами данных:

- PostgreSQL — для продакшн-решений с хранением пользователей, логов и данных;
- SQLite — для лёгких прототипов и оффлайн-приложений;
- MongoDB — для хранения гибких, вложенных и динамических структур данных.

Контейнеризовал проекты с использованием Docker, настраивал docker-compose для локальной и серверной разработки.

Разрабатывал AI-агентов:

- Агент на базе Transformers и OpenAI API для анализа тональности телефонных звонков и выдачи рекомендаций по стилю общения;
- Telegram-бот для анализа психологических проблем и оказания ментальной помощи с применением NLP-моделей.

Писал техническую и пользовательскую документацию для разработанных решений.

Образование

Бакалавр

2027

Белгородский государственный технологический университет имени В.Г. Шухова, Белгород

Повышение квалификации, курсы

2025	Анализ больших данных БГТУ им. В. Г. Шухова Цифровая Кафедра, Анализ данных, машинное обучение, BigData
2025	Горютиновый golang WB Техношкола, Go developer

Электронные сертификаты

2024	Intermediate Machine Learning
------	-------------------------------

Навыки

Знание языков	Русский — Родной
	Английский — B2 — Средне-продвинутый
Навыки	Python Go C/C++ SQL Numpy pandas Scikit-learn PyTorch
	Matplotlib XGBoost CatBoost Deep Learning Docker
	Теория вероятностей Математическая статистика CI/CD MLflow
	PySpark

Дополнительная информация

Обо мне	Я — увлечённый Data Scientist с глубокими знаниями в области анализа данных, машинного обучения и разработки интеллектуальных решений.
	https://github.com/HoshiBatista — репозитории с моими проектами.
	https://www.kaggle.com/crissyro — участие в соревнованиях, публикации ноутбуков и аналитические работы.
	Стремлюсь применять знания для развития бизнеса, решая сложные задачи с помощью данных.