

Галина Тузникова — Data Scientist



- Junior
- Россия, Санкт-Петербург
- Полная занятость
- Удаленная работа
- Контакт: geeklink.io/cv-83792



Навыки

AIogram, Bash, Catboost, Docker, Git, Google Colab, lightgbm, Linux, Matplotlib, Numpy, Pandas, PostgreSQL, PySpark, Python, Pytorch, Scikit-learn, scipy, Seaborn, Streamlit, XGBoost, Yandex Cloud.

Опыт работы

• ООО "Elbrus bootcamp"

02.2023/06.2023 - Data Scientist

В составе команды участвовала в проектах по анализу данных: поиск и предобработка данных, проверка статистических гипотез, проведение A/B тестов, анализ временных рядов. Я создавала модели машинного обучения, а также занималась построением и обучением нейросетей для задач классификации и регрессии. Также я разрабатывала streamlit-приложения и телеграм-ботов.

Стэк: Python, Pandas, NumPy, SciPy, Scikit-learn, CatBoost, XGBoost, LightGBM, PyTorch, Matplotlib, Seaborn, PySpark, Yandex Cloud, GoogleColab, Streamlit, Aiogram, Git, Bash, Docker, PostgreSQL

Пример проекта.

AristocratBot - telegram-bot, различающий более 30 классов столовой утвари по фотографии.

Ссылка на репозиторий: <https://github.com/GalkaMT/AristokratBot>

• АО "НБИ"

11.2021/01.2023 - Дата-аналитик

В составе команды участвовала во внедрении, тиражировании и расширении функционала системы оптимизации режимов ТЭЦ при заданных рыночных условиях и технологических ограничениях. Критерием оптимизации являлась минимизация расходов энергетических ресурсов. Система позволила увеличить маржинальную прибыль ТЭЦ до 5%.

Задачи:

-Анализ массивов данных о режимах работы станций (более 100 часовых показателей) с использованием статистических методов (оценка качества, структуры, заполнение пропусков, работа с базами данных и технических показателей).

-Использование инструментов визуализации для контроля и интерпретации данных.

-Подготовка презентаций, участие в совещаниях, семинарах.

Стэк:

MS Excel, Python, Pandas, PostgreSQL, Matplotlib, Seaborn, Power BI

• ПАО "ТГК-1"

03.2019/11.2021 - Аналитик службы оптимизации режимов

В составе команды провела ретроспективный анализ технико-экономических показателей и оценку

маржинальной доходности всех ТЭЦ филиала за 20 лет. Результаты исследования используются при формировании инвестиционной и ремонтной программ.

Задачи:

- Анализ технических и технико-экономических показателей работы ТЭЦ с использованием статистических методов и алгоритмов.
- Выявление и анализ ключевых факторов, оценка влияния сезонности на прибыльность и технические показатели ТЭЦ .
- Создание макетов и автоматизация расчетов для упрощения и ускорения анализа данных, связанных с выбором состава оборудования, формирование отчетов, подготовка презентаций и участие в совещаниях.
- Менторство и обучение студентов и стажеров работе с данными, включая обработку "сырых" данных и применение аналитических методов.

Стэк:

MS Excel, Access, Power Query, PowerPoint, Power BI Python, Pandas, Matplotlib

Образование

- **Магистр**

Санкт-Петербургский политехнический университет Петра Великого
2007/2009

Энергомашиностроительный факультет, Теплоэнергетика

Обо мне

- Data Scientist с опытом работы в обработке и анализе данных более 2х лет.
- Имею высшее техническое образование, опыт работы в энергетике, в том числе- на производстве, с "сырыми" данными разного качества и из разных источников.
- Имею практический опыт использования методов машинного обучения, прогнозного моделирования и статистического анализа.
- Обладаю навыками сбора, агрегации и подготовки исходных данных для обучения моделей, интерпретации и визуализации результатов.
- Имею опыт выступлений перед большой аудиторией (чтение лекций, участие в семинарах, совещаниях).
- Люблю работать с табличными данными и различными визуализациями.
- Постоянно повышаю свою профессиональную квалификацию с помощью образовательных курсов, читаю статьи по темам, связанным с дата-аналитикой, принимаю участие в хакатонах, соревнованиях на kaggle.
- Увлекаюсь SUP-бордингом и акварельным скетчингом.

Технологический стек: Python, Pandas, NumPy, SciPy, Scikit-learn, CatBoost, XGBoost, LightGBM, PyTorch, Matplotlib, Seaborn, PySpark, Yandex Cloud, Google Colab, Streamlit, Aiogram, Git, Bash, Docker, Linux, PostgreSQL