

НОВИЧКОВ СЕРГЕЙ

DATA SCIENTIST

+7(910)-437-29-18 | novichkovsd@ya.ru | Telegram | VK | LinkedIn | GitHub

ОБО МНЕ

Data Scientist с опытом работы более двух лет. Мой опыт включает руководство командой разработки из 5 человек, развертывание рекомендательных систем и NLP-моделей, а также разработку и внедрение ETL-процессов с использованием передовых технологий.

НАВЫКИ

Programming Languages: Python, SQL (MySQL, PostgreSQL), R, Bash, NoSQL (HBase)

Developer Tools: Git, Docker, PyCharm, Jupyter Notebook, Google Colab

Libraries: Scikit-learn, Statsmodels, Transformers, Pandas, NumPy, Matplotlib, Plotly, Seaborn, Aiogram3, Asyncio, XGBoost, CatBoost, LightGBM, NLTK, SpaCy, Requests

Other Systems : Hadoop (HDFS, YARN), ETL (Sqoop, Flume, NiFi, Kafka)

Languages : Russian (Native), English (Intermediate)

ОПЫТ РАБОТЫ

Data Scientist

10.2023 – настоящее время

Startup, SalaryScan - ML-платформа для расчета заработной платы по резюме

Москва, Россия

- Разработал систему парсинга резюме с открытых источников, что ускорило сбор датасета в 6 раз
- Использовал NER для выявления навыков, что позволило размечать данные автоматически, а не вручную
- Разработал и обучил ML модели для прогноза вилки з/п по резюме

Стек: Python, Sklearn, MySQL, Linux, Bash, Docker, Requests, Psycopg2, SpaCy, Git

Team Lead

03.2023 – 10.2023

Startup, b2c marketplace в telegram

Москва, Россия

- Руководил командой из 5 человек, распределял задачи в kanban, проводил код-ревью
- Разработал рекомендательную систему, что повысило конверсию продаж на 37%
- Спроектировал и разработал SQL базу данных, что сделало выгрузку данных быстрее на 60%
- Задеплоил проект с помощью docker на удаленный сервер

Стек: Python, PostgreSQL, Linux, Bash, Docker, Aiogram 3, Psycopg2, Asyncio, Git

Data Scientist

09.2021 – настоящее время

РТУ МИРЭА

Москва, Россия

- Выступил на 4 митапах по Data Science (50+ слушателей)
- Вывел рекомендательную систему в production с использованием Docker
- Разработал две SQL базы данных (MySQL/PostgreSQL), что повысило скорость работы с данными в 3 раза
- Внедрил систему машинного обучения для классификации дорожных знаков, что позволило предоставить готовое решение для навигационной CV системы автомобиля. Ассигнатуру нейронной сети: 0.98.
- Выстроил ETL процессы с помощью NiFi, Kafka, Sqoop и Flume, что ускорило выгрузку данных на 70%. Внедрил использование Hadoop (HDFS, YARN) и NoSQL БД HBase

Стек: Python, SQL, Linux, Bash, Docker, Git, Hadoop, Sklearn, ETL, RecSys, XGBoost, CatBoost, Keras

ПРОЕКТЫ

LesMeh

09.2022 – 03.2023

- Разработал асинхронного телеграм-бота и парсер Excel-расписания, который ежедневно обновляет расписание. Решение снизило среднее время получения расписания в 5 раз
- 860+ человек используют бота, 300+ пользуются им ежедневно
- Собрал датасет и обучил на нем RuGPT-3 модель LLM, собрав цифрового двойника студента
- Задеплоил проект в Docker-контейнере на удаленном сервере

Стек: Python, Aiogram2, Asyncio, MySQL, Docker, Bash, Miro, BeautifulSoup4, Transformers, Linux

ОБРАЗОВАНИЕ

РТУ МИРЭА

Москва, Россия

Бакалавриат, прикладная математика, анализ данных, 4.8/5.0

2021 – настоящий момент